

Understanding the Influence of Adsorption-Mediated Processes on Antibody Aggregation in Bioprocessing

A thesis submitted to University College London for the degree of
Doctor of Philosophy in Biochemical Engineering

Alice Rachel Mazzer

April 2015

Declaration

I, *Alice Rachel Mazzer*, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. The thesis does not exceed the prescribed word limit, and none of the work presented in this thesis has been submitted for a qualification at any other institution.

Signature.......... Date.....15 / 04 / 16.....

Project Supervisors:

Dr Daniel G. Bracewell

Prof Paul A. Dalby

Dr John O'Hara

Acknowledgements

I would like to acknowledge my supervisor, Dan Bracewell, without whom I may not have chosen to pursue this PhD. Thank you Dan for helping to make this PhD rather exciting, at times. Thank you for your advice and your gentle approach.

John, Jenny, Xavier and the UCB team, thank you for your time and enthusiasm.

Tatiana, Chris, Paul and Hirsh, thank you for assisting me in my pursuits at the NCNR, and for showing me how it rolls across the pond.

Luke and Filip, thank you for all your help with the neutrons; I couldn't have done it without you.

John and Jeremy, thank you for your very insightful inputs; if such a thing can exist, you really were charming examiners.

UCL crowd, what a great bunch you are. Doing a PhD is a lot more fun when surrounded by such a lively, lovely, diverse group of people. Anand, you were there with me to the end, thanks for being awesome.

Alex, you have always been there for me. We have been on quite a journey over the last four years. Thank you for your endless love and devotion.

Marianne, Kirsty and Ollie, you are my best friends. Without you, life would be lacking.

Ma, Pa and Art, you make all this worthwhile.

Abstract

Affinity chromatography is an indispensable method of protein purification used during the manufacture of many therapeutic antibodies. The protein A ligand is a popular choice for selective purification of immunoglobulin G (IgG) molecules. Aggregated product is often found in protein A elution pools; this is generally attributed to the effect of low pH (elution) on protein structure. However, there is evidence that other facets of the chromatography process influence aggregation phenomena. Physical and chemical transitions, such as concentration of protein on the adsorbent surface and change in buffer composition, may challenge the structural integrity of proteins, increasing their propensity for aggregation.

The influence of elution from protein A on the aggregation rate of an IgG4 during a subsequent low pH hold was investigated. IgG4 was incubated in elution buffer after protein A chromatography and the monomer concentration in neutralised samples was quantified by size exclusion chromatography. Rate constants for monomer decay over time were determined by fitting exponential decay functions to the data. Similar low pH experiments were implemented in the absence of a chromatography step. The IgG4 demonstrated highly pH-dependent and apparently concentration-independent aggregation behaviour. The findings suggested that aggregation was governed predominantly by a pH-dependent unfolding/ re-folding equilibrium. Elution from protein A was found to increase aggregation rates by half an order of magnitude, while other aspects of the aggregation kinetics appeared un-affected.

In order to advance understanding of how adsorption processes might impact antibody stability, neutron reflectivity was used to characterise the structure of adsorbed IgG on model surfaces. In the first model system IgG was adsorbed directly to silica and demonstrated a side-on orientation with high surface contact. A maximum dimension of 60Å in the surface normal direction and high density surface coverage were observed under acidic pH conditions. In the second model system protein A was attached to a silica surface to produce a configuration representative of a glass chromatography resin. Interfacial structure was probed during

sequential stages from ligand attachment, through IgG binding and elution. Adsorbed IgG structures extended up to 230Å away from the surface and showed dependence on surface blocking strategies. The data was suggestive of two IgG molecules bound to protein A with a somewhat skewed orientation and close proximity to the silica surface. The findings provide insight into the orientation of adsorbed antibody structures under conditions encountered during chromatographic separations.

The outcomes of this work provide a broad scope for future investigations. Based on preliminary work, using neutron measurement techniques to monitor aggregate formation inside glass chromatography resins is suggested as an interesting direction.

List of Figures

Figure 3.1 Common protein secondary structures: right-handed alpha-helix and anti-parallel beta-sheet	22
Figure 3.2 Schematic of antibody structure.	25
Figure 3.3 Non-native aggregation and native polymerisation pathways shown schematically	29
Figure 3.4 Schematic depictions of protein A ligands	36
Figure 3.5 Crystal structure of the C domain of staphylococcal protein A, <i>SpA_C</i> , in complex with the Fc monomer fragment of human IgG. RCSB Protein Data Bank ID: 4WWI (biological assembly 2).	38
Figure 3.6 The stages of protein A affinity chromatography and typical buffer conditions for each stage.	43
Figure 3.7 An SEC chromatogram. This chromatogram was generated using a Tosoh Tsk-Gel G4000SWXL column and an aggregated IgG	48
Figure 3.8 DSC thermograms identifying phase transitions for different domains of various IgG molecules	53
Figure 5.1 Schematic of experimental design where incubation of IgG4 at low pH follows protein A chromatography	74
Figure 5.2 SEC chromatograms for IgG4 incubated under a range of low pH conditions.	76
Figure 5.3 Native-PAGE of untreated IgG4 and HMW species formed in solution or after elution from protein A at low pH.....	78
Figure 5.4 IgG monomer loss over time in solution.....	80
Figure 5.5 Semi-log plot of monomer decay rate, R_0 , against pH in solution at three different IgG concentrations.	81
Figure 5.6 The effect of maltose concentration on the rate of IgG4 monomer loss.....	82
Figure 5.7 Rates of IgG4 monomer loss under different pH conditions in solution only and after elution from a protein A chromatography column	84

Figure 5.8 Y_0 (plateau point parameter) from exponential decay curve fits plotted against pH in solution, and after protein A chromatography.	87
Figure 5.9 Semi-log plot of monomer decay rate, R_0 , against pH for solution-only experiments and experiments including a protein A chromatography step.....	89
Figure 6.1 Schematic representation of a neutron reflectivity instrument, and a photograph of the NG7 reflectivity rig at the NIST Center for Neutron Research.	97
Figure 6.2 Schematic representation of the flow cell used for reflectivity experiments.	99
Figure 6.3 Momentum transfer in neutron reflectivity, and constructive and destructive interference patterns.....	101
Figure 6.4 Simulated reflectivity data for a silicon-D ₂ O interface with no material present at the solid-liquid interface, and with a 55Å homogenous film of material	103
Figure 6.5 Example of output from bootstrap error analysis using RasCal software.....	111
Figure 6.6 Monomer reduction over time after dialysis into 30mM sodium citrate.	113
Figure 6.7 Reflectivity against momentum transfer vector, Q , for sequential IgG adsorption and wash steps at pH 4.1, and corresponding profiles of scattering length density, SLD, against distance from the silicon surface.....	116
Figure 6.8 The data in <i>Figure 6.7</i> was re-fitted giving improved chi-squared values and good physical feasibility; an additional layer located adjacent to the SiO ₂ layer was added to the model.	123
Figure 6.9 Graphical output of bootstrap error analysis for selected parameters from the re-fitted SLD profile of IgG adsorbed at pH4.1, seen in <i>Figure 6.8</i>	125
Figure 6.10 Reflectivity against momentum transfer vector, Q , for sequential IgG adsorption and wash steps at pH 6.2 and pH 3.7, and corresponding profiles of SLD against distance from the silicon surface	129
Figure 7.1 Components used to immobilise protein A on the silicon wafer and an NMR structure of the b domain of protein A.....	137
Figure 7.2 Amino acid sequence of recombinant staphylococcal protein A (rSPA) from Repligen.	140

Figure 7.3 Simplified depiction of an ATR-FTIR sample and infrared beam configuration, with the sample cell in flow-through mode.....	143
Figure 7.4 Flow chart and cartoon diagrams to illustrate a procedure used to assess rSPA non-specific binding to silica, rSPA cross-linking, IgG adsorption and IgG non-specific binding .	145
Figure 7.5 FTIR spectra for sequential stages of silica surface modification	146
Figure 7.6 FTIR spectra showing stages of IgG adsorption to protein A, elution from protein A, and IgG non-specific binding and subsequent elution	148
Figure 7.7 Annotated FTIR spectra for immobilised rSPA and adsorbed IgG showing amide peak areas and sub-peaks identified for rSPA and IgG4.....	153
Figure 7.8 Reflectivity data for attachment of rSPA to an amino-silylated silica surface, addition of the blocking agent BSA and corresponding profiles of scattering length density against distance from the silicon surface.	158
Figure 7.9 Reflectivity against Q at three solution phase contrasts for IgG4 adsorbed at pH 6.7 to rSPA-modified silica with BSA blocking.....	163
Figure 7.10 Profile of scattering length density against distance for IgG4 adsorbed at pH 6.7 to protein A-modified silica with BSA blocking.	164
Figure 7.11 Reflectivity data for attachment of rSPA to an amino-silylated silica surface, and addition of the blocking agent PEG ₆₀₀₀ , and corresponding profiles of scattering length density against distance from the silicon surface.	168
Figure 7.12 Reflectivity data for IgG4 adsorbed at pH 6.7 to rSPA-modified silica with PEG ₆₀₀₀ blocking, and corresponding profiles of scattering length density (\AA^{-2}) against distance from the silicon surface.....	169
Figure 7.13 Reflectivity data for IgG4 adsorbed at pH 6.7 to rSPA-modified silica with PEG ₆₀₀₀ blocking in <i>Cell C</i> , and corresponding profiles of scattering length density against distance from the silicon surface.....	171
Figure 7.14 SLD-Distance profiles for IgG adsorbed to protein A with BSA surface blocking and PEG ₆₀₀₀ surface blocking in <i>Cell B</i> and <i>Cell C</i>	172

List of Tables

Table 3.1 Selected resin characteristics, subsequent column conditions and predicted effect regarding aggregation phenomena.	42
Table 3.2 Selected recombinant and natural protein A media and their basic compositions. ...	42
Table 3.3 Further methods for characterising samples and proteins.	56
Table 6.1 Scattering lengths of elements commonly found in biological molecules	105
Table 6.2 Molecular volumes, sum of scattering lengths ($\sum b$) and scattering length densities for a range of amino acids and for a dipeptide in H ₂ O and in D ₂ O where the amide and hydroxyl hydrogens are exchanged to deuterium.....	106
Table 6.3 Scattering length density for IgG4- κ in H ₂ O and in D ₂ O at a range of H-D exchange levels	106
Table 6.4 Theoretical SLD values for IgG at different H-D exchange levels; scattering length densities of the mobile phase and inner and outer protein layers from fitted models for IgG adsorption data, and estimated maximum protein volume fractions for each layer.....	119
Table 7.1 Chi-squared values for fitted models for different sample configurations across the three SLFCs (A, B and C).....	156
Table 7.2 Volume fraction estimates for layers of different sample configurations in <i>Cell A</i> .	165

Nomenclature and Abbreviations

Å	angstrom
APTS	3-aminoprpyltriethoxysilane
ATR	attenuated total reflectance
BSA	bovine serum albumin
CBS	consensus binding site
CD	circular dichroism
CDR	complementarity determining region
C _H	constant region of an immunoglobulin heavy chain
C _L	constant region of an immunoglobulin light chain
D ₂ O	deuterium oxide
DLS	dynamic light scattering
EDTA	ethylenediaminetetraacetic acid
DSC	differential scanning calorimetry
Fab	fragment antigen-binding
Fc	fragment crystallisable
FDA	Food and Drug Administration (Federal agency)
FTIR	Fourier transform infrared spectroscopy
HCP	host cell protein
HMW	high molecular weight
HPLC	high performance liquid chromatography
IgG	immunoglobulin G
Mab	monoclonal antibody
NCNR	NIST Center for Neutron Research
NIST	National Institute of Standards and Technology
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PEG	polyethylene glycol
pI	isoelectric point
Q	momentum transfer vector
Q _z	momentum transfer vector
RAL	Rutherford Appleton Laboratory
R _g	radius of gyration
R _h	hydrodynamic radius

R _M	reactive monomer
rSPA	recombinant staphylococcal protein A
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
SLD	scattering length density
SLFC	solid-liquid flow cell
SMW	silicon matched water
SpA _C	C domain of staphylococcal protein A
SPDP	succinimidyl 6-(3'-(2-pyridyldithio)propionamido) hexanoate
T _m	thermal denaturation midpoint temperature
V _H	variable region of an immunoglobulin heavy chain
V _L	variable region of an immunoglobulin light chain

Table of Contents

1 Thesis Description	17
2 Aims and Objectives.....	19
2.1 Context and Motivation	19
2.2 Objectives	20
3 Literature Review.....	21
3.1 Antibodies	21
3.1.1 Protein structure	21
3.1.2 Antibody structure.....	24
3.2 Aggregation.....	26
3.2.1 Aggregate classification.....	26
3.2.2 Aggregation kinetics	27
3.2.3 Intrinsic factors	30
3.2.4 Extrinsic factors	32
3.3 Chromatography	34
3.3.1 Chromatography in bioprocessing	34
3.3.2 Protein A	35
3.3.2.1 Protein A as an affinity ligand	35
3.3.2.2 Structural characteristics of protein A and Fc.....	35
3.3.2.3 Fc-protein A interaction: potential to induce IgG aggregation	39
3.3.2.4 Protein A supports – influence on adsorption.....	40
3.3.3 Stages of affinity chromatography.....	43
3.3.3.1 Product loading.....	43

3.3.3.2	Column wash	44
3.3.3.3	Product elution	44
3.4	Analytical Methods	45
3.4.1	Analytical methods and aggregation	45
3.4.2	Size exclusion chromatography (SEC)	46
3.4.2.1	How SEC works	46
3.4.2.2	Sample preparation and operating ranges	48
3.4.2.3	Preventing column-protein interactions	49
3.4.3	Dynamic light scattering (DLS)	51
3.4.4	Differential scanning calorimetry	52
3.4.5	Circular dichroism	54
3.4.6	Further methods	56
3.5	Conclusion	57

4 Materials and Methods58

4.1	Materials	58
4.2	Equipment	58
4.3	Methods	59
4.3.1	IgG4 aggregation rate in solution	59
4.3.2	IgG4 aggregation rate after chromatography	60
4.3.3	Size exclusion chromatography	61
4.3.4	Aggregation rates data analysis	61
4.3.5	Neutron reflectivity experiments at the NIST Center for Neutron Research	61
4.3.5.1	Sample preparation	61
4.3.5.2	Reflectivity cell	62

4.3.5.3	Reflectivity experiments	62
4.3.5.4	Data analysis	64
4.3.6	Experiments at the ISIS Neutron and Muon Source	64
4.3.6.1	Sample preparation: protein A	64
4.3.6.2	Sample preparation: BSA and PEG	65
4.3.6.3	Sample preparation: IgG	65
4.3.6.4	ATR-FTIR experiments	66
4.3.6.5	Reflectivity experiments	67

5 Determining the Influence of Elution from Protein A on the Rate of IgG4 Aggregation69

5.1	Introduction and Objectives	69
5.2	Experimental Design.....	72
5.3	Rate of monomer loss at low pH in solution.....	79
5.3.1	Negligible effect of maltose.....	82
5.4	Effect of column step on rate of monomer loss at low pH.....	83
5.4.1	Monomer leading shoulder	85
5.4.2	Variation in monomer plateau.....	85
5.4.3	Influence of column step on R_0	88
5.5	Conclusion	90

6 Using Neutron Reflectivity to Examine the Structure of IgG4 Adsorbed to a Model Surface92

6.1	Route to reflectivity	92
6.2	Neutron reflection with IgG at the NCNR.....	93

6.3	Chapter objectives.....	93
6.4	Neutron measurements for biological materials	94
6.5	Neutron Reflectivity.....	96
6.6	Neutron Reflectivity Theory and Data Analysis.....	100
6.6.1	Basic reflectivity theory	100
6.6.2	Experiment design.....	104
6.6.3	Data analysis methodology	107
6.6.3.1	Data reduction.....	107
6.6.3.2	Building a constrained model	107
6.6.3.3	Evaluating the model	109
6.6.3.4	Adjusting the model.....	110
6.7	Examining the conformation of IgG4 adsorbed to silica	112
6.7.1	Sample preparation and stability assessment.....	112
6.7.2	Conformation of IgG4 adsorbed at pH 4.1.....	113
6.7.2.1	Reflectivity plots.....	114
6.7.2.2	Scattering length density-distance profiles: part 1	117
6.7.2.3	Scattering length density-distance profiles: part 2.....	122
6.7.3	Conformation of adsorbed IgG4 under chromatographic conditions.....	127
6.8	Conclusions.....	131

7 Characterisation of IgG4 Bound to Immobilised

Protein A..... 134

7.1	Introduction.....	134
7.2	Chapter objectives.....	135
7.3	Generating the model system.....	135
7.3.1	Cross-linker chemistry	135

7.3.2	Cross-linking staphylococcal protein A to silica	138
7.3.3	Blocking agents.....	141
7.4	Monitoring surface modifications with ATR-FTIR.....	142
7.5	Neutron reflectivity: characterisation of IgG adsorbed to protein A at the solid-liquid interface	155
7.5.1	Reflectivity experiments and data modelling.....	155
7.5.2	IgG adsorption to rSPA-modified surface with BSA blocking.....	157
7.5.3	IgG adsorption to rSPA-modified surface with PEG blocking.....	167
8	Conclusion.....	175
9	Future Work.....	178
10	References	180
11	Appendix	192
11.1	Publication	192
11.2	Conferences.....	192

1 Thesis Description

Protein aggregates often form during or after chromatographic separations, but the mechanisms driving aggregation in adsorption processes are not well-understood. A typical case is affinity capture of monoclonal antibodies by protein A chromatography. Here, elution of the antibody product is effected by passing low pH buffer through the chromatography column. Low pH is known to affect protein stability and cause aggregation. Thus, it is generally assumed that protein aggregates found in these elution pools form solely as a result of pH adjustment. However, emerging evidence indicates that other facets of the chromatography process can influence protein aggregation phenomena.

The overall objective of this work is to improve understanding of how adsorption-mediated processes affect the aggregation behaviour of protein biologics. In this work a single model protein will be investigated: an IgG4. The purified IgG was kindly donated by UCB Celltech (Slough) and represents a common class of biotherapeutic molecules.

Inspiration for the first results chapter, *section 5*, was drawn largely from work by Shukla et al. (2007a). They found that, for an Fc fusion protein, the rate of aggregation at low pH was influenced by a prior protein A chromatography step. The objective of this chapter is to produce data that further shows the effect of affinity-based separation on subsequent aggregation behaviour of the protein product. Attempting to quantify the extent and elucidate the nature of the effect for a monoclonal antibody (mAb) will represent an original contribution to understanding in this field.

In *section 6* a new direction for the work is established. The focus is shifted towards understanding how adsorption at solid-liquid interfaces affects protein structure, and subsequent aggregation behaviour, by direct measurement of adsorbed IgG structures using neutron reflectivity methods. Neutron measurements are used relatively rarely in bioprocessing and biotechnology research. However, these methods are potentially useful tools for obtaining unique structural information on biological molecules in their native state in bioprocess-related

environments. In *section 6* a simple model system is characterised using neutron reflectivity: IgG adsorbed to silica under chromatographically relevant buffer conditions. Here, the objective is to find out how the IgG orients in the presence of a charged surface and the related effect of buffer conditions.

In *section 7* the complexity of the model system is increased significantly. The model surface is built-up to include a covalently attached protein A ligand and one of two different blocking molecules. After characterisation of the modified surface, the organisation of adsorbed IgG molecules is probed in this new model system.

A promising direction for future work evolves from *sections 6* and *7*. Potential use of small angle neutron scattering to track the formation of aggregate species inside packed resin particles is discussed in *section 9: Future Work*.

2 Aims and Objectives

2.1 Context and Motivation

In biotherapeutics manufacture, chromatography operations are used to remove contaminants, increase product purity and reduce process volume during downstream processing of many therapeutic molecules. Product value is high at this relatively late stage in the process, so obtaining maximum product yield is essential. Process conditions encountered during chromatography can be detrimental to protein stability and can cause product molecules to aggregate, reducing product yield. In affinity chromatography, the products are eluted from chromatography columns at low pH. Aggregation induced by low pH is a well-documented phenomenon (Cromwell et al. 2006, Ejima et al. 2007, Hari et al. 2010, Shukla et al. 2007a). However, it is suggested that low pH is not the sole cause of aggregation in affinity chromatography; rather, a number of factors relating to resin-product interactions may be involved. For example, Shukla et al. (2007a) showed that the aggregation rate of an Fc fusion protein at pH 3.6 increased when protein A chromatography was carried out immediately prior to low pH incubation (Shukla et al. 2007a).

The changes in protein conformation induced by chromatographic separation, and the subsequent effect on aggregation propensity when further stresses are introduced (such as low pH), have been the subject of only limited investigation.

The overall aim of this work is to improve understanding of how adsorption-mediated processes affect the aggregation behaviour of protein biologics. In this work a single model protein is investigated: an IgG4. The purified IgG was kindly donated by UCB Celltech (Slough, UK) and represents a common class of biotherapeutic molecules. A protein A-based chromatography medium is used as a model resin; this has significant industrial relevance, as will be highlighted throughout *section 3 – Literature Review*.

Many of the following objectives are specific to a system involving IgG4 and protein A affinity chromatography. However, it is hoped that the outcome of this research will be applicable to a

range of sorbents and therapeutic proteins, and lead the way to a more comprehensive understanding of adsorption-related stresses and how to limit their impact on product quality and process costs.

2.2 Objectives

- To develop a quantitative method for determining whether elution from a protein A column affects the kinetics of IgG aggregation beyond the (expected) effect of low pH alone.
 - Generate kinetic data for IgG aggregation in protein A elution pools, and at low pH without a chromatography step.
 - Ensure the above-described data-sets are suitable for direct comparison with one another, that is, the influence of the *column step* can be clearly identified and all other experimental factors are well-controlled.
- To gain insight into the IgG aggregation mechanism with a view to ascertaining how the chromatography process affects aggregation from a mechanistic perspective.
- To identify process parameters that may impact column-related aggregation behaviour.
- To improve understanding of the effect of adsorption processes on IgG structure, adding to mechanistic insights derived from kinetic data.
- To generate data regarding the behaviour or structure of IgG adsorbed at surfaces under chromatographic conditions, in order to advance understanding of how surface conditions might influence aggregation behaviour.

3 Literature Review

The following review of the literature will describe how the native structure of an antibody is brought about, discuss the basic principles of protein aggregation and its causes, assess the mechanism of affinity chromatography and how it might contribute to protein aggregation, and explore some of the analytical techniques most commonly used to quantify and characterise protein aggregates.

3.1 Antibodies

3.1.1 Protein structure

Antibodies are protein molecules, and so share the same basic structural characteristics as all proteins. The structure of a native protein can be described in three to four levels: primary, secondary, tertiary and sometimes quaternary. Primary structure refers to the sequence of amino acids, or polypeptide chain. Often, proteins contain a number of polypeptide chains that subsequently form distinct elements of secondary or tertiary structure. The individual amino acids of a polypeptide chain are called residues. Residues are held together in a polypeptide chain by covalent peptide bonds, which are planar bonds that do not rotate (Cozzzone 2001). Primary structure is important because it defines the location of certain residues whose unique properties allow tertiary structures to form and, importantly, the protein to interact with other molecules to carry out its function. For example, cysteine residues contain the sulfhydryl group which is necessary for disulphide bond formation (Cozzzone 2001). Secondary structure refers to the local interaction of amino acids within a polypeptide chain, or one section of it, to form an ordered three dimensional structure. One reason for the order and rigidity of secondary structures is the planar nature of peptide bonds. Two major types of secondary structure are the alpha helix and the beta sheet. There are many variants of these classes of secondary structure including various helices, turns (where the polypeptide chain bends back on itself) and motifs such as the Greek key and the beta hairpin (specific beta strand arrangements) (Cozzzone 2001, Hubbard and Kamran Haider 2001). Unique protein motifs usually have highly specialised biological functions, however, these will not be described here.

Alpha helices form due to hydrogen bonding between polypeptide backbone carbonyl and amide (N-H) groups located 3.6 amino acids apart from one another (Cozzone 2001). The alpha helix is generally a stable structure. Alpha helices are nearly always right-handed because in this arrangement large side chains are not too close to the main chain; in left handed helices the close proximity of side chains and the main chain causes instability (Cozzone 2001, Hubbard and Kamran Haider 2001), see *Figure 3.1*

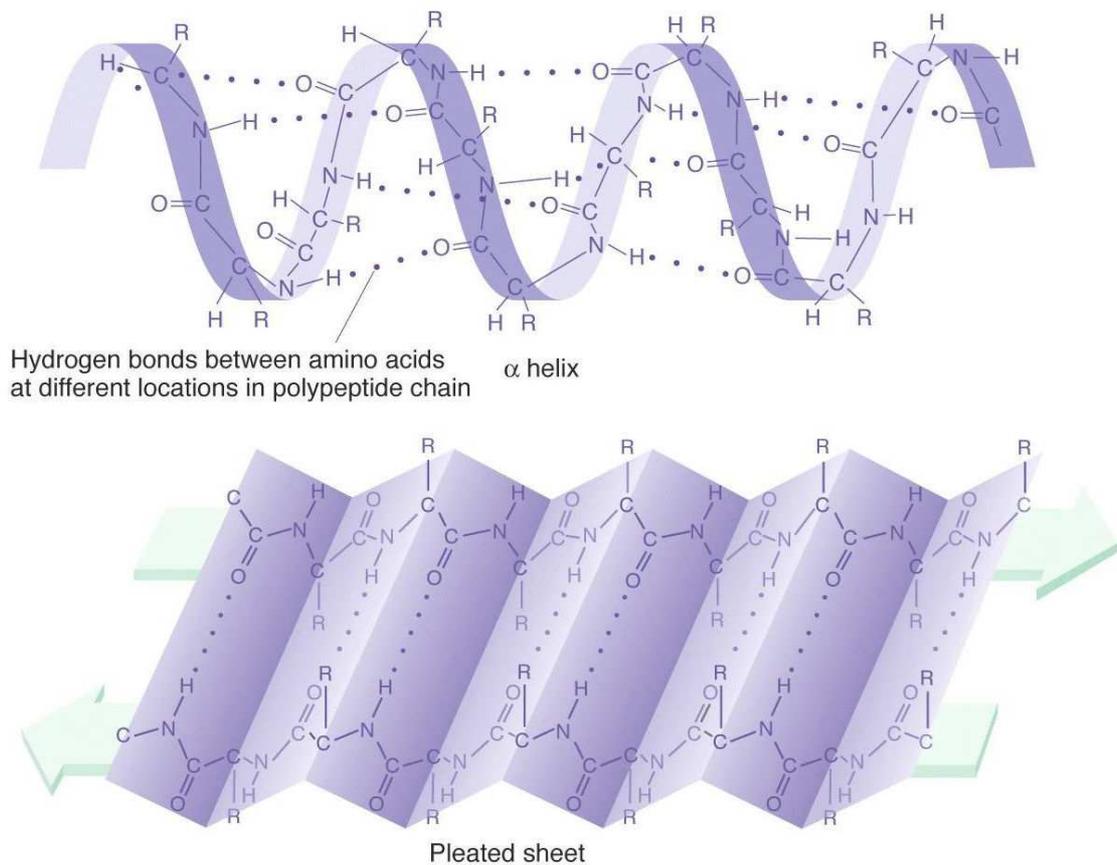


Figure 3.1 Common protein secondary structures: right-handed alpha-helix (top) and anti-parallel beta-sheet (bottom). Hydrogen bonds are represented by consecutive dots.

Figure 3.1 was taken from http://www.mun.ca/biology/scarr/MGA2_03-18b.html.

Beta sheets consist of parallel or antiparallel beta strands (commonly part of a continuous polypeptide chain) that are bound adjacently by backbone carbonyl-amide hydrogen bonds; successive adjacently bound strands result in a sheet-like structure (see *Fig.3.1*). Amino acid side chains are orientated roughly perpendicular to the sheet. Beta sheets have “pleated” surfaces because the chiral C_α atoms exhibit tetrahedral bonding within each residue, i.e. residues are not flat (see *Fig.3.1*). The key structural motif native to immunoglobulins is the immunoglobulin fold in which 7-9 anti-parallel β strands are folded around a hydrophobic core (Bork et al. 1994).

Tertiary structure is the specific folding arrangement of different elements of secondary structure and their orientation in three dimensional space. Tertiary structures are held in place by interaction of residue side-chains (Lumry and Eyring 1954). Tertiary structures often form as a means of hydrophobic burial –hydrophobic residues orient towards the centre of the molecule and polar residues dominate at the surface. This is because physiological fluids are aqueous and contain ions, meaning they interact favourably with polar residues but not hydrophobic ones (Chan 2001, Hubbard and Kamran Haider 2001). Tertiary structure is crucial to biological function because it generates the three dimensional form that contributes to mediating interaction with other molecules, e.g. a binding pocket, and also places specific side chains needed to facilitate these interactions in the right locations.

A fourth level of protein structure, quaternary structure, involves grouping together of a number of protein molecules to form larger multimeric entities. In these instances the individual protein molecules are referred to as subunits. Multimeric proteins usually require all subunits to be present to carry out their biological functions.

Most protein structures can be divided into units called domains. Domains are fairly small tertiary folded structures, commonly around 150 residues, which constitute distinct functional units of a protein (Yeats and Orengo 2001). Protein domains function cooperatively. For example, one domain may bind a receptor and prompt changes in its neighbouring domain

allowing it to bind a molecule, and so on. In biology, two major forms of domain include globular and transmembrane.

3.1.2 Antibody structure

Antibodies are made up of four polypeptide chains. There are two ~50kDa *heavy chains* flanked by two ~23kDa *light chains* (Lydyard et al. 2011). The four chains are arranged in a Y-shape held together by di-sulphide bonds. Each chain contains a constant region and a variable region; constant regions vary between but not within antibody subclasses, while variable regions differ in their millions (Goding 1996). Each chain contains a number of highly homologous globular domains known as immunoglobulin domains, each containing ~110 amino acids. Each heavy chain contain 3-4 constant domains denoted C_{H1} , C_{H2} , C_{H3} , C_{H4} and one variable domain denoted V_H . The light chains each consist of two domains, V_L and C_L . Heavy chains are roughly adjacent up to the “hinge” region, which lies between C_{H2} and C_{H1} , where the two chains significantly diverge to form the Y shape. Hinge regions vary in length and flexibility depending on the antibody class and subclass (Vermeer and Norde 2000); this can significantly affect antibody properties. A schematic representation of antibody structure is shown in *Figure 3.2*.

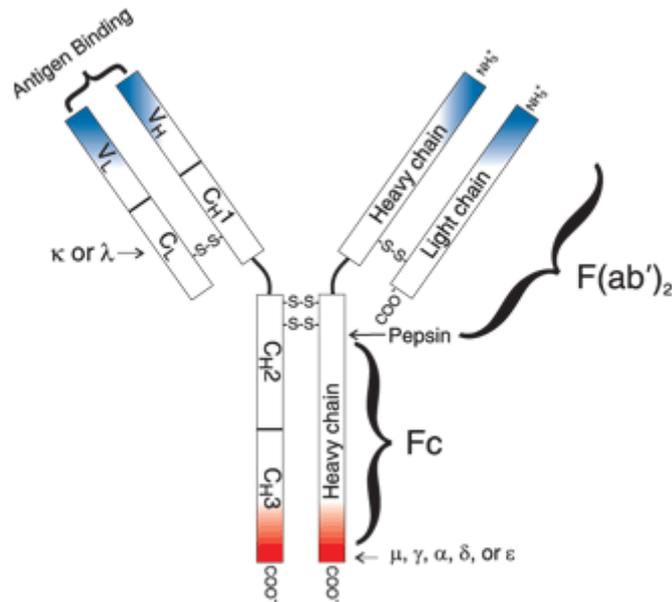


Figure 3.2 Schematic of antibody structure. F(ab')₂ refers to the two joined Fab fragments that would be produced if pepsin digestion were to occur at the marked location (arrow). μ, γ, α, δ and ε denote the different heavy chains found in different antibody classes. All IgG antibodies have γ heavy chains. K and λ represent the different possible light chain types.

Figure 3.2 was adapted from <http://www.invitrogen.com/site/us/en/home/References/Molecular-Probes-The-Handbook/Technical-Notes-and-Product-Highlights/Antibody-Structure-and-Classification.html>.

Individual IgG domains all have roughly the same characteristic fold consisting of two beta sheets and barely any α-helices. Non-polar residues are orientated towards the centre of globular domains to reduce contact with water (Vermeer and Norde 2000).

A carbohydrate moiety is attached to each C_{H2} domain. This helps to maintain certain limitations on antibody conformation (Deisenhofer 1981) and may improve stability by bridging the gap between the two C_{H2} domains. It is suggested that the two oligosaccharides keep the Fc region “open” with the two C_{H2} domains pushed apart, allowing the antibody to carry out its binding functions (Krapp et al. 2003).

3.2 Aggregation

In protein science, aggregation is a general term that refers to the generation of non-native protein species consisting of at least two originally-native protein monomer units. Large protein aggregates may contain hundreds of monomers. The individual units of an aggregate may be native or non-native state protein, and the aggregation mechanism(s) may be reversible or irreversible. Where aggregated protein units remain in their native state, the mechanism is likely reversible; alteration of solution conditions could reverse aggregation in such cases (Chi et al. 2003, Cromwell et al. 2006). Conversely, irreversible aggregation often occurs when proteins partially denature, and subsequently aggregate as a means of gaining stability (Shukla et al. 2007a). This is a form of physical degradation which compromises the protein's chance of successful manufacture and sale. In the biopharmaceutical industry, aggregates are classed as contaminants and must be cleared to acceptable levels if a protein formulation is to be used therapeutically (Chi et al. 2003, Cromwell et al. 2006).

3.2.1 Aggregate classification

Aggregates classified as “small” are usually soluble oligomers consisting of ≤ 10 units; oligomers classified as “large” may range in diameter from $\sim 20\text{nm} - 1\mu\text{m}$, depending on the size of the monomeric protein. Insoluble aggregate levels can be quantified in terms of turbidity, though particles contributing to turbidity that are not visible by eye may range in size from $1\mu\text{m} - 50\mu\text{m}$. These are often termed “subvisible particles” (den Engelsman et al. 2011, Mahler et al. 2009, Philo 2006, Sharma et al. 2009).

The forces that hold aggregates together may include weak forces such as electrostatic or hydrophobic interactions and Van der Waals, or covalent bonds such as di-sulphide bridges and di-tyrosine cross-links (Mahler et al. 2009, Sharma et al. 2009). Covalently bound aggregates are unlikely to be reversible, and often it is the smaller aggregate species that are reversible (Philo 2006). Non-native aggregation can be defined as aggregation phenomena that result in changes in secondary structure content of the monomers involved, which contribute to stabilisation of the aggregate therein (Roberts 2007).

3.2.2 Aggregation kinetics

Experiments exploring aggregation kinetics often assume that monomer loss is directly proportional to aggregate formation (Hari et al. 2010, Shukla et al. 2007a); that is, any monomer lost during an experiment must have become part of an aggregate in the same solution. This simplifies determination of aggregation rate, without the need to characterise the aggregates themselves. Determining the rate of monomer loss can be a useful tool for elucidating extrinsic factors that affect aggregation rates; size exclusion chromatography (SEC) is the usual method. However, other methods must be implemented to identify intrinsic inducers of aggregation. Such methods would ideally characterise individual mAbs in terms of key predictors of aggregation propensity, such as melting temperatures of the different constant domains which could be related to unfolding rates induced by external factors.

Size distribution data is useful for understanding aggregation kinetics. Determining how the size distribution of particles (including monomers) changes over time and may help to identify rate-limiting steps and so contribute to building or support of kinetic models for aggregation mechanisms.

Many early models of aggregation kinetics were based on native aggregation. It was thought that extensive aggregation could occur without significant alteration of the native monomer. These models describe a process more akin to polymerisation, whereby small reversible native oligomers form and grow in size as more monomers associate at a constant rate. Presently, models of non-native aggregation are more widely accepted and are the subject of much investigation (Roberts 2007, Roberts 2014).

The Lumry-Eyring model is frequently used to describe aggregation kinetics (Chi et al. 2003, Roberts 2007, Shukla et al. 2007a); this model assumes two basic stages. Prior to aggregation the native protein undergoes destabilisation which may constitute a level of unfolding; unfolded molecules then form irreversible aggregates. This could imply second order kinetics; however, in this model only the first stage is in equilibrium (native protein \leftrightarrow unfolded protein). Subsequent aggregate formation is considered relatively instantaneous and irreversible. Thus

kinetics should be first order (Roberts 2007, Shukla et al. 2007a). It is, therefore, the rate of protein unfolding that becomes of great interest when using this model. Moreover, the level of unfolding that necessitates aggregation must be queried as this will contribute to both the rate of aggregation and its potential for reversibility.

Generally two major stages of aggregate formation are discerned in all models: nucleation followed by growth. In non-native aggregation nucleation is thought to constitute the following events:

(i) Monomers fold and unfold in equilibrium; a large range of conformations may be generated and reversed.

(ii) Certain non-native or partially unfolded conformations tend to associate reversibly, forming reversible oligomers.

(iii) The monomeric units of reversible oligomers adopt conformations that favour stabilisation of the aggregated assembly; this may involve alteration of protein secondary structure as well as tertiary.

(iv) Simultaneous to event (iii), the oligomer is stabilised and so the aggregate becomes irreversible (Chi et al. 2003, Roberts 2007).

Monomers described in stage (ii) are often termed *reactive monomers*.

The growth stage involves addition of the relevant non-native monomers to the aggregate *nucleus* to form larger aggregates that may be filamentous or amorphous in structure. This could involve initial reversible association of the additional monomer(s) with the aggregate nucleus, followed by some form of re-arrangement that irreversibly binds the two. Growth is usually indicated to be faster than initial nucleation, so a bias towards a forward irreversible mechanism is often presumed.

Large filamentous aggregates may associate further among themselves to form gels or fibrils, while amorphous aggregates may form insoluble amorphous precipitates (Roberts 2007). The stages of non-native and native aggregation are depicted schematically in *Figure 3.3*.

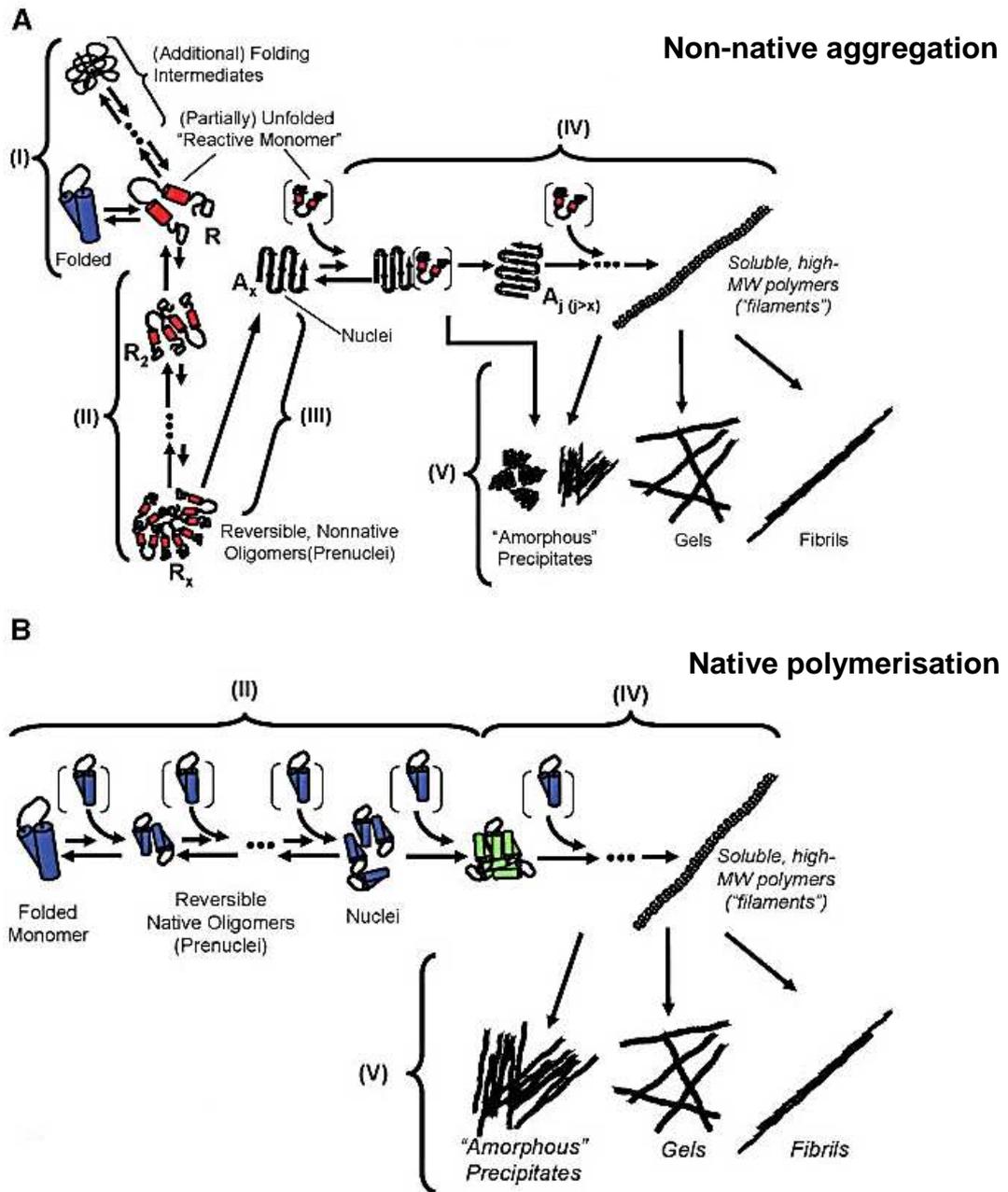


Figure 3.3 (a) Non-native aggregation and (b) native polymerisation pathways shown schematically. In (a) (I) native monomer reversibly forms reactive monomer (R) (II) R reversibly dimerises (R_2) and dimers in turn form reversible oligomers (R_x), (III) oligomers become irreversible (A_x), completing the nucleation phase. Growth ensues either (IV) by addition of R to form soluble filamentous polymers which may associate further to form gels or fibrils, or (V) by association of irreversible oligomers (A_x) to form amorphous precipitates. In (b) similar stages occur while the monomer remains in its native state.

Adapted from Roberts (2007). Copyright © 2007 Wiley Periodicals, Inc.

Some models argue that association of non-native monomers is the rate-limiting step, rather than unfolding. Unstable structures have high free energy, and because high free energy is thermodynamically unfavourable, such structures are expected to rapidly adopt new conformations with lower free energy. However, the unfolded monomer state has only marginally more free energy than the folded; unfolding also causes an increase in conformational entropy which is energetically favourable (Chi et al. 2003, Roberts 2007). Equilibrium between unfolded and native or re-folded monomer may be rapidly established under certain conditions, and association of non-native monomers may become rate-limiting.

3.2.3 Intrinsic factors

The probability of protein aggregation is a product of intrinsic and extrinsic factors. Where intrinsic factors are concerned, some proteins lend themselves to aggregation whilst others are more stable. Elements of protein primary, secondary and tertiary structure all contribute to aggregation phenomena. Secondary structure is suggested to be crucial; aggregated species often contain significantly altered segments of secondary structure, a commonly adopted non-native form being the β -sheet (Andrews and Roberts 2007, Chi et al. 2003, Fernandez et al. 2003). Beta-sheets may form as part of the aggregation mechanism, or be pre-existing and undergo moderate alteration during aggregation (Andrews and Roberts 2007). Strong roles of primary structure, that is, certain sequence motifs, have also been indicated (Wang et al. 2009). Wang et al. (2009) used computational tools to identify potential aggregation-prone motifs in 21 commercial mAbs and 20 non-commercial mAbs. The predicted aggregation-prone motifs consisted of 5-10 consecutive residues located primarily in or around the CDR loops of the antibodies. CDR-proximal motifs identified as aggregation prone tended to differ between different antibodies; while motifs located elsewhere in the sequences were usually highly homologous. Hydrophobic and aromatic residues were common in aggregation-prone regions. This is supported by results of a study by Biddlecombe et al. (2009) who found three different IgG4 molecules to aggregate at different rates. The three IgG4 sequences were highly homologous in all regions apart from the antigen-binding regions. It was suggested that antigen-

binding regions may be susceptible to destabilisation simply because they are more “exposed to the solution environment” (Biddlecombe et al. 2009).

Conversely, Hari et al. (2010) found the Fab fragment (that containing CDR loops) to be highly stable when subjected to aggregation-inducing conditions. They purified native Fc and Fab fragments of an IgG2 to try and elucidate the intrinsic cause of IgG2 aggregation at low pH. Fc and Fab fragments and the full length antibody were incubated at pH 3.5 for 50 hours (also in the presence of 0.5M NaCl). Samples were taken at various time points and aggregation was quantified in terms of monomer loss using cation exchange chromatography. They found that the Fab fragment was highly resistant to aggregation, whereas the Fc fragment had the highest aggregation rate –greater than 90% monomer loss after 50 hours (Hari et al. 2010). Interestingly, the rate of full-length antibody aggregation was not far behind that of the isolated Fc fragment. This suggests that instability in one region of the IgG strongly influences aggregation propensity of the whole antibody, and that the presence of apparently stable domains does not necessarily exert a stabilising effect.

Furthermore, a study by Ishikawa et al. (2010) supports the theory that specific intramolecular interactions between variable and constant domains influence Fab stability. They used differential scanning calorimetry to show that IgG molecules of different subclasses with identical variable regions varied in their Fab melting temperatures, indicating a role of constant domains in Fab stabilisation (or lack thereof). The T_m of IgG1 Fab was approximately 6°C higher than that of IgG4 Fab. The authors found the same IgG molecules to have near-identical CD-spectra, suggesting that differences in Fab stability were not due to variation in secondary structure, but possibly the result of different patterns of di-sulphide bonding and other inter-domain interactions (Ishikawa et al. 2010).

However, another study showed that an IgG1 and an IgG4 with the same variable regions had very similar Fab T_m values (Garber and Demarest 2007). The results of Garber and Demarest (2007) indicate far greater variation in Fab T_m values within subclasses (here, IgG1); the authors suggest that this variation is due to V_H domain sequence irregularities.

Some may argue that tertiary structure is the key determinant of a protein's aggregation propensity because it creates the protein's surface properties such as charge distribution, hydrophobic patches, and thermal stability (Ishikawa et al. 2010).

3.2.4 Extrinsic factors

There are numerous external factors that may significantly accelerate aggregate formation in protein solutions. Mechanical stresses caused by agitation or pumping motions, or shear at a solid-liquid interface solution conditions such as temperature, pH, protein concentration and the concentration of other substances; liquid-air interfaces; adsorption and desorption events; protein-surface/-material interactions –all of these factors, alone or in combination, have been purported to influence aggregation pathways (Biddlecombe et al. 2009, Cromwell et al. 2006).

Key points of interest in this work will be antibody aggregation at low pH and antibody aggregation as a result of adsorption to chromatographic media. Low pH is frequently cited as a cause of aggregation (Chi et al. 2003, Cromwell et al. 2006, Hari et al. 2010, Ishikawa et al. 2010, Shukla et al. 2007a), though it is not just acidic pH that can induce aggregation. Different proteins have different, often fairly narrow, pH ranges within which they are stable in their native state and thus have low propensity for aggregation (Chi et al. 2003, Townsend and Deluca 1990).

Solvent composition can significantly affect both the rate and extent of antibody aggregation. Hari et al. (2010) demonstrated that not just low pH, but specifically increased acid concentration increased the rate of monomer loss in solutions of IgG2.

Vermeer and Norde (2000) found that protein unfolding induced by different methods resulted in formation of aggregates with different structural characteristics. This highlights the importance of extrinsic factors in inducing aggregation, as different forms of stress may induce aggregates of varied character from a single monomer population.

It is quite clear from the literature that the influence exerted by any extrinsic inducer of aggregation depends greatly on the protein. Nevertheless, external factors can have extreme

effects, and it may indeed be an easier task to optimise processing conditions for every antibody (or other protein) candidate produced than to generate only antibodies that will survive well under narrowly defined conditions. For this reason, it is important to understand how external factors and conditions can affect aggregation profiles in order to limit protein aggregation during bioprocesses.

Shukla et al. (2007a) used size exclusion chromatography to investigate on-column protein aggregation in affinity chromatography. They wanted to determine whether the adsorption process itself (particularly regarding protein A), rather than just the low pH used for protein elution, contributed to subsequent protein aggregation. The protein in question was an Fc fusion protein. The protein was loaded onto a protein A column, washed and then eluted with citrate buffer, pH 3.6. The elution fraction was left to incubate at pH 3.5 and samples were taken at various time points and neutralised. After neutralisation samples were analysed using SEC to quantify formation of high molecular weight (HMW) species, i.e. oligomers. Parallel “solution experiments” were carried out to mimic the pH 3.5 incubation, but without the prior Protein A column step. Finally, several additives were tested for their ability to reduce aggregation rates in both scenarios (*solution* and *on-column*). Percentage HMW species was plotted against time and a rate constant was derived; this was used to compare data sets. Preliminary experiments indicated that rate constants were not affected by uncontrolled parameters such as concentration and elution time, so provided a useful quantifier for all conditions tested. The differences in rate constants obtained for solution and on-column experiments were striking. In the absence of additives, the rate constant for aggregation was approximately 5 times greater after Protein A chromatography than in solution only. All post-column rate constants were greater than the in-solution equivalents. Urea was found to be an effective additive in reducing on-column aggregation; it was most effective at the lowest concentration tested, 0.5M. However, in solution urea was most effective at 1M. Urea may act as a stabiliser by “stiffening” backbone amide bonds and so decreasing rotational freedom (Shukla et al. 2007a).

Shukla et al. (2007a) point out that those additives which stabilise proteins in solution may have destabilising effects during adsorption/desorption events. This was the case for sucrose (10%) in their experiments.

These are important results that show extrinsic aggregation inducers to be additive, but suggest that the aggregation mechanism depends on the inducer(s). Moreover, the aggregation inducing effect of protein A desorption is clearly highlighted, benefiting further investigation.

3.3 Chromatography

3.3.1 Chromatography in bioprocessing

Chromatography is an indispensable method of purification used in downstream processing of biopharmaceuticals. Commonly, a number of chromatographic separations are carried out as part of a bioprocess in order to achieve adequate product purity and contaminant removal. Bioprocesses are often designed so that unit operations prior to chromatography achieve a high level of particulate removal, and so prevent column fouling and maximise resin lifetime. This is important because chromatography resins can be expensive, and their performance is imperative to product safety. Thus, when designing unit operations preceding chromatography emphasis is also placed on reducing process stream volume, minimising viscosity and maximising product concentration (viscosity increases with protein concentration) . Some engineering considerations:

- Viscosity is inversely proportional to the rate of mass transfer.
- Concentration is proportional to the rate of mass transfer.
- Lowering the flow rate will increase loading efficiency (less capacity wastage).
- Bed compression increases with flow rate...subsequently either reducing flow rate or increasing pressure drop.

It is important to note that when it comes to large scale biopharmaceutical manufacture, time is of the essence; manufacturers will often chose (or rather, be left with no choice) speed and simplicity and, even, familiarity over efficiency in order to meet productivity demands.

Chromatography has a number of modes including affinity, ion exchange and hydrophobic interaction. Affinity chromatography is the most costly of these, but this has not affected its popularity in industry. Affinity chromatography uses high selectivity ligands to capture product molecules; the ligands are covalently bound to a solid phase which is normally packed into a column. The high specificity of affinity chromatography makes it ideal for product “capture” in the early stages of purification; only the product binds the ligand while many protein and non-protein impurities pass through the column.

3.3.2 Protein A

3.3.2.1 Protein A as an affinity ligand

A near-ubiquitous ligand used for product capture in commercial bioprocesses is protein A, a protein that was originally purified from *Staphylococcus aureus*. The popularity of protein A lies in its ability to capture a range of whole antibodies by binding to their Fc region, along with its relative stability under the strong basic conditions used for column cleaning (Hober et al. 2007). Protein A has a strong affinity for human IgG subclasses 1, 2 and 4 and various non-human IgGs. The large majority of FDA-approved commercial mAbs fall into the categories: human, humanized or chimeric IgG1, IgG4 and, less commonly, IgG2; murine IgG1 and IgG2a (Hober et al. 2007, Wang et al. 2009). Protein A binds full length IgG with a high level of selectivity and minimal interaction with the Fab region (Salvalaglio et al., 2009). Though protein A has been shown to bind Fab domains of IgG, Fc binding is strongly preferential in full length IgG.

3.3.2.2 Structural characteristics of protein A and Fc

Protein A is found naturally in the cell walls of *Staphylococcus aureus*. It has a molecular weight of 42 kDa and contains five highly homologous Fc-binding regions located towards the N-terminal of the protein; the C-terminal region differs significantly in sequence from the Fc-

binding regions and makes up approximately 38% of the total residues (Moks et al. 1986, Sjobahl 1977). The domain structure of native staphylococcal protein A is depicted schematically in *Figure 3.4*, along with some of its engineered variants which are used in various commercially available protein A chromatography resins.

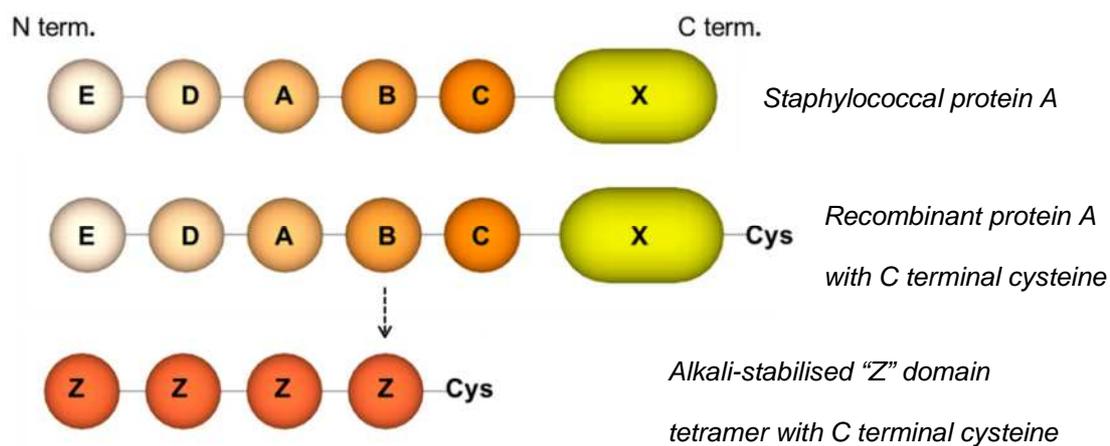


Figure 3.4 Schematic depictions of protein A ligands. Domains E, D, A, B and C are capable of binding the Fc of IgG molecules; the C terminal X domain is involved in cell wall attachment in the bacterial protein. The commercial MabSelect ligand consists of full length recombinant protein A with a C terminal cysteine for oriented coupling. The MabSelect Sure ligand is a tetramer of Z domains; the Z domain is derived from the native B domain and contains a number of asparagine substitutions (Lain 2013, Moks et al. 1986, Nilsson et al. 1987).

Protein A can bind at least two IgG molecules simultaneously (Sjodahl, 1977). The Fc-binding B domain denoted *Fragment B* was crystallographically refined in complex with a human IgG Fc fragment at 2.9Å and 2.8Å resolution by Deisenhofer (1981). The structure of Fragment B consists of three-helix bundle containing two antiparallel α -helices; these helical regions are involved Fc binding.

Deisenhofer (1981) elucidated some important characteristics of the protein A-binding region of the IgG Fc fragment. C_{H2} and C_{H3} domains were found to exhibit typical immunoglobulin domain structures: two layers of antiparallel β -sheets enclosing a hydrophobic core (Deisenhofer 1981). Deisenhofer observed no surface contact between the dimeric C_{H2} domains. The interior face of each C_{H2} domain has a carbohydrate moiety attached; the two carbohydrate moieties are thought to form a weak bridge between C_{H2} domains, perhaps substituting for protein surface contact. Surface contact exists between the dimeric C_{H3} domains, and the area of contact is approximately double that of the C_{H2} carbohydrate contact area (Deisenhofer 1981).

The C_{H2} domains of the Fc dimer had higher temperature factors than their respective C_{H3} domains, indicating greater disorder in C_{H2} domain structure. This corresponds with the numerous studies that show C_{H2} domains to have the lowest melting temperature compared to other immunoglobulin domains (Hari et al. 2010).

Recently, Deis et al. (2015) published work detailing the crystal structure of the C domain of staphylococcal protein A, SpA_C, in complex Fc at 2.3Å resolution. The new crystal structure showed that helix III of SpA_C was properly folded when bound to Fc. The work also revealed a second binding site for SpA_C to Fc at the helix 2/3 interface; however, isothermal titration calorimetry and simulation data suggested that this was a low affinity interaction (Deis et al. 2015). The structure of SpA_C bound to Fc at the main high affinity helix 1/2 SpA_C interface is shown in *Figure 3.5*. The residues on the protein A domain involved in Fc binding reside largely on helix I, with half as many on helix II (Deis et al. 2015). Interacting residues on the Fc side highlighted in *Fig.3.5* were described by Deis et al. (2015) and DeLano et al. (2000).

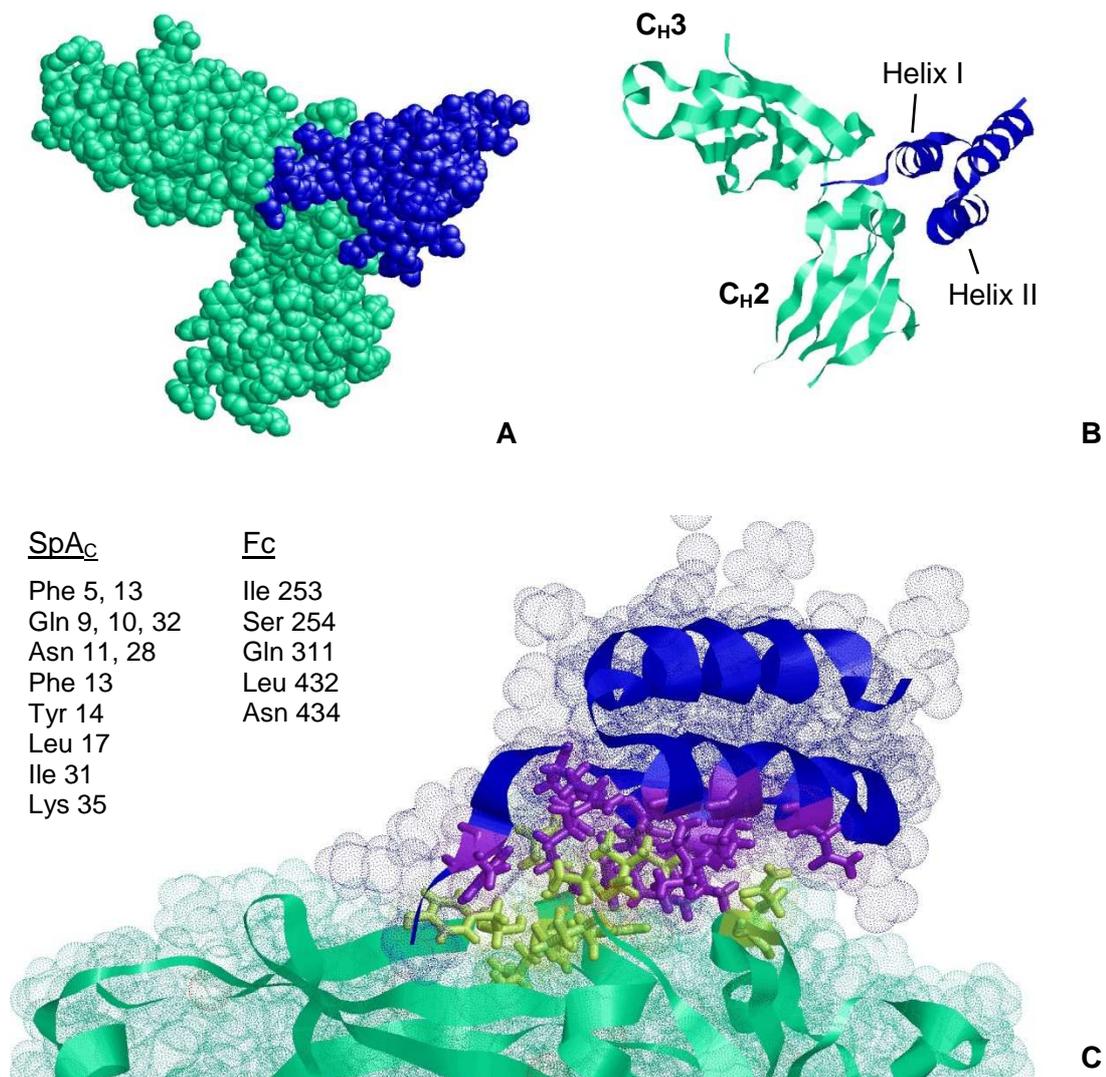


Figure 3.5 Crystal structure of the C domain of staphylococcal protein A, SpA_C , in complex with the Fc monomer fragment of human IgG. SpA_C is shown in blue and Fc in turquoise in spacefilling (a) and ribbon (b) models. In (c) the structure has been rotated and individual residues from SpA_C (purple) and Fc (green) involved in the binding interaction are shown as sticks and listed to the left of the image. In SpA_C , helix II begins at approximately residue 22. RCSB Protein Data Bank ID: 4WWI (biological assembly 2).

The crystal structure in *Figure 3.5* was obtained by Deis et al. (2015); the visual representation was edited using RasMol software (Bernstein 2000, Sayle and Milner-White 1995).

3.3.2.3 Fc-protein A interaction: potential to induce IgG aggregation

A study by DeLano et al. (2000) clarifies many features of the binding interaction between the IgG Fc region and the B1 domain (Fragment B) of protein A. Binding takes place about the hinge between the C_{H2} and C_{H3} domains of the Fc (see *Figure 3.5*). This precise region is known to bind with high affinity to a number of molecules other than protein A, including domain C2 of protein G, neonatal Fc receptor and rheumatoid factor. Hence, the region has been termed *consensus binding site* (CBS) (DeLano 2000). The CBS is largely hydrophobic in character, contains relatively few polar residues and has a high level of solvent accessibility. These features indicate burial of hydrophobic residues as a strong driver of binding (Deisenhofer 1981, DeLano 2000). Interestingly, work by Fernandez et al. (2003) speculates that high order protein folds are orientated such that backbone hydrogen bonds are protected from water attack by hydrophobic groups in what is termed the “desolvation domain”. The “desolvation domain” is defined as the space within two overlapping 7Å spheres, one centred on each of the alpha carbons of a hydrogen bonding pair of residues. It is supposed also that the presence of hydrophobic rather than polar or charged groups in the hydrogen bond desolvation domain reduces the charge screening effect on the H-bonding pair, strengthening the bond (Fernandez et al. 2003). This mechanism may contribute to the preservation of protein secondary structures.

It may be possible then that adsorption to a ligand by hydrophobic interaction could result in destabilisation of secondary structure, as the precise orientation of hydrophobic residues is altered to allow for their interaction.

Deisenhofer (1981) noted that C_{H2} disorder was greater in the Fc-Fragment B complex than in the unbound Fc, implying potential Fc destabilisation on adsorption to protein A. However, this may not have been the case had the Fab fragment been present, due to potentially stabilising C_{H2}-Fab contact.

According to Delano et al. (2000), the CBS undergoes considerable conformational changes when binding to a ligand. In fact, the nature of the change in conformation depends on the

ligand, highlighting the flexibility of this region, and explains why this non-polar accessible region of the Fc is the preferred binding site for a number of different ligands (DeLano 2000). The interaction is not purely hydrophobic; certain highly conserved hydrogen-bonding sites also exist in the CBS (Shukla and Hinckley 2008).

A different study indicated that electrostatic and Van der Waals forces contributed similarly to the binding energy. It was shown that hydrophobic interactions occur largely at the protein A- C_{H2} interface, whilst electrostatic forces represent the main interaction with the C_{H3} region of Fc (Salvalaglio et al. 2009).

The binding mechanism does not depend exclusively on ligand chemistry; properties of the support material (e.g. agarose) and spacer arm also influence IgG binding (Lowe 2001, Salvalaglio et al. 2009). Moreover, chemistry is not the only force at work. As noted by Lowe (2001), “geometrical and steric constraints imposed by the complex three-dimensional matrix environment” cannot be overlooked.

3.3.2.4 Protein A supports – influence on adsorption

For a resin to achieve commercial success it must have certain basic properties, aside from ligand chemistry, that allow application of general affinity chromatography methods without limitation. Non-specific interactions between resin and protein, and other molecules, must be strictly minimal for purity standards to be met. The load material should flow freely through the resin and become evenly distributed throughout the column, both between particles and within pores, and large impurities must not become trapped within the beads. Thus, achieving high porosity whilst minimising dead volume is essential, and an even distribution of roughly spherical particles is ideal (Cuatrecasas 1970). Other important characteristics include hydrophilicity and sufficient binding sites for (spacer and) ligands (Busini et al. 2006). Compressibility is another significant factor because it governs pressure limitations which in turn limit process productivity (Fahrner et al. 1999).

Ligand leaching can be particularly problematic in resins containing natural receptors like protein A. Thus, in recent years considerable effort has been made to develop synthetic ligands,

and to modify natural ligands using recombinant technology. The engineering challenge is to produce affinity resins that retain the excellent selectivity and affinity of natural ligands whilst improving stability against high flow rates and harsh clean in place (CIP) protocols in order to reduce ligand leakage and extend column lifetime. Such improvements would reduce costs considerably. Higher capacity resins are also increasingly sought-after (Hober et al. 2007, Lowe 2001, Salvalaglio et al. 2009, Zamolo et al. 2008). Concomitantly, investigation into the effect of the resin support material, commonly agarose, on ligand-product binding is on the rise.

Salvalaglio et al. (2009) used a modelling approach to assess the effect of the support, namely agarose, on the protein A-IgG interaction. The IgG-protein A interaction model was based on the crystal structure by Deisenhofer (1981); the agarose support molecular model was generated by molecular dynamics simulation. A minor shift was identified in the predicted binding energies when protein A was bound to an agarose support; electrostatic interactions increased and Van der Waals forces decreased. However, the overall IgG-protein A binding energy was not significantly affected by the presence of agarose (Salvalaglio et al. 2009). Zamolo et al. (2008) report similar findings regarding overall binding energy. However, in molecular dynamics simulations generated by Zamolo et al. (2008), transient interactions between Fc fragment and agarose were observed; Fc backbone and side-chain carbonyls were seen to temporarily hydrogen-bond with agarose hydroxyl groups. These bonds were not present once the system had stabilised, but such incidents could affect later aggregation behaviour.

As mentioned previously, the three-dimensional matrix environment is an important determinant of a resin's performance in binding its target molecule. Physical characteristics of a resin, in combination with parameters like flow rate and concentration of load material, might also impact aggregation phenomena. A summary of typically desirable characteristics for affinity chromatography resins and their predicted effect on aggregation phenomena is provided in *Table 3.1* (Hober et al., 2007).

Table 3.1 Selected resin characteristics, subsequent column conditions and predicted effect regarding aggregation phenomena.

Resin characteristic	Resultant column conditions	Potential to affect aggregation phenomena
High dynamic binding capacity	Greater protein binding	Greater concentration effect could increase aggregation.
Stability at increased flow rates	Higher flow rates	May increase shear stress – can induce aggregation.
High porosity (retaining rigidity)	Greater protein binding	Easier transit through column may reduce physical stresses – potential to reduce aggregation risk. However, more protein binding could cause concentration effect noted above.
Small particle size	Improved dynamic binding capacity / higher flow rates	Possible concentration effect. Shorter diffusion distance could reduce stress.
Large pore size	Improved mass transfer	No expected effect of mass transfer alone, but longer diffusion distance could result in increased stress.

Table 3.2 Selected recombinant (r-) and natural (n-) protein A media and their basic compositions.

Trade Name	Ligand	Support matrix
MabSelect SuRe	Engineered rProtein A	Cross-linked agarose
ProSep A	rProtein A	Controlled pore glass
POROS 50	rProtein A	Polystyrene/divinylbenzene
Affi-gel-Protein-A support	nProtein A	Polymeric
Protein A cellthru 300	rProtein A	Agarose

Table adapted from Hober et al. (2007).

Agarose is a popular choice of support for affinity resins, and modern agarose affinity media use cross-linked agarose for improved rigidity and stability, but other types of support have their benefits (Hober et al. 2007). For example, controlled pore glass resins have low compressibility. Some examples of different affinity media are given in *Table 3.2*.

3.3.3 Stages of affinity chromatography

The basic principles of affinity chromatography are displayed schematically in *Figure 3.6*, along with buffer conditions employed for typical protein A chromatography processes. The following sections give a brief overview of some of the process parameters, optimisation strategies and potential links with aggregation phenomena.

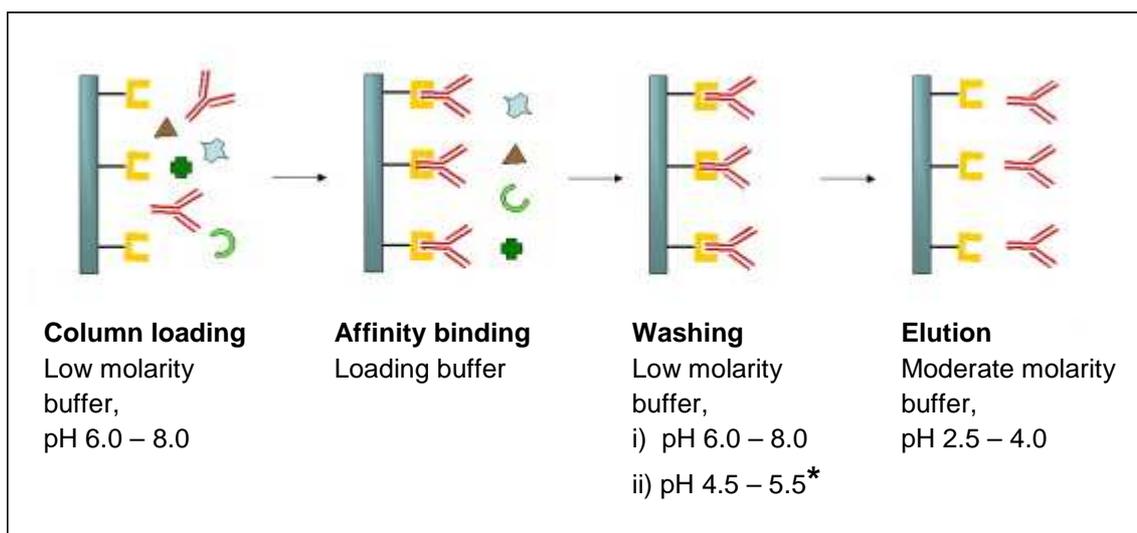


Figure 3.6 The stages of protein A affinity chromatography and typical buffer conditions for each stage. Squared-horseshoe shapes represent the affinity ligand, Y shapes represent the mAb and remaining small shapes represent impurities. For washing, the loading buffer may be used (pH 6.0 – 8.00), but commonly a different buffer, one which doesn't cause elution but does improve impurity removal, is used. The star symbol (*) indicates that the pH range shown is common but not always optimum

Figure adapted from Roque et al. (2007), reprinted with permission. Copyright © 2007 Elsevier B.V. All rights reserved.

3.3.3.1 Product loading

Protein must be loaded onto the column under conditions that favour binding of the protein product to the ligand. Typically pH conditions between 6.0 and 8.0 are effective, though optimum binding pH varies from product to product. Major factors that can be varied during protein loading are the concentration of product in the load material and its flow rate; these two

factors affect dynamic binding capacity (Swinnen et al. 2007) and so may contribute to variations in aggregation behaviour (see *Table 3.1*).

3.3.3.2 Column wash

After the clarified cell culture broth has been passed through the column allowing the target molecule to bind the affinity ligand, the column is usually washed with several column volumes of buffer to facilitate the removal of impurities like host cell proteins (HCPs). The buffer used for the wash step is usually optimised for the particular process, and will depend on the nature of the product, the chromatography medium and the types of impurity present in the load material. In protein A chromatography operations, the pH of the wash buffer often lies between that of the loading and elution buffers, with the aim of using the lowest pH possible that will maximise impurity removal without eluting product molecules (Shukla and Hinckley 2008). Shukla and Hinckley (2008) performed a number of experiments which aimed to improve understanding of HCP removal by washing in protein A chromatography. They found that when intermediate pH buffers (here, pH 4.4) were used, improvements in purity coincided with reduced product yield. Other results showed that HCP interaction with the column was reduced in the absence of IgG, indicating a co-interaction or co-elution mechanism for HCP and product. Based on these findings, high pH wash buffers containing various additives were investigated as an alternative to the traditional intermediate low pH. The results were positive; HCP removal was increased and yield maintained or increased for a number of different molecules compared to various control systems. Therefore, though intermediate low pH buffers are a common choice for washing, alternatives can prove beneficial (Shukla and Hinckley 2008). This could provide a route of investigation regarding the effect of wash conditions on aggregation propensity.

3.3.3.3 Product elution

Elution involves dissociation of the product-ligand interaction by a change in buffer conditions allowing the purified product to be collected. The elution buffer used depends on the nature of the product-ligand interaction, and also relative product stability. In protein A chromatography

low pH buffers are used, typically between pH 2.5 and pH 4. Shukla and Hinckley (2008) describe the elution mechanism as follows: Highly conserved ionisable residues that exist in both the CBS and the proximal protein A surface both become protonated at low pH. The positive charges generated repel one another resulting in a decrease in hydrophobic contact area such that the interaction is broken and product elutes. Lower pH may improve product yield or production rate, but could cause product destabilisation so a balance must be struck (Shukla and Hinckley 2008).

3.4 Analytical Methods

3.4.1 Analytical methods and aggregation

Aggregates present a unique challenge in analytical methodology due to their broad size range and varied morphologies. Protein aggregates can range in size from nanometres to millimetres in diameter, so there is no single piece of analytical equipment capable of detecting the full range of aggregate species that may be present in a single sample (den Engelsman et al. 2011, Philo 2006). The aggregation mechanism and aggregate character are often unknown, so it can be difficult to discern whether aggregated species are affected by the physical processes that take place during analyses, or the chemical environments generated. Some aggregation models stipulate that once an aggregation mechanism is triggered the process becomes very rapid; large aggregates may form, leaving relatively few intermediates (Roberts 2007). Various scenarios may exist in which a range of phenomena that confound the analyses could occur such as (i) aggregates continue to form after the reaction has been “stopped” (e.g. after neutralisation of acid environment); (ii) monomers are affected by experimental conditions but do not aggregate –they may undergo more subtle conformational changes that affect results; (iii) rates of dissociation for reversible aggregates are very slow –once the reaction is “stopped”, re-equilibration of the sample in the new environment takes hours or even days (Philo 2006). These are just some imaginative examples, but are meant to illustrate the sometimes unpredictable behaviour of relatively complex proteins in solution.

For these reasons researchers commonly employ orthogonal methods, whereby several fundamentally different analytical techniques are used to gather a variety of data pertaining to one characteristic of a sample, or identification of a single subset of aggregates, say, by size (den Engelsman et al. 2011, Mahler et al. 2009). For methods to differ orthogonally, their fundamental modes of detection or separation of different species must be different.

Any piece of analytical instrumentation should provide the following utilities: (a) detection – molecules or particles are “seen” by the instrument; (b) separation – the instrument can distinguish molecules/particles that are different from one another (by size / shape / charge / electron density, etc.), or facilitate such distinction by the user.

There are a range of techniques frequently used to study aggregation phenomena, each with different advantages and disadvantages. Questions worth considering when choosing methods of investigation include: What type of data is being sought – can a method of statistical analysis be clearly defined? How long will the method take – will the data required be gathered quickly enough, is it an efficient method? Will the sample be degraded by the analysis, and is this important? What quantity of sample is available for analysis? Is a very robust method needed, or is the objective to implement more advanced techniques? What are the associated costs; are they justifiable? Most analytical technologies have their caveats relating to how sample material behaves in the given system (also over time) or pre-treatment requirements (Philo 2006). Selected technologies will be described and their limitations discussed in the following sections.

3.4.2 Size exclusion chromatography (SEC)

3.4.2.1 How SEC works

An extremely common technique used to monitor aggregate formation is size exclusion chromatography (SEC) (den Engelsman et al. 2011, Ejima et al. 2005). In SEC different species in a sample are separated on the basis of size, and detected by a UV detector. Samples are injected into a column containing a resin or matrix. A common resin material is silica in the form of beads or particles. The particles are roughly spherical and porous; the pores in the particles are not straight-through, they are convoluted channels. A mobile phase (i.e. buffer) is

continuously passed through the column under high pressure; the sample is transported through the column in the mobile phase which permeates the void volume between and within all particles. As the sample moves through the column, large molecules travel between the silica particles, they are too large to enter the pores. However, small molecules can enter the pores and so take a more convoluted route through the resin. Hence, large particles travel more quickly through the column, arriving at the UV detector before small particles. The UV detector relays information to a live chromatogram where peaks are generated relative to molecule size (retention time) and concentration (signal strength /peak height).

So, in a typical sample the peak that appears last (on the right) in the chromatogram represents the monomeric protein, the peak to the left of the monomer may represent dimers, and subsequent peaks to the left of this (earlier elution time) could be larger oligomers. Peaks succeeding the monomer peak (to its right) could be the result of fragmented protein, or high concentrations of buffer components. Glycine, for example, can be detected by absorbance at 220nm. *Figure 3.7* shows an SEC chromatogram for aggregated IgG; here, peaks due to fragment material are present.

SEC is an excellent method for determining rates of monomer loss by comparison of monomer peak areas in control samples to those in aggregated samples. Dimer formation can also be reliably monitored, though care should be taken when assigning molecular weight to suspected dimer species, as the hydrodynamic radius of a dimer is not usually directly proportional to that of a monomer (Cromwell et al., 2006).

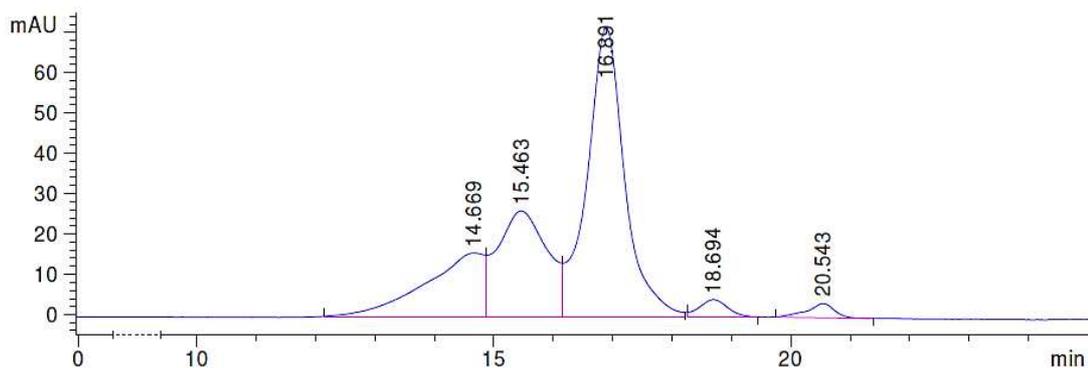


Figure 3.7 An SEC chromatogram. This chromatogram was generated using a Tosoh Tsk-Gel G4000SWXL column and an aggregated IgG. The mobile phase flow rate was 0.65mL/min due to pressure limitations. The x-axis is time (min) and the y-axis is absorbance at 220nm (mAU). Rotated values above peaks represent elution time at maximum peak height. From left to right: two overlapping peaks likely representing small oligomer, 14.7 min, and dimer, 15.5 min, species; monomer peak at 16.9 min; large and small IgG fragments at 18.7 min and 20.5 min, respectively. The left-most oligomer peak appears to contain approximately three unresolved sub-peaks tailing out to the left.

3.4.2.2 *Sample preparation and operating ranges*

A clear advantage of SEC is that the quantity of sample required is small – 10s of micrograms or less per injection – and minimising the protein load maximises resolution (den Engelsman et al. 2011, Hagel 2011). However, dilution of samples for injection may cause weakly associated aggregates to dissociate, or new aggregates to form due to the change in conditions (Arakawa et al. 2010, den Engelsman et al. 2011, Philo 2006). Time taken for SEC analyses using high performance liquid chromatography (HPLC) instruments is moderate, with single runs lasting for between 5 and 20 minutes. However, a sequence of many samples could take several hours to run, so non-reactive conditions would need to be ensured.

SEC analyses are limited to fairly small oligomers ($\leq 50\text{nm}$); larger species tend to be “filtered out” by the column (den Engelsman et al. 2011). Such species may include subvisible aggregates that are not removed by centrifugation. Samples are usually centrifuged at

12000rpm, or passed through a 0.2 μ m pore filter membrane, prior to injection onto the column in order to remove very large particles that could damage the column or HPLC system. Remaining smaller particles may be too large to pass through the column, or they may adsorb to the resin and seemingly disappear from the analysis (Arakawa et al. 2010, den Engelsman et al. 2011). High molecular weight (HMW) species that are able to traverse the column may not produce well resolved peaks; a single peak to the left of the chromatogram could be the product of oligomers of a range of sizes (Arakawa et al. 2010). Purification of HMW species and re-injection at a lower concentration could help to resolve peaks.

3.4.2.3 Preventing column-protein interactions

A common problem in SEC is non-specific protein adsorption; hydrophobic and charge based interactions are known to occur between resin and sample molecules. Interactions may range in strength from irreversible to rapidly reversible. It should be noted that larger particles like aggregates have a greater propensity for adsorption to resins than monomers simply because they have more potential contact sites, in other words a greater surface area (Arakawa et al. 2010). Resin-sample interactions can result in misleading chromatograms. For example, when reversible protein-resin adsorption phenomena take place, the monomer peak can appear normal apart from a longer retention time. Longer retention time is meant to indicate smaller molecules, so underestimation of monomer size could occur, or the inaccuracy could simply go unnoticed (Arakawa et al. 2010).

Understanding of the system can be improved by generation of a calibration curve using samples of known molecular weight and concentration. Molecular weight markers must be of very similar shape to the sample protein because hydrodynamic radius is the true parameter of separation (den Engelsman et al. 2011, Hagel 2011). Of course, it is important to ensure that calibration molecules do not interact with the column. It is often possible to prevent such interactions by mobile phase buffer optimisation; moderate to high salt concentration buffers tend to work well, though too much salt can have adverse effects due to “salting out” of the

protein (Arakawa et al. 2010). Similarly, mobile phase buffers must be optimised for experimental runs.

Typically, if the dissociation constant for resin-sample interaction is low protein will not elute within the expected SEC run time. Lowered protein recovery is observed as reduction in total peak area on a chromatogram, compared to controls.

One study found that monomer recovery as well as aggregate recovery in SEC was significantly diminished after stressing of a sample by pH reduction. Addition of 0.2M arginine to the mobile phase buffer was found to prevent resin-sample interactions, revealing that 62% of aggregate species and 18% of monomers in the samples had previously been lost by adsorption to the SEC column matrix (Arakawa et al. 2010, Ejima et al. 2005). Use of arginine as a mobile phase additive for the prevention of resin-protein interactions was found to be highly effective and necessary for two different types of SEC column, one silica-based and one polymer-based (Ejima et al. 2005). This demonstrates the importance of mobile phase optimisation in SEC.

On the flip side, arginine has been shown to suppress aggregation. It is possible then that a portion of the additional aggregated and monomeric species recovered in the presence of 0.2M arginine were actually dissociation products of much larger aggregates that would otherwise have been too large to traverse the column. This theory relies on the assumption that arginine's ability to suppress aggregation is synonymous with an ability to reverse it. Organic solvents present similar benefits and issues: they can be effective additives for the prevention of hydrophobic sample-resin interactions, but may induce dissociation of aggregates held by hydrophobic association (Arakawa et al. 2010).

In order to prevent generation of erroneous data, it is prudent to carry out characterisation assays that will complement SEC analyses –suppositions based on SEC data alone will either be supported by extended characterisation data, or contradicted by it. These “extended characterisation” assays are usually less robust than SEC-based assays, yet they can increase understanding of a given system, helping with further optimisation of SEC.

3.4.3 Dynamic light scattering (DLS)

Dynamic light scattering is a qualitative method used to estimate the hydrodynamic radius of particles. Measurement is based on fluctuations in scattered light intensity due to Brownian motion (den Engelsman et al. 2011). A laser beam is shone into the sample and its light is scattered by particles in the sample. When particles move (under Brownian motion), the intensity of the light they scatter changes; this is detected by the instrument which subsequently measures the rate of particle movement over time. The rate of Brownian motion is inversely proportional to particle size. This method has been cited as effective in detecting particles as small as 1nm and as large as 10 μ m. Scattering intensity is proportional to the product of particle weight concentration and molecular mass. Thus large particles like aggregates produce high-intensity light scattering (Philo 2006). Obtaining quantitative data using DLS is not possible and resolution is not very high (den Engelsman et al. 2011, Mahler et al. 2009). According to Mahler et al. (2009) size differences of five-fold or greater can be identified. Philo (2006) states that a two-fold difference in particle radii, and an eight-fold difference in molecular weight are required for distinction between particles.

The volume of sample required depends on the exact instrument being used; modern instruments may require only a few microliters of sample, but older models need millilitres of sample. Fortunately, DLS is a non-destructive technique so samples can be analysed further after DLS, presuming they are stable over time. This could be particularly useful if sample variability were a problem. By the same token, buffer conditions do not need to be changed for DLS analysis; it is a non-perturbing method (den Engelsman et al. 2011).

In theory DLS instruments can generate size distribution data from polydisperse samples, producing resolved peaks representing particles with differing hydrodynamic radii, because different particle populations have unique rates of Brownian motion and scatter light with varying intensity. However, in practice polydisperse samples produce less than excellent data and better resolution can be achieved by using mono-disperse samples (Mahler et al. 2009, Philo 2006). This is a notable limitation of DLS as it somewhat negates the advantage of being

able to use “un-altered” samples –significant perturbation may occur during generation of mono-disperse samples from polydisperse ones.

Finally, it should be noted that the technique of DLS is sensitive to large contaminants like dust, and to bubbles (Mahler et al. 2009).

3.4.4 Differential scanning calorimetry

Differential scanning calorimetry (DSC) is a method for determining the thermal stability of proteins. The technique is routinely used to find T_m values for individual protein domains and identify phase transitions. DSC was used by Hari et al. (2010) to investigate thermal transitions in a range of IgG1 and IgG2 antibodies incubated at pH 3.5. The method was successful in identifying lower C_H2 melting temperatures in IgG2 antibodies than in IgG1. This was fitting with SEC data that found IgG1 to be less susceptible to aggregation at pH 3.5 than IgG2. See *Figure 3.8*. It was also found that for aglyco-IgG1, Fab and C_H3 transitions were shifted to lower temperatures (*Fig. 3.8*), possibly due to the already de-stabilised C_H2 domain (destabilised due to lack of N-linked glycan group).

DSC is a useful tool for probing inherent structural integrity of proteins and could help to determine rates of unfolding under other stress conditions at ambient temperature. One disadvantage of the method is that phase transition peaks in thermograms can overlap, making thermodynamic analysis of individual transitions more difficult (Vermeer and Norde 2000).

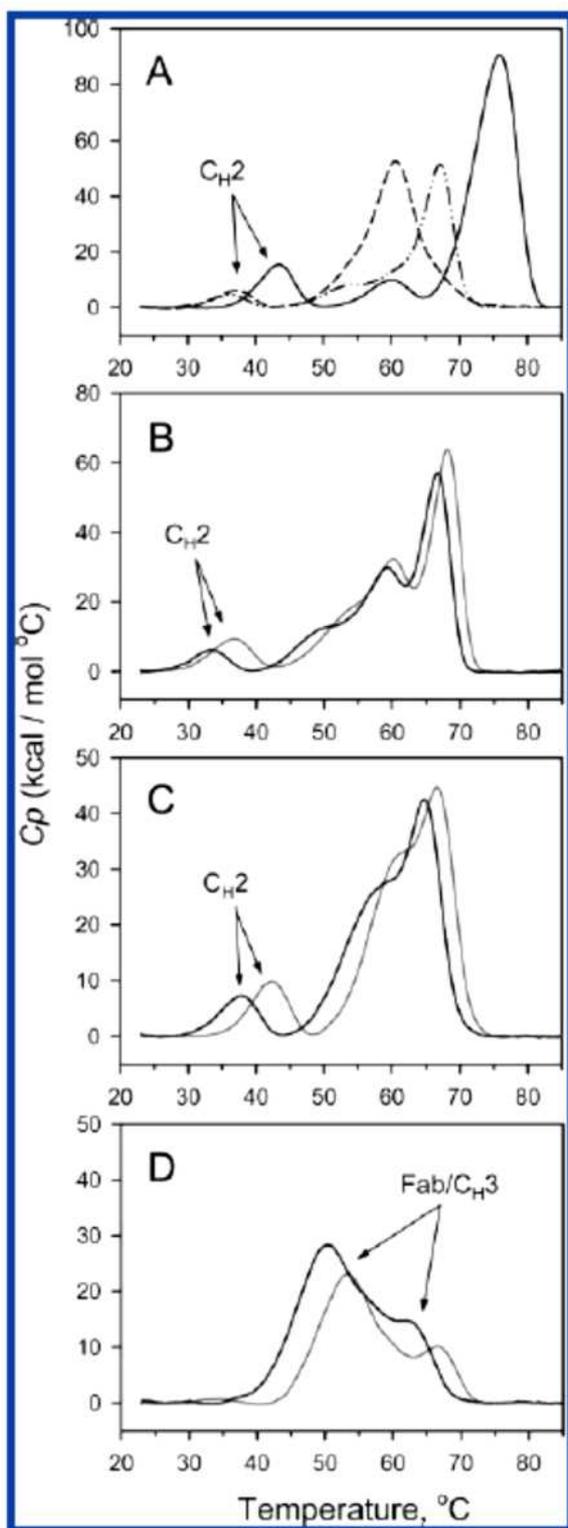


Figure 3.8 DSC thermograms identifying phase transitions for different domains of (A) IgG1-A (solid line), IgG2-A (dashed line) and IgG2-B (dashed and dotted line) at pH 3.5; (B) IgG2-G, (C) IgG2-F and (D) aglyco-IgG-1 in the presence of either acetic acid (black line) or citric acid (grey line), pH 3.5.

Figure reprinted with permission from Hari et al. (2010). Copyright © 2010, American Chemical Society.

3.4.5 Circular dichroism

CD spectroscopy can be used to study protein secondary and tertiary structure and detect protein conformational changes (Tetin et al. 2003).

CD spectroscopy works by detecting differences in absorbance of left- and right-circularly polarised light. In optically active molecules (i.e. chiral molecules like amino acids) the amount of left-handed circularly polarised light and right-handed circularly polarised light absorbed is not equal. This results in rotation of the plane of polarized light which is detected by the instrument. Spectra produced by CD instruments are plots of differential molar extinction coefficient against wavelength (Drake 1994, Pradyot 2004).

The part of a molecule that absorbs light is called the chromophore. Regular protein secondary structure results in repetitive, specific, asymmetric orientation of chiral residues and so specific chromophore-chromophore orientation (Drake 1994). Chromophores that follow similar orientation patterns to one another will all contribute to a unique CD signal. Due to this, different types of secondary structure can be identified.

Different types of chromophore absorb light at different wavelengths, so full CD spectra for proteins are generated by measurement within three different wavelength ranges. Secondary structure characteristics are measured in the far-UV region (170-240nm) (as described, due to peptide-bonded backbone orientation); di-sulphide bonds and tryptophan residues in the mid-UV region (220-235nm); and aromatic side chain interactions relating to tertiary structure in the near UV region (240-320nm) (Drake 1994, Tetin et al. 2003). Different types of secondary structure, e.g. beta sheet, alpha helix, beta turn, each produce a characteristic CD signal in the far-UV region. According to Tetin et al. (2003) all aromatic residues can rotate light extensively in the far-UV, as well as in the near-UV, region so affect the shape of far-UV spectra.

CD spectra immediately confirm whether ordered secondary structure is present in the protein because defined signals are produced only when ordered structure is present. CD spectroscopy is a useful method for determining relative proportions of various types of secondary structure and has been cited as a reliable technique (Drake 1994, Tetin et al. 2003).

According to Vermeer and Norde (2000) CD signals are not affected by solution conditions, this is an obvious advantage for reasons described previously.

A very simple method of exploiting CD spectra is by overlaying control spectra with experimental spectra. One can easily see whether two spectra are identical or not; if they are identical then it is safe to say that the experimental conditions have not induced any changes in protein secondary or tertiary structure. A more complex task is the relative quantification of different types of secondary structure from CD spectra. A combination of comparison to reference spectra and use of mathematical algorithms can be used for this purpose (Tetin et al. 2003). Once achieved, this can be extremely useful information to have because alterations and shifts in secondary structure content can be strong indicators of aggregation propensity. For example, an increase in beta-sheet content could precede irreversible aggregation.

Tetin et al. (2003) compared data obtained from CD spectra analysis with X-ray crystal structures of three different Fab fragments in order to gauge the accuracy of CD spectra analysis. Fab CD spectra in the far-UV region were typical of beta-sheet dominated globulins, though there were also marked differences between the three Fabs. Tetin et al. (2003) attributed the differences to interference from aromatic side chains in the far-UV region. Indeed, spectra in the near-UV region varied significantly between the three Fabs, indicating diverse tertiary arrangements due to varied side-chain interactions. On the whole, CD data was highly comparable to X-ray data in terms of the number of units of each type of secondary structure and their average lengths (Tetin et al. 2003). CD data was more accurate for beta strands than for helical structures, but this could be due to the fact that immunoglobulins contain little regular helical secondary structure.

CD is useful for studying protein-ligand binding because it is quantitative. Changes in CD spectra on binding of protein to a ligand will be directly proportional to the amount of the protein that is affected by either direct (surface to surface) or indirect (resultant distal conformational change) involvement in the binding mechanism.

3.4.6 Further methods

Various other analytical methods that may be used to study protein aggregates or adsorbed proteins are displayed in *Table 3.3*. These methods range from very well-known bench techniques to highly specialised methods.

Table 3.3 Further methods for characterising samples and proteins.

Method	Measurement basis	Result	Ease of Implementation
SDS-PAGE	Proteins coated in SDS, charged relative to surface size, travel through gel towards +ve electrode	Proteins separated on basis on MW Quantitative data	Easy, fast, reliable
Native-PAGE	Non-denatured proteins run on gel in high pH buffer, -ve charge on proteins relative to isoelectiric point	Proteins separated in native state on basis of MW and isoelectric point.	Easy to implement, can be difficult to interpret. MW not quantitative.
Hydrogen exchange mass spectrometry	Protein incubated with excess of D ₂ O, H→D exchange rate is greater in more exposed H. MW _D > MW _H .	Maps solvent exposed regions of protein, can track unfolding.	Laborious
Fluorescence spectroscopy	Fluorescent labelling detected	Track unfolding by increase in fluorescence	Easy
Neutron Reflectivity	Protein adsorbed to solid-liquid interface, neutron beam reflected at interface; reflected intensity depends on structure and composition of adsorbed material	Obtain detailed structural information on adsorbed material in chosen buffer system.	Intensive and expensive, but high quality and quantity data obtained.
Small angle neutron scattering	Neutron diffraction method; flexible sample environment.	Obtain native structure of protein and identify HMW species.	Intensive and expensive, but high precision data.

3.5 Conclusion

Aggregation is a highly undesirable route of protein degradation that can occur as a result of intrinsic and extrinsic factors. Aggregation mechanisms are not well defined and a range of different models attempting to describe them exist. Antibodies are fairly delicate molecules often susceptible to destabilisation under certain conditions; such conditions may be encountered during typical bioprocesses. Protein A chromatography, an ostensibly indispensable method of mAb purification, uses conditions that are known to cause aggregation in some antibodies, low pH elution. Studies have shown that the combination of low pH and protein desorption may exacerbate mAb aggregation. Understanding the interaction between mAb, namely IgG, and Protein A provides some clues as to the role of adsorption and desorption in inducing aggregation; the interaction is largely hydrophobic and requires a conformational change in the antibody. Chromatography parameters like load material concentration, and resin characteristics like dynamic binding capacity, are expected to have an effect on aggregation phenomena, though this is not well-documented in the literature. As aggregates range in size from 10s of nanometers to a few millimetres a broad range of techniques are needed to detect and analyse them. Aggregate assessment often comprises use of orthogonal methods to understand their various characteristics. Simple quantitative methods like SEC can be of great use in determining rates of aggregate formation, while more complex biophysical methods like CD can help to shed light on protein structural changes that may lead to aggregation.

4 Materials and Methods

4.1 Materials

The IgG molecule used for all experiments was kindly donated by UCB Celltech (UCB Celltech, Slough, UK). It is a purified IgG4 kappa antibody; the hinge region is not mutated. The IgG was formulated at 17.8 mg/mL in 270mM glycine, 1% maltose, pH 5.0. This IgG4 kappa has an isoelectric point between 6.85 and 8.15.

Deuterium oxide (99.9 atom % D) was from Sigma Aldrich. Silicon wafers for reflectivity at the NIST Center for Neutron Research (NCNR) were from El-Cat (New Jersey, USA); other NCNR reflectivity cell components were custom made for laminar flow reflectometry.

Silicon wafers and solid-liquid flow cells (SLFCs) for reflectivity at the ISIS Neutron and Muon Source were purpose built. 3-aminopropyltriethoxysilane (APTS) and sulfosuccinimidyl 6-(3'-(2-pyridyldithio)propionamido) hexanoate (sulfo-LC-SPDP) were from Thermo Scientific Pierce (Thermo Fisher Scientific, Leicestershire, UK).

Recombinant staphylococcal protein A was from Repligen (Lund, Sweden).

4.2 Equipment

Size exclusion chromatography (SEC) was carried out using an Agilent 1100 HPLC system with ChemStation software (Agilent Technologies UK Ltd, Berkshire, UK) with a TSK Gel 7mm × 300mm G3000SWXL column (Tosoh Bioscience LLC, Montgomeryville, USA). Affinity chromatography was carried out using an AKTA Avant 25 liquid chromatography system with Unicorn 6 software (GE Healthcare, Uppsala, Sweden), with a 1mL HiTrap MabSelect Xtra column (GE Healthcare UK Ltd, Buckinghamshire, UK) attached.

Attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR) was done using a Thermo Nicolet Nexus instrument (Madison, WI, USA) that was fitted with a monolayer/grazing angle accessory (Specac 19650 series, Kent, UK), a mercury cadmium

telluride detector, and an air dryer to purge water vapour and carbon dioxide from the instrument (Clifton et al. 2012).

Reflectivity measurements at the NCNR (Gaithersburg, MA, USA) were performed on the NG7 horizontal neutron reflectometer.

Reflectivity measurements at the ISIS Neutron and Muon Source (Rutherford Appleton Laboratory, Harwell Science and Innovation Campus, Oxfordshire, UK) were performed on the INTER time-of-flight reflectometer. Buffer flow through the SLFCs was achieved using an L7100 HPLC pump from Merck-Hitachi (Kent, UK).

4.3 Methods

4.3.1 IgG4 aggregation rate in solution

Solution experiments were carried out initially to determine an appropriate pH range to work in with this IgG4 molecule, and timescales for aggregation. The IgG was mixed with various solutions of glycine-HCl to produce solutions of IgG at concentrations ranging from 0.9 mg/mL to 4.5 mg/mL, pH conditions ranging from pH 2.78 to pH 3.03, and a buffer concentration of 0.15M glycine-HCl. Proportions of reagents needed to give the desired pH were calculated using the Henderson-Hasselbalch equation with an adjustment for fixed ionic strength (Beynon et al. 1996). PH was measured using a standard laboratory pH probe; pH adjustment was not implemented and reported pH values represent the *measured* values. Concentrations recorded are calculated values based on the initial concentration reported for the drug substance and the subsequent level of dilution in low pH buffer. Samples were taken from each solution at various time points and neutralised with a 10% volume of 0.8M tris-HCl, pH 8.45. Neutralised samples were diluted in SEC mobile phase buffer. Samples were stored at 4°C overnight before analysis by SEC. All conditions were run in duplicate. Controls consisted of IgG4 diluted to various concentrations in SEC buffer and run on the SEC column. The order of analysis of all samples

(including controls) by SEC was randomised. Peaks were detected by absorbance at 280nm. The following equation was used to determine percentage monomer recovered:

$$R = (A_S/A_C) \times 100 \quad \text{Equation 1}$$

Where R is the percentage monomer recovered, and A_S and A_C are the sample and control monomer peak areas (mAU²), respectively. Percentage monomer recovered was plotted against incubation time to determine rate of monomer loss.

4.3.2 IgG4 aggregation rate after chromatography

A 1mL HiTrap MabSelect Xtra column was equilibrated with 3 column volumes (CV) 0.02M sodium phosphate, pH 6.7, and 2mL of 5.6 mg/mL IgG4 was loaded onto the column using a sample loop. The column was washed with 3CV (sufficient to return baseline to zero) 0.02M sodium phosphate, pH 6.7, and step elution of the protein was effected with 5CV 0.15M glycine-HCl, pH 2.93. The column effluent was monitored by absorbance at 280nm and 0.5mL fractions were collected throughout elution. The delay volume between the detector and the fraction collector on the AKTA system was 594 μ L. Protein concentration in each fraction was measured offline and showed that fraction location was consistent with the UV trace from the AKTA software. The pH was monitored using the in-line pH meter supplied with the AKTA instrument which was calibrated for the elution pH range (between pH 1.68 and pH 4.00); in-line pH measurement was used only as an indicative measure before offline measurement of individual fractions. The majority of the elution peak was collected into three central fractions. The third fraction, containing the tail end of the peak, was used for subsequent experimentation as this fraction was at an appropriate pH to induce aggregation. This will be referred to as *fraction 3*. Fraction 3 was mixed to ensure homogeneity and incubated at the elution pH. The pH of all elution fractions was verified at the end of the incubation using a PHR 146B Micro Combination pH Electrode (Lazar Research Laboratories, Los Angeles, CA, U.S.A.). Chromatography “dry runs” were also carried out in the absence of IgG and the pH of the elution fractions were measured individually. Fraction 3 was found to be at the pH of the

original elution buffer, +/- 0.03 pH units. Samples of fraction 3 were neutralised at various time points, then diluted in SEC mobile phase buffer and stored overnight before SEC analysis as in *section 4.3.1*. The protein concentration in these samples was determined by measuring the absorbance of each sample at 280nm using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, DE, U.S.A.), and using Beer's law and an extinction coefficient $E_{1cm}^{0.1\%} = 1.61$ (experimentally determined by UCB Celltech, Slough). Control samples, for calculation of percentage monomer loss, consisted of the column feed (IgG diluted in equilibration buffer), the concentration of which was also verified by UV absorbance measurement. Controls for "effect of column alone" consisted of samples from the leading side of the elution peak; on exiting the protein A column, these samples were not at a pH low enough to induce aggregation, this was confirmed by SEC analysis.

4.3.3 Size exclusion chromatography

0.1M sodium phosphate, 276mM NaCl, pH 6.3 was used as the mobile phase buffer for SEC. The flow rate was 1 mL/min and the injection volume was 20uL. Peaks were detected by absorbance at 280nm. Before injection onto the SEC column all samples were spun at 13000rpm to remove any precipitated protein material.

4.3.4 Aggregation rates data analysis

Statistical analysis and data fitting were carried out using OriginLab® OriginPro 9.0 software.

4.3.5 Neutron reflectivity experiments at the NIST Center for Neutron Research

4.3.5.1 Sample preparation

IgG4 was prepared at in H₂O-based 0.03M sodium citrate, pH 4.03 and pH 6.21, and D₂O-based 0.03M sodium citrate, pH 4.10 and pH 6.24. Buffer exchange from the IgG4-κ formulation buffer into sodium citrate buffers was done using Amicon Ultra-15 centrifugal filter units with a 10kDa nominal molecular weight limit (NMWL) (Merck Millipore, Darmstadt, Germany). Buffer exchange was carried out in three stages resulting in a final dilution factor of at least 3000 from the original buffer. The IgG concentration in the samples was determined by

measuring the absorbance at 280nm using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, DE, U.S.A.), and using Beer's law and an extinction coefficient $E_{1\text{cm}}^{0.1\%} = 1.61$ (experimentally determined by UCB Celltech, Slough). All samples were adjusted to an IgG4 concentration of 9.5 mg/mL. Samples were shipped to Gaithersburg, USA, at 4°C one week before commencement of reflectivity experiments. Before injection into the reflectivity cell, samples were filtered through a 0.2µm membrane and diluted in their respective buffers to a final IgG concentration of 5.6 mg/mL.

For stability experiments, the same methods were used to buffer exchange IgG4 into H₂O-based 0.03M sodium citrate, pH 6.3, pH 4.3, pH 3.8 and pH 3.2. Stability experiments were carried out in the UK. After buffer exchange, sample concentration was adjusted to 10mg/mL and samples were incubated either at room temperature or at 4°C. Controls were IgG4-κ diluted to 10mg/mL in the formulation buffer (0.27M glycine, 1% maltose, pH 5.0) and incubated either at room temperature or at 4°C. Samples were analysed at various time points by SEC using the method described in *section 4.3.3*. Sample pH was not adjusted before injection onto the column, and the injection volume was either 5µL or 10µL.

4.3.5.2 Reflectivity cell

The reflectivity cell was prepared using the method described by Perevozchikova *et al.* (2015) in the section *Materials and Methods – NR Measurements – Substrate Preparation*. In brief: circular silicon wafers approximately 76mm in diameter and 5000µm thick were cleaned using a Piranha etch solution (3:1 concentrated sulphuric acid: 30% hydrogen peroxide) and rinsed with ultrapure water before assembly into the SLFCs. The assembled reflectivity cell contained a chamber, created by a Viton spacer, with a volume of approximately 1.5mL and a height of 100µm (Perevozchikova et al. 2015).

4.3.5.3 Reflectivity experiments

Reflectivity measurements were performed on the NG7 horizontal neutron reflectometer at the NIST Center for Neutron Research, Gaithersburg, MA, USA. The reflectivity cell was placed on the sample stage and connected to a syringe pump with valved tubing, meaning sample changes

could be made without introducing air into the reflectivity cell. In total two separate reflectivity cells were used for the experiments. At the beginning of experiments each cell was aligned and slit scans taken.

Measurements were taken at Q_z values ranging from 0.008\AA^{-1} to 0.250\AA^{-1} . For each sample configuration neutron data collection took place over a minimum of 6h, this provided sufficient counting statistics to resolve signal over background counts (Perevozchikova et al. 2015); counting times were weighted towards the high Q region ($0.025\text{-}0.25\text{\AA}^{-1}$). For each sample configuration background scans were programmed to run periodically between sample analyses; background scans ran over a short time period (20 min each) for a selection of Q values across the full Q range. All scans were run in duplicate and in some cases triplicate (e.g. overnight runs). The reflectivity instrument counts the number of neutrons that hit the detector for a given Q value. The counts are normalised based on an intensity of 1 for total internal reflection of the neutron beam. For each scan, error bars equal the square root of the number of counts before normalisation.

The cell was washed with 7.5mL (5 cell volumes) buffer and neutron measurements were made for the bare silicon wafer for two solution phase contrasts, D_2O and H_2O . In the first SLFC, after washing the cell with D_2O -based buffer, 3mL (2 cell volumes) IgG4 in 0.03M sodium citrate, pH 4.1 was flowed into the reflectivity cell at 6 mL/h. The IgG solution remained in the cell during reflectivity measurements. Reflectivity data in the D_2O solution phase contrast was acquired for the following additional sample configurations: after IgG solution rinsed from cell; IgG solution (fresh) re-introduced into cell; IgG solution rinsed out of cell (again); after another rinse. Reflectivity data in the H_2O contrast was obtained for one IgG adsorption step and after one rinse step.

In the second SLFC reflectivity data was obtained in D_2O and H_2O solution phase contrasts for the following sample configurations: IgG adsorbed in 0.03M sodium citrate, pH 6.2; IgG solution washed out with the same buffer; surface washed with 0.03M sodium citrate, pH 3.7.

4.3.5.4 Data analysis

Initial data processing was done using Reflred software, and NCNR in-house software (Kienzle et al. 2000-2006). Low Q and high Q scans were stitched together and averages taken for replicate scans. Slit and background scans were averaged and subtracted from the stitched sample scans to give the final reflectivity curves.

Model fitting was done using RasCal software (Hughes 2013) which fits Abeles layer models to the data using an optical matrix formalism (Clifton et al. 2015). The details of building and constraining the model are described in *section 6.6.3*.

Fitting error analysis was performed using the *Bootstrap error analysis* function in RasCal (Hughes 2013). This is described in more detail in *section 6.6.3*.

Models were selected based on a combination of lowest chi-squared values and most robust key fitted parameters (results of the bootstrap error analysis).

4.3.6 Experiments at the ISIS Neutron and Muon Source

4.3.6.1 Sample preparation: protein A

Recombinant staphylococcal protein A (rSPA) was dialysed into D₂O-based *PBS-EDTA* buffer which consisted of 0.14M NaCl, 0.003M mM KCl, 0.01M phosphate buffer solution and 0.001M EDTA, pH 7.4. Dialysis was done using a Slide-A-Lyzer™ Dialysis Cassette with a 3.5kDa molecular weight cut off (Thermo Fisher Scientific, Leicestershire, UK) in two stages with a total dilution factor of 100000 for original buffer constituents. rSPA concentration was determined using Beer's law and a percent solution extinction coefficient $\epsilon^{1\%} = 1.65$ (275nm), i.e. molar extinction coefficient = 0.165.

LC-SPDP was attached to rSPA by following the protocol in the product manual from Thermo Fisher Scientific, this is based largely on methods by Carlsson et al. (1978). The linker-modified rSPA was separated from reaction by-products using PD-10 Desalting Columns (GE Healthcare, Buckinghamshire, UK) which contain a Sephadex G-24 medium. The identity of desalting

fractions was confirmed using spectrophotometric methods and reduced SDS-PAGE with SigmaMarker wide range molecular weight markers (Sigma-Aldrich, Dorset, UK).

Modified rSPA was stored at 4°C for up to 5 days.

The molar ratio of protein to linker was determined using the method in the product manual, which is based on a method by Carlsson et al. (1978). The method estimates the quantity of pyridine-2-thione produced when the sample is reduced by measuring the change in absorbance at 343nm. The following equation was used:

$$\text{Moles of SPDP per mole of protein} = \frac{\Delta A}{8080} \times \frac{\text{Protein molecular weight}}{[\text{protein}]}$$

Equation 2

where the value 8080 is the extinction coefficient of pyridine-2-thione at 343nm ($M^{-1}cm^{-1}$) and $[\text{protein}]$ is in units of mg/mL.

4.3.6.2 Sample preparation: BSA and PEG

A 1% w/v bovine serum albumin solution was prepared by dissolving BSA (Bovine Serum Albumin Fraction V, Sigma-Aldrich (Dorset, UK)) in D₂O-based PBS-EDTA buffer (see section 4.3.6.1). A 1% w/v PEG₆₀₀₀ solution was prepared by dissolving PEG₆₀₀₀ (VWR Chemicals, Leicestershire) in PBS-EDTA buffer.

4.3.6.3 Sample preparation: IgG

IgG4-κ was dialysed into D₂O-based 0.03M sodium phosphate, pH 6.7 using Slide-A-Lyzer™ Dialysis Cassettes with a 20kDa molecular weight cut off (Thermo Fisher Scientific, Leicestershire, UK) in two stages with a total dilution factor of 100000 for original buffer constituents. Dialysed material was prepared at 18 mg/mL IgG4 and stored at 4°C for up to 5 days. The dialysed IgG was diluted to 1mg/mL in the same buffer for FTIR and reflectivity experiments.

4.3.6.4 ATR-FTIR experiments

The silicon crystal was cleaned by ozone cleaning, creating a highly hydrophilic surface. The crystal was amino-silylated by immersion in a 2% solution of APTS in extra dry acetonitrile* for 30 seconds. The crystal was rinsed with extra dry acetonitrile* and allowed to dry.

*the reagent was labelled *extra dry acetone* (Fisher, Leicestershire), which was the intended solvent. However, the product was re-called by Fisher several months later after the bottles were found to contain acetonitrile (extra dry), not acetone. This was not expected to have a considerable adverse effect on the amino-silylation reaction.

The amino-silylated crystal was assembled in a solid-liquid flow cell custom made to fit the FTIR instrument. All subsequent surface modifications were carried out in flow-through mode and monitored by FTIR. Flow was controlled by a peristaltic pump at a flow rate of approximately 1 mL/min. For reagent additions at least 5 flow-cell volumes were used and for rinse steps at least 8 cell volumes (the cell volume, including tubing, was approximately 1.5mL). All FTIR experiments were carried out using D₂O-based buffers.

Data collection was accomplished using Omnic software (Thermo Scientific). Spectra were recorded at a resolution of 4cm⁻¹ across a wavenumber range of 4000-1480 cm⁻¹. For each measurement 128 interferograms were collected, averaged and converted to absorbance units. Background scans were collected at the beginning of each experiment, when the buffer was changed and at other relevant points between experimental stages. The background was automatically subtracted from the spectrum during data collection. In most cases the most recent background file was used for subtraction, but in certain cases earlier background files were selected to reflect buffer changes and to reveal peaks that had previously been blanked out. At each stage of surface modification (including rinses) spectra were collected continuously for the first 15 minutes, and once every 15 minutes thereafter. The instrument was continuously purged of water vapour and carbon dioxide. Each time the instrument was switched on and once the flow-cell was filled with buffer a background was taken and spectra collected immediately –

these spectra contained increasingly negative peaks in regions 3950-3500 cm^{-1} and 2000-1480 cm^{-1} due to H_2O and CO_2 purging. These spectra were saved as *subtraction spectra*.

Modified rSPA was cross-linked to the amino-silylated surface using methods described by Carlsson et al. (1978). The cross-linking reaction was allowed to proceed overnight. After rSPA cross-linking the surface was incubated with 1% BSA solution to block any reactive surface patches. The buffer was changed from PBS-EDTA to 0.03M sodium phosphate, and a new background taken before addition of 1mg/mL IgG4. The IgG solution was rinsed out after 10 minutes. IgG was eluted with 0.1M Glycine-HCl, pH 3.4. Variations of this experimental procedure were executed using different amino-silylated silicon crystals; these are described in *section 7.4*.

Spectra were analysed using Omnic software (Thermo Scientific). *Subtraction spectra* were used to manually subtract $\text{H}_2\text{O}/\text{CO}_2$ peaks from each spectrum – in most cases subtraction was minimal; only long sample incubation times resulted in significant changes in $\text{H}_2\text{O}/\text{CO}_2$ content. No further processing was performed. Peaks were identified using the *peak detection* tool at a sensitivity of 50%, the threshold was varied depending on signal to noise ratio. Peaks were integrated using the *peak area* tool which automatically performs a baseline correction.

4.3.6.5 Reflectivity experiments

Silicon wafers (50 x 80 x 20mm; 20mm = thickness) were cleaned using the method described in *section 4.3.5.2* and amino-silylated using the method described in *section 4.3.6.4*. Wafers were wet assembled into solid-liquid flow cells (SLFCs) and fixed to the reflectometer sample stage. Three separate SLFCs were used.

Reflectivity measurements were performed on the white beam INTER time-of-flight reflectometer at the ISIS Neutron and Muon Source (Rutherford Appleton Laboratory, Harwell Science and Innovation Campus, Oxfordshire, UK). Measurements were taken at Q_z values ranging from 0.01\AA^{-1} to 0.20\AA^{-1} . Glancing angles of 0.7° and 2.3° were used to acquire data in the Q ranges $0.01\text{-}0.1\text{\AA}^{-1}$ and $0.033\text{-}0.20\text{\AA}^{-1}$, respectively. Neutron counts are normalised based

on an intensity of 1 for total internal reflection of the neutron beam. For each scan, error bars equal the square root of the number of counts before normalisation.

For each SLFC, the inlet was connected to an HPLC pump for flowing buffer into the cell; the pump was operated remotely. Reagents were injected manually. Each sample cell was analysed before any surface modifications were made (apart from amino-silylation) using at least one solution phase contrast, the D₂O contrast always being the first choice. Reflectivity data was collected during and after various stages of surface modification such as cross-linker attachment, rSPA cross-linking, surface blocking, IgG adsorption, IgG elution and surface regeneration. Cross-linking was done under conditions described in *section 4.3.6.4*, but the buffer was changed to 0.03M sodium phosphate, pH 6.7 (the buffer used during IgG adsorption) before neutron measurements were made. Buffer conditions were not expected to influence neutron measurements, but would likely affect protein conformation and orientation. In two of the SLFCs PEG₆₀₀₀ was used as a blocking molecule instead of BSA. For protein-containing sample configurations, data was collected for three solution phase contrasts: D₂O, H₂O and *silicon matched water* (SMW). SMW contains 38% D₂O and 62% H₂O. Buffers were prepared in D₂O and H₂O; the SMW contrast was achieved by running a mixing program on the HPLC pump.

Initial data processing was done using MantidPlot (Taylor et al. 2012). The data was processed similarly to that collected at the NCNR (*section 4.3.5.4*). Background subtraction was not implemented for the raw data and was instead integrated into the model fitting process using RasCal (Hughes 2013). Model fitting and bootstrap error analysis was done as described in *section 6.6.3*. The models were more complex than those for NCNR data; this is discussed in *section 7.5*.

5 Determining the Influence of Elution from Protein A on the Rate of IgG4 Aggregation

5.1 Introduction and Objectives

As described in *Section 2*, the objective of the initial work is to demonstrate the effect of chromatographic separation on protein (product) aggregation behaviour. The focus will be on protein A chromatography; this is a highly pervasive chromatographic step used to purify and concentrate IgG products from freshly clarified cell culture supernatants or homogenates. Staphylococcal protein A binds all IgG molecules of subclasses 1, 2 and 4 (Hober et al. 2007) with high selectivity and minimal interaction with the Fab region (Salvalaglio et al. 2009), the active region of the drug molecule. Product molecules are eluted from protein A resins by lowering the pH; a typical elution buffer is 0.1M sodium citrate, pH 3.3. Low pH is often maintained for a period of time for the purpose of viral inactivation (Ejima et al. 2007, Shukla et al. 2007a). However, for many antibodies, acidic conditions and sudden pH changes can result in aggregation (Cromwell et al. 2006, Ejima et al. 2007, Hari et al. 2010, Hober et al. 2007, Shukla et al. 2007a).

Protein aggregation induced by pH has been the subject of much investigation in bioprocess development, and low pH is typically cited as the cause of product aggregation occurring during or after protein A chromatography (Paul et al. 2014); it is also acknowledged that protein aggregation often occurs more readily at high protein concentrations. Further to this, a more limited pool of evidence suggests that low pH may not be the sole cause of aggregation in protein A chromatography, rather, the adsorption and desorption events themselves may contribute significantly (Gagnon et al. 2015, Shukla et al. 2007a, Shukla et al. 2007b).

A typical model for protein aggregation consists of four stages: reversible destabilisation of native structure or partial unfolding to form the “reactive monomer” (R_M) (Roberts 2007); reversible or irreversible association of R_M yielding a more thermodynamically favourable

“aggregate” state; association of a critical number of R_M to form a nucleus; addition of R_M or small oligomers to the nucleus to form larger amorphous or ordered aggregates. Different theories argue different stages as the rate limiting step (Chi et al. 2003, Roberts 2007). For the first stage, the destabilising effect of low pH on IgG4 was shown by Ejima *et al.* (2007) using differential scanning calorimetry (DSC), a method for assessing thermal unfolding. At pH 6.0, two major endothermic transitions were seen; at pH 3.5 similar transitions were seen at significantly lower temperatures; at pH 2.7 a very different transition profile was seen, with major transitions occurring at relatively low temperatures (Ejima et al. 2007).

A study by Shukla et al. (2007a) attempted to elucidate the collective effects of low pH, protein concentration and chromatographic separation on an Fc-fusion protein. Shukla et al. found that protein A chromatography significantly increased the rate constant for formation of high molecular weight species at low pH. Rate constants were determined graphically using a derivation of a Lumry-Eyring-based kinetic model for monomer loss/aggregate formation. In this instance rate constants were found to be concentration-independent. Different additives included in the chromatography elution buffer significantly altered aggregation rates. Interestingly, in some cases, additives that stabilised proteins at low pH in solution had destabilising effects in chromatography experiments. Urea was an effective additive in reducing on-column and in-solution aggregation at concentrations of 0.5M and 1M, respectively (Shukla et al. 2007a). In other work, arginine was found to prevent protein aggregation on elution from protein A (Shukla et al. 2007b). Recent work by Gagnon et al. (2015) showed that elution from protein A affects the conformation of IgG1; namely, a significantly reduced hydrodynamic radius and changes in secondary structure content were observed. Notably these effects did not occur at low pH in the absence of the elution event.

Protein A binds to the Fc region of IgG about the hinge between the C_{H2} and C_{H3} domains (see *Figure 3.5, section 3.3.2 Protein A*). The precise region of the Fc that binds to protein A can also bind a number of other molecules; thus, it has been termed *consensus binding site* (CBS) (DeLano 2000). The CBS is largely hydrophobic in character, contains relatively few polar

residues and has a high level of solvent accessibility. According to Delano et al. (2000), the CBS undergoes considerable conformational changes when binding to a ligand. In fact, the nature of the change in conformation depends on the ligand, highlighting the flexibility of this region. Though flexibility implies good structural recovery after conformational change, under antagonistic conditions such as low pH there may be greater vulnerability to detrimental levels of structural alteration. X-ray crystallography studies by Deisenhofer (1981) found that C_{H2} domain disorder was greater in an Fc-Fragment B complex (Fragment B being a single protein A domain) than in the unbound Fc, implying potential Fc destabilisation on adsorption to protein A (Deisenhofer 1981, Gagnon et al. 2015). However, this may not have been the case had the Fab fragment been present, due to C_{H2}-Fab contact (Deisenhofer 1981). Conversely, it has been shown that some IgG molecules retain a folded structure at pH conditions as low as pH 2.7 after elution from a protein A column, though it is not identical to the native structure. In this case, IgG aggregation after low pH elution was instead attributed to conformational disturbances occurring during neutralisation of low pH samples (Ejima et al. 2007).

Predictably, the literature indicates great variation in protein aggregation behaviour depending on the protein. Though aggregation in protein A elution pools is well-acknowledged, its mechanism and impact is not well-understood. The aim of this chapter is to clearly present a method showing that the process of protein A chromatography increases the rate of aggregation of an IgG4 molecule at low pH. Experiments are designed to mimic the part of a standard mAb bioprocess where protein A chromatography is followed by incubation at low pH for the purpose of virus inactivation, and subsequent neutralisation of this pool. MabSelect Xtra (GE Healthcare), which consists of agarose particles with an engineered recombinant protein A ligand (see *Figure 3.4, section 3.3.2 Protein A*), will be used as this is a modern industrially relevant resin. Importance is placed on production of reliable kinetic data and strong comparability between data sets where a chromatography step is included, and those done in solution only at low pH.

5.2 Experimental Design

In *Figure 5.1.a* the experimental flow is shown. The terms *on-column* and *column experiments* will be used interchangeably to describe experiments that correspond to that shown in *Fig 5.1.a*, including incubation of the eluted material in *Fraction 3* at the elution pH. The terms *solution experiments*, *in-solution* and *solution-only* will be used interchangeably to refer to experiments in which there was no chromatography step and incubation at low pH was initiated by adjusting the pH manually.

Initial experiments were done in solution under a range of low pH conditions and IgG4 concentrations in order to determine the basic behaviour of the mAb at low pH and identify a useful pH range to work in. Typical conditions and time-courses for protein A chromatography and virus inactivation were taken into account. Time-consumption and reproducibility of experiments was also considered. Additionally, the formulation buffer of the purified mAb contributed to experimental design decisions; as the formulation buffer already contained a high concentration of glycine (0.27M), glycine-HCl buffers were used to lower the pH in solution experiments, and for product elution in column experiments. Glycine-HCl is often used as an affinity chromatography elution buffer, though citrate or acetate buffers are more common in bioprocessing when a slightly higher elution pH is required (Bottomley et al. 1995, Firer 2001, Hahn et al. 2003, Pabst et al. 2014). In the case of this work, using glycine-HCl reduced the number of sample manipulation steps and potential confounding effects this might have on aggregation data. The pH range chosen for experiments covers a larger range than would usually be considered for industrial purification (Ejima et al. 2007, Hober et al. 2007, Shukla et al. 2007a, Shukla and Hinckley 2008, Shukla et al. 2007b), as it is required to plot correlations between aggregation rates and low pH.

Column experiments were carried out as shown in *Figure 5.1*. Incubation at the elution pH was assumed to begin at the first sign of the elution peak (increase in absorbance at 280 nm), denoted t_0 in *Fig. 5.1.a*. It is assumed that the mAb has been exposed to the elution buffer at this stage. The column effluent does not approach pH 3 until the tail end of the peak, i.e. *Fraction 3*,

but marking t_0 earlier eliminates the possibility of exaggerating the column effect on aggregation rate. The IgG4 concentration in *Fraction 3* (*Fig. 5.1.b*) was within the range tested in solution, supporting comparability between in-solution and on-column data.

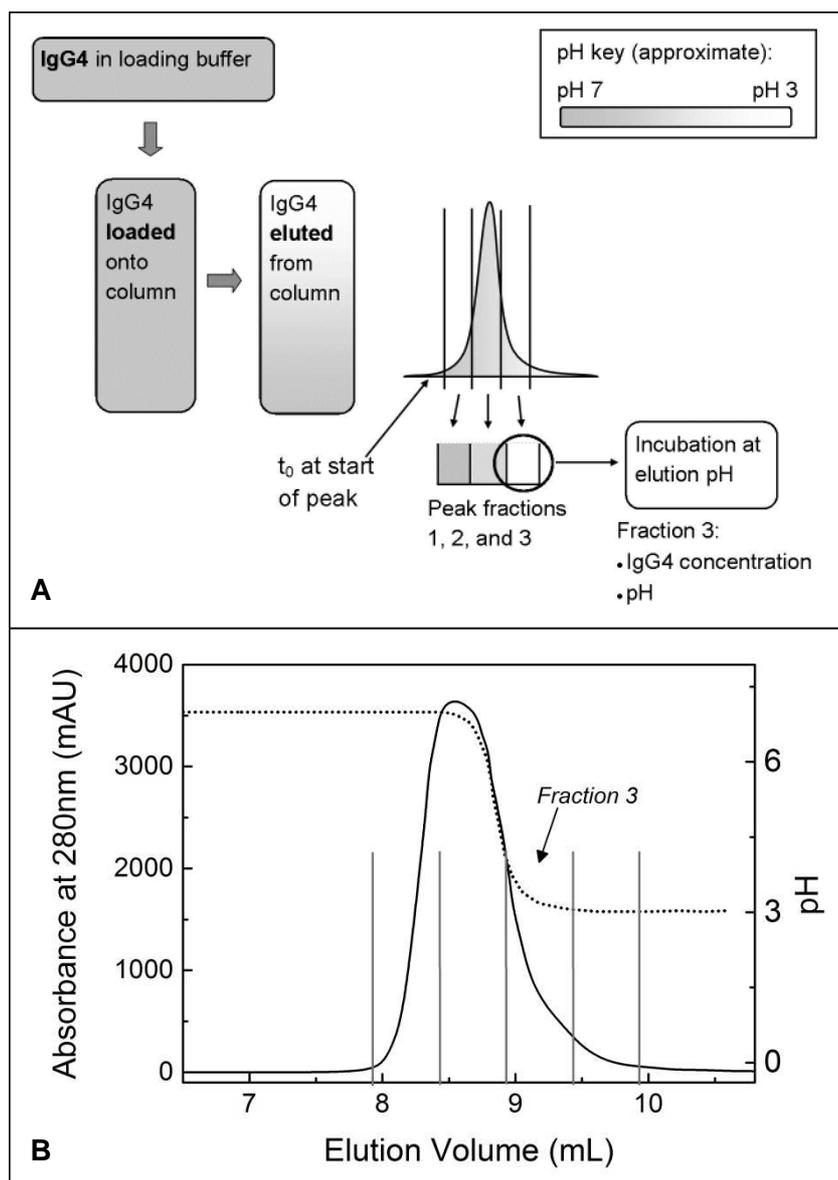


Figure 5.1 Schematic of experimental design where incubation at low pH follows protein A chromatography (a). Incubation is assumed to begin when the elution peak begins, denoted t_0 . Fractions containing approximately the first two thirds of the elution peak were at pH conditions considerably higher than the elution buffer pH. IgG concentration and pH in fraction 3 were the conditions reproduced in solution experiments for comparability with column experiments. A precise representation of the experimental outcome can be seen in b, a chromatogram from a protein A chromatography run. The solid line (left y-axis) represents the IgG elution peak and the dotted line (right y-axis) indicates the pH of column effluent, as measured by the in-line pH probe. Here, flow of elution buffer begins at 7mL Elution Volume. The last two fractions of the elution peak, indicated within vertical lines in b, were at pH conditions low enough to induce aggregation.

Size exclusion chromatography was used to quantify IgG monomer and high molecular weight (HMW) species in the test samples. Monomer loss over time was chosen as the basis for quantification of aggregation rates. Under all sufficiently harsh conditions tested, monomers were partially converted to a range of high molecular weight species over time as can be seen in *Figure 5.2*. A few major observations were made. Firstly, regardless of (perceived) harshness of conditions, with sufficient time, incubation at low pH resulted in an increasing front shoulder to the monomer peak and a small shift of the whole peak in the direction of shorter elution time. This is indicative of a monomer either with increased molecular weight, or with altered physical character that could accelerate its transit through the SEC column (Fekete et al. 2014). Harsher conditions and/or longer incubation times instigated these changes in the monomer peak. The *pronounced* front shoulder on the monomer peak (*Fig. 5.2*, top) may represent an individual population of monomer in a molten globule state (Filipe et al. 2012). Alternatively, it may be due to an oxidised species such as the oxidised tryptophan IgG2 monomer “pre-peak” species characterised by Wong et al. (2012). Another observation was that decrease in monomer peak area was proportional to increase in HMW species peak area, up to a point. After this point, (see top and middle plots *Fig. 5.2*), the peak area for HMW species did not increase and the monomer front shoulder developed further; in the bottom plot (*Fig. 5.2*) the monomer peak continued to decline without either significant development of the monomer front shoulder or an increase in HMW species. These results indicate that aggregates whose molecular weights exceed the exclusion limit of the SEC column may form to varying extents depending on the conditions. Such species may have been retained inside the column or HPLC system, or been removed by centrifugation prior to sample injection onto the column. Further, large species may subsequently “fall apart”, creating a range of species including reversible dimers – such species could contribute to a fronting shoulder on the monomer peak.

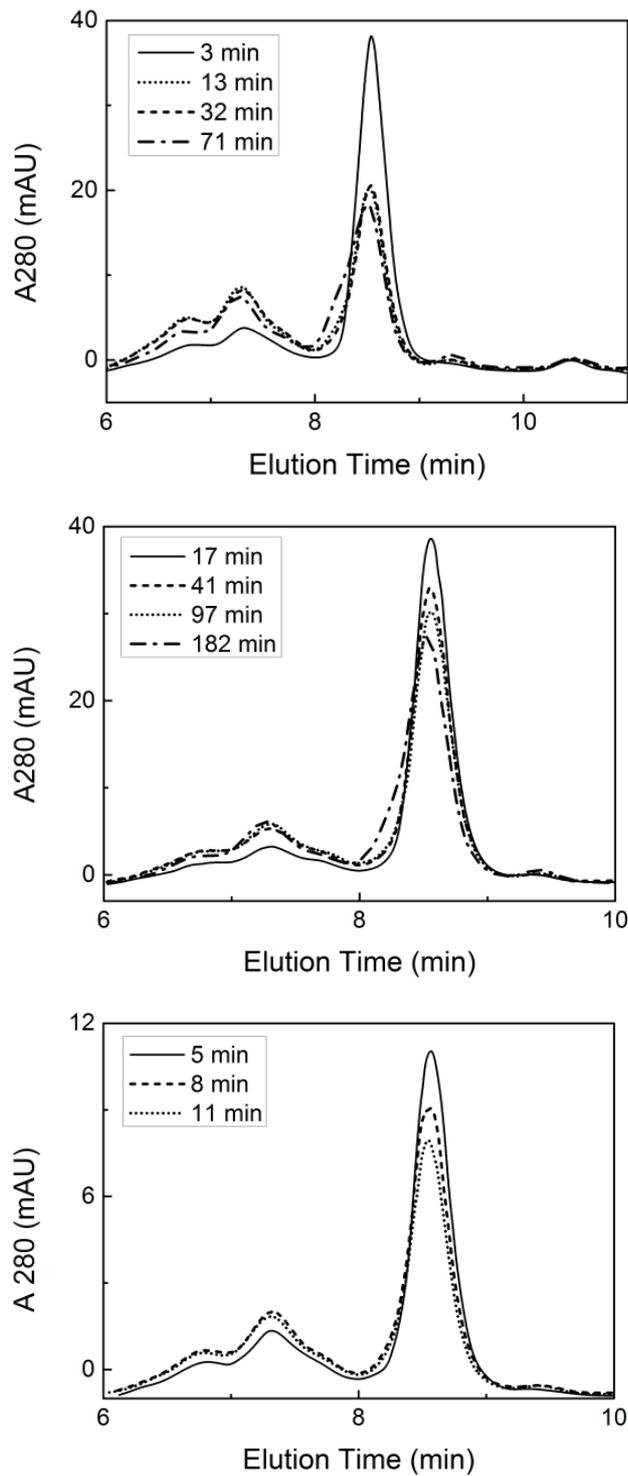


Figure 5.2 SEC chromatograms for IgG4 incubated under a range of low pH conditions. Top: pH 2.78 in solution. Middle: pH 3.05 after elution from column loaded with 12mg IgG. Bottom: pH 2.92 after elution from column loaded with 25mg IgG. Legends show incubation times to the nearest minute.

In separate experiments, aggregate species formed at low pH in solution or after elution from protein A were separated into fractions by SEC. Species eluting from the SEC column at 6-7 min, 7-8 min and 8-9 min were designated *HMW large*, *HMW small* and *monomer*, respectively, and collected in separate fractions. The fractions were concentrated and then analysed by native polyacrylamide gel electrophoresis (native-PAGE); the result of which is shown in *Figure 5.3*. Monomer fractions were not run on the gel shown in *Fig. 5.3*, but the untreated IgG4 monomer was run for comparison to HMW fractions. In native-PAGE, molecular weight markers cannot be used for accurate sample molecular weight determination, because the rate at which native proteins travel through the gel depends on their shape and isoelectric point as well as molecular weight. The markers show that the gel has run normally. Assuming that untreated IgG4 remained in the monomeric state, the monomer can be seen as a double band in lanes 2 and 11. Some streaking below the main bands may be due to fragment species. In lanes 3-6, the dimer and trimer bands are in distinct locations and it can be seen that there was little conversion of HMW species back to monomer. In lanes 7-10 (solution experiments), the same dimer and trimer bands are present. There has clearly been significant conversion back to smaller species and monomer from *HMW large* and *HMW small* fractions. However, these fractions had been stored for a long period (months) before running on native PAGE, while the column experiment fractions were run days after collection.

The results of native-PAGE are in good agreement with the SEC results, in that two major groups of HMW species, as well as IgG monomer, can be identified. The HMW species are also shown to be robust but not indefinitely irreversible.

Overall, it was concluded that reduction in monomer peak area as determined by SEC provided an accurate measure of aggregation rate. In cases of long or harsh incubation, the leading shoulder developing on the monomer peak was counted as normal monomer when plotting the data; the implications of this are discussed in *Section 5.4.1*.

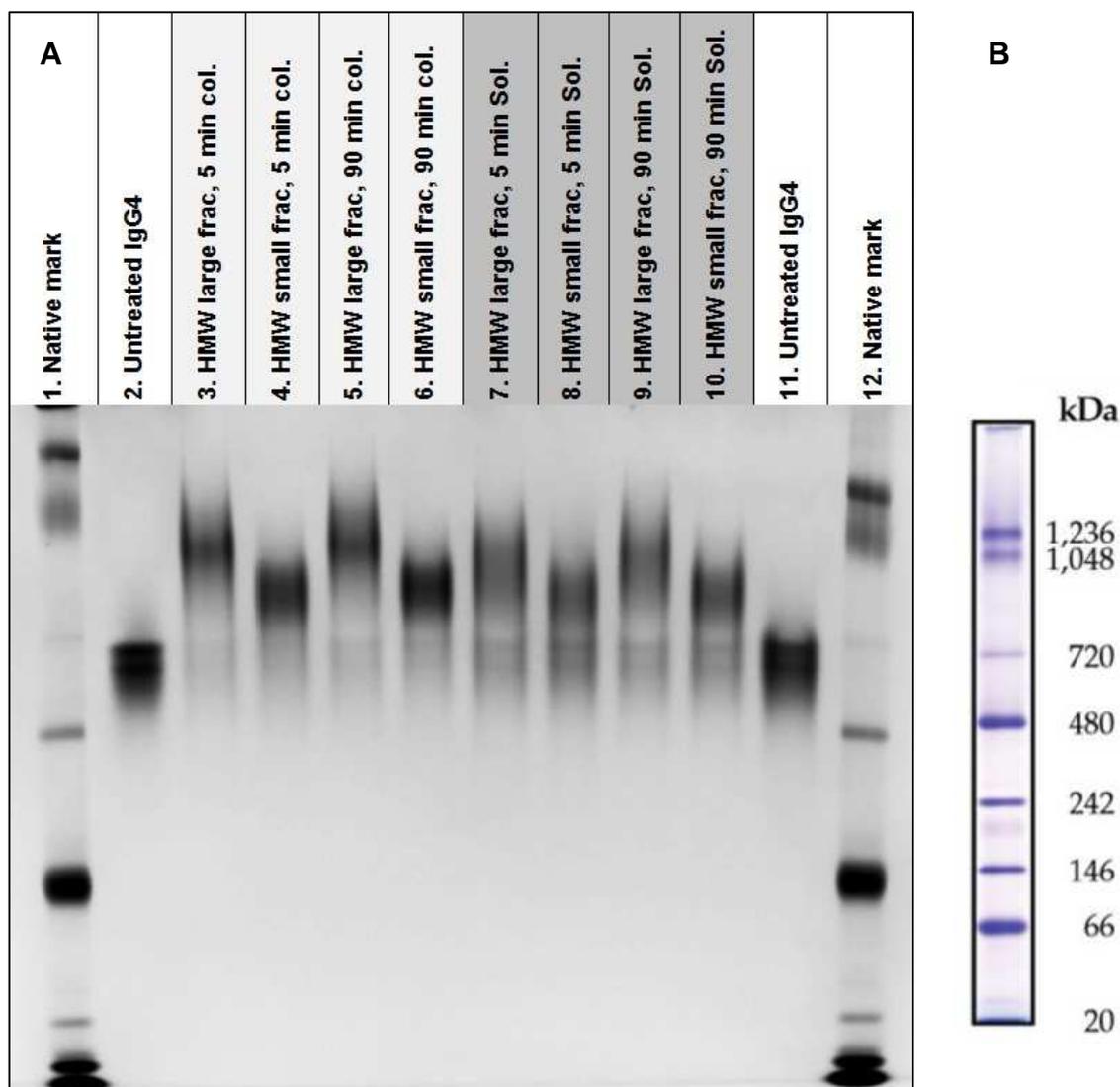


Figure 5.3 (a) Native-PAGE of untreated IgG4 (lanes 2 and 11) and HMW species formed in solution (lanes 7-10) or after elution from protein A (lanes 3-6) at low pH; samples were neutralised after either 5 min or 90 min incubation. Samples were run on an SEC column and separated into fractions designated monomer, *HMW large* and *HMW small* (monomer fractions not were not run on the gel displayed). The molecular weight marker *NativeMark* was run in lanes 1 and 12. The manufacturer's example of *NativeMark* run on a 3-12% Bis-Tris gel with molecular weight designations (kDa) is shown in (b).

5.3 Rate of monomer loss at low pH in solution

The IgG4 was found to be highly sensitive to pH within a critical range. At pH conditions lower than pH 2.8 aggregation was rapid; plateau occurred in less than 30 minutes. However, at > pH 3.0 completion took approximately 4 hours. It should be noted that the plateau was not found at the point of complete monomer loss, and varied depending on pH conditions; lower pH conditions resulted in greater total monomer loss, this can be seen in *Figure 5.4*. It is possible that what appears as a plateau in monomer depletion actually represents a point of reversible equilibrium between species at one or more stages of the aggregation process. An ancillary experiment was carried out in which monomer purified from initial aggregation runs was re-exposed to low pH under the same conditions as in the initial runs. Recovered monomer was found to display near-identical aggregation behaviour to initial monomer (data not shown). This indicates that the surviving monomer population is not distinct from the aggregating portion, supporting the possibility of an equilibrium mechanism. Despite the fairly large IgG concentration range tested (five-fold), concentration appeared to have little effect on aggregation kinetics, and observed differences in plateau did not follow a clear trend. Thus we hypothesise that aggregation plateau was determined by pH-dependent equilibrium between unfolded and native or re-folded monomers. While this could provide an interesting topic for more detailed investigation, it will not be the focus of this chapter. Instead, we concentrate primarily on initial aggregation *rates* as well as apparent total monomer loss.

All data sets were fitted with exponential decay curves using the following equation:

$$y = y_0 + A \cdot e^{-R_0 \cdot x} \quad \text{Equation 3}$$

where y_0 is the plateau point (total monomer remaining), A is the initial value (total monomer lost), R_0 is the rate constant (h^{-1}) and x is incubation time (h).

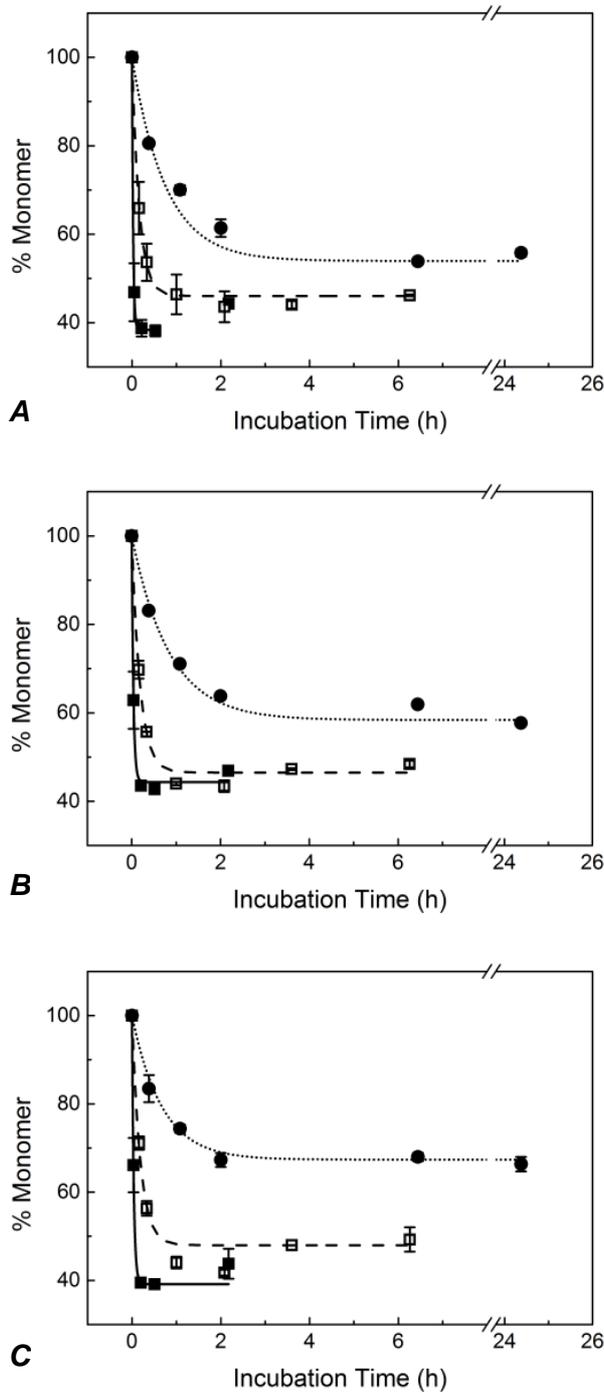


Figure 5.4 IgG monomer loss over time in solution at (a) 0.9mg/mL, (b) 2.7mg/mL and (c) 4.5mg/mL. Different symbols represent different incubation pH conditions: circles, pH 3.03; open squares, pH 2.95; filled squares, pH 2.78. Error bars show the standard deviation for each point based on full experimental repeats, $n = 2$. Exponential decay curves were fitted to the data using the equation $y = y_0 + Ae^{-R_0 \cdot x}$ (see Section 3.2 for equation specifics). In (a), for pH 2.78 the last time point (2.18h) was excluded from the curve fit (see section 5.4.1). All fits were significant with adjusted $r^2 > 0.99$ and $P < 0.01$.

The value of interest resulting from fitting curves to the data is the rate constant, R_0 . The rate constant provides a convenient comparator for chromatography experiments, similarly to that used by Shukla et al. (2007a). Rate constants varied most with pH, while IgG4 concentration had a limited effect on R_0 values, within the concentration range tested, as seen in *Figure 5.5*. To verify concentration effect, or lack thereof, R_0 was plotted against pH for the three IgG concentrations tested and linear regression with 95% confidence limits was applied to each data set. There were few data points at 0.9mg/mL so this data set was not fitted. Confidence intervals of the linear fits for 2.7mg/mL and 4.5mg/mL overlapped, indicating that IgG concentration did not significantly affect R_0 in solution (*Fig. 5.5*). This is useful in terms of allowing some flexibility in IgG concentration for column experiments, as elution fractions may not be of precisely the same concentration for every run.

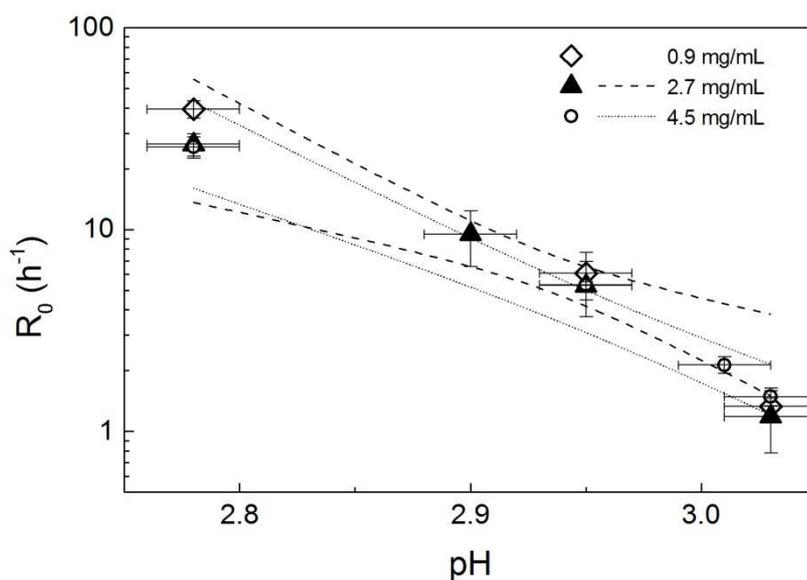


Figure 5.5 Semi-log plot of monomer decay rate, R_0 , against pH in solution at three different IgG concentrations. Y-error bars show the standard error for the R_0 parameter (*equation 3*) obtained from least squares fitting; x-error bars represent pH measurement error. Linear regression was applied for concentrations 2.7mg/mL and 4.5mg/mL; for clarity, *95% confidence intervals only* are shown for these fits. A trend has not been fitted to the 0.9mg/mL data because there are few data points.

5.3.1 Negligible effect of maltose

Maltose has been known to stabilise proteins and reduce aggregate formation (Bolli et al. 2010). The IgG4 initial formulation contained 1% maltose; subsequently, low concentrations of maltose were present in solution experiments (0.05% - 0.25%). Maltose was not added to elution buffers for column experiments so was not present during those incubations. Thus it was necessary to confirm that maltose was not reducing aggregation rates in solution. Additional solution experiments were carried out at pH 2.90 with maltose at concentrations from 0.15% to 4.0%, and 2.7 mg/mL IgG4. Maltose had no significant effect on IgG aggregation rate in these experiments, as can be seen in *Figure 5.6*. Low maltose concentrations (0.25% – 1.5%) also had no discernible effect in solution from 0.9 – 4.5 mg/mL IgG4, pH 3.0 (data not shown). Subsequently, maltose was not added to protein A elution buffers.

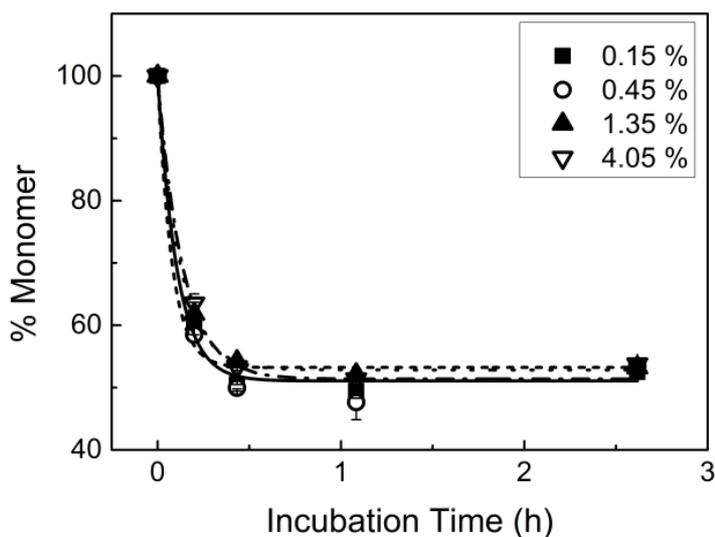


Figure 5.6 The effect of maltose concentration on the rate of IgG4 monomer loss at pH 2.90, 2.7mg/mL IgG4 is shown. Maltose concentration for each data set is shown in the graph legend. Curves were fitted with the exponential decay function in *equation 3*.

5.4 Effect of column step on rate of monomer loss at low pH

Monomer loss over time under various pH conditions is plotted for column and solution experiments side by side in *Figure 5.7.a-c*. All column experiment data was fitted with exponential decay curves as was done for solution data. *Fig.5.7.d* shows all column experiment curves side by side on a single unbroken scale.

Examining the plots in *Fig.5.7*, the initial rate of monomer consumption and the position and behaviour of the plateau can be compared for the different test conditions. Due to the sharp drop-off in pH sensitivity between pH 2.7 and pH 3.0, the data in *Figure 5.7.a-c* is displayed in three narrow ranges: pH 2.78 – 2.86 (*a*), pH 2.90 – 2.95 (*b*) and pH 3.03 – 3.11 (*c*). Note that the x-axis scales differ for each plot and also contain scale breaks. The pH values indicate the pH at which the IgG was incubated after elution from the column, which was the same as the pH of the original elution buffer. In all cases, inclusion of the protein A chromatography step immediately before incubation resulted in a faster rate of monomer decay at any given pH. Exponential decay kinetics to a point of apparent equilibrium was maintained in column experiments as in solution. This suggests that the chromatography step does not significantly change the basic aggregation mechanism, but does accelerate it.

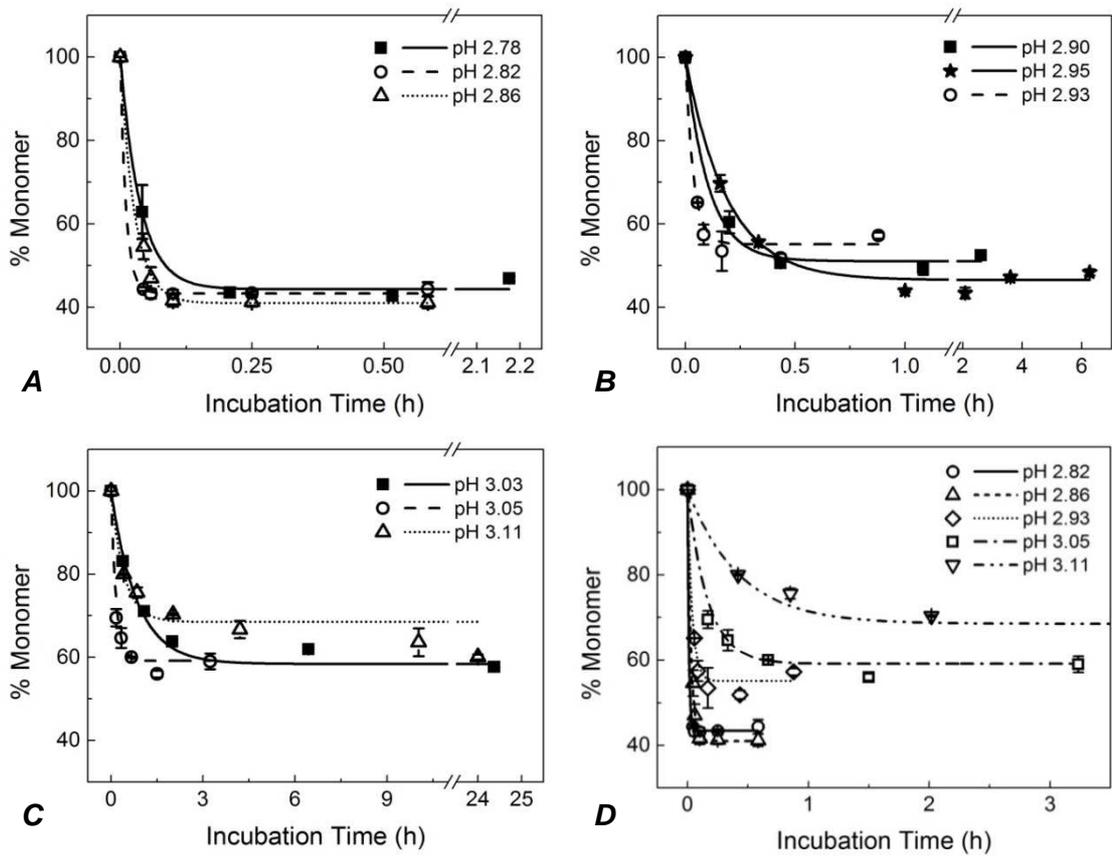


Figure 5.7 Rates of monomer loss at 2.7mg/mL IgG under different pH conditions in solution only (filled symbols and solid lines) and after elution from a protein A chromatography column (open symbols, dashed lines and dotted lines) (a-c). Different pH conditions are shown in the plot legends. Curves for column runs across the full pH range are shown together in (d). Error bars show the standard deviation for each time point based on full experimental repeats, $n = 2$. All data sets were fitted with exponential decay curves as was done for initial solution only data. All fits were significant with adjusted $r^2 > 0.98$ and $P < 0.01$.

5.4.1 Monomer leading shoulder

As can be seen largely in *Fig. 5.4* and *Fig. 5.7*, though a general plateau was reached in all experiments, it was not completely stable. In some cases (typically the harsher conditions) monomer concentration appeared to drift up again. This may be explained by the apparent formation of an altered monomer species, as seen in *Fig. 5.2*. The (extra) area contributed by this peak shoulder was counted as normal monomer in “% Monomer” calculations (see *equation 1, section 4.3.1*). Exclusion of the shoulder area would result in stabilisation of % Monomer values at the plateau, rather than upward drift up. On the other hand, whether the monomer shoulder species should be considered a precursor to the aggregated form, or even one of its break-down products, is contestable (Filipe et al. 2012). Thus, data points that drift up after the plateau were counted as “normal” for the purpose of curve fitting *unless* the fit was rejected by the software (see *Fig. 5.4.a*, pH 2.78). Nonetheless, it is worth noting that the shoulder may represent a degraded form of the product whose bioactivity differs from that of the native monomer (Wong et al. 2012); this could provide an interesting subject for further investigation but will not be addressed in this work. In other cases, after a period of stability monomer concentration began to drop (*Fig. 5.7.c*, pH 3.11), this is discussed below.

In column experiments the shoulder developed to a similar extent but more quickly than in solution experiments, supporting the model of acceleration of the mechanism of aggregation when the chromatography step was present.

5.4.2 Variation in monomer plateau

Within the lowest pH range, pH 2.78-2.86, the monomer plateau consistently fell between 40% and 45% of the initial concentration for both solution and column experiments (*Figure 5.7.a, d*). Within the highest pH range, pH 3.03-3.11, plateau points were similar for near-matching pH conditions (e.g. solution pH 3.03 versus column pH 3.05). The trend of reduced total monomer loss with increasing pH was matched by the column experiment data at pH 3.11, all in all the highest pH tested, with the plateau being at the highest percentage monomer compared to all other test conditions (*Fig. 5.7.c-d*). For the pH 3.11 column experiment a slow linear drop in

monomer continues between 4 hours and 24 hours, meaning that exponential decay may not be the best fit available for the data in this region and a different mechanism of aggregation may prevail. For the purpose of this study, an exponential decay fit was used as it described the initial period of rapid monomer loss well, giving an adjusted r^2 value of 0.985 with significance $P < 0.01$ for pH 3.11 (*Fig. 5.7*). It is possible that above a certain pH, or even after a certain duration, the system transitions to a different kinetic mechanism.

In other cases upward drift in the plateau occurred, but this can largely be attributed to the presence of monomer leading shoulder species described above.

In summary, the plateau point was not considered a robust parameter to quantify the aggregation that occurred under the given conditions. This is further illustrated in *Figure 5.8*, where the relationship between pH, y_0 values and IgG concentration for solution-only and column experiments is displayed. The general trend of decreasing monomer loss with increasing pH can be seen, and it is also apparent that the effects of IgG concentration and the column step on y_0 are not well defined.

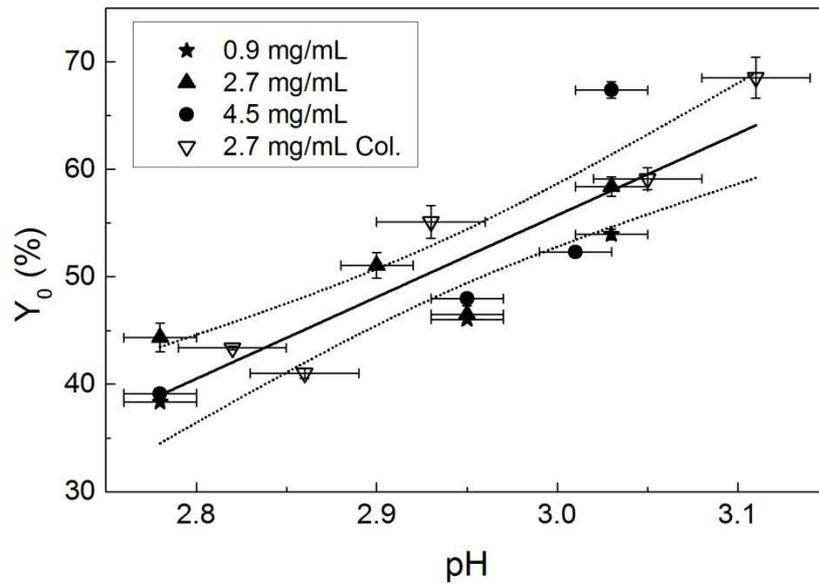


Figure 5.8 Y_0 (plateau point parameter) from exponential decay curve fits plotted against pH at 0.9mg/mL, 2.7mg/mL and 4.5mg/mL in solution, and at approximately 2.7mg/mL after protein A chromatography. Y-error bars show the standard error for the Y_0 parameter (*Equation 3*) obtained from least squares fitting; x-error bars represent pH measurement error. The data was pooled and fitted with a linear trend (solid line); error weighting was not used for the fitting. The 95% confidence interval is shown as dotted lines. The fit was significant with an adjusted r^2 of 0.75 and $P < 0.01$.

5.4.3 Influence of column step on R_0

Focussing on the initial period of rapid monomer loss under the various conditions displayed in *Figure 5.7*, it can be seen that the column step had a considerable effect on the rate at which monomers were converted to aggregate species during the low pH hold. Looking at *Fig. 5.7.b*, between pH 2.90 and pH 2.95 it is clear that that rate of monomer loss is fastest for the column experiment. Comparing pH 3.03 solution data to pH 3.11 column data (*Fig.5.7.c*), the initial decline in monomer progresses at a similar rate for both experiments; this is significant considering the extreme pH sensitivity of the system.

Figure 5.7 allows the kinetics of individual cases to be assessed and compared. Due to the pH sensitive nature of the experiments, direct comparison of column and solution data is difficult because of small differences in pH. Thus, individual fitted parameters, R_0 and y_0 , were plotted against pH to more clearly delineate the effect of the chromatography step.

Figure 5.9 illustrates how the protein A chromatography step causes a shift in the rate of monomer loss compared to low pH incubation alone. It is clear that two separate trends exist for R_0 against pH for solution-only and on-column experiments. In *Fig.5.9*, 95% confidence bands were calculated using the standard error for the R_0 parameter (obtained from least squares fitting) for weighting. The linear fit lines for column and solution-only R_0 values are almost parallel; this supports the theory that the aggregation mechanism is accelerated, but not significantly altered, by the chromatography step. The data represents the minimum difference between after-column and solution-only rates, in that the earliest reasonable start time for low pH incubation was used for column experiments (see section *5.2 Experimental design*).

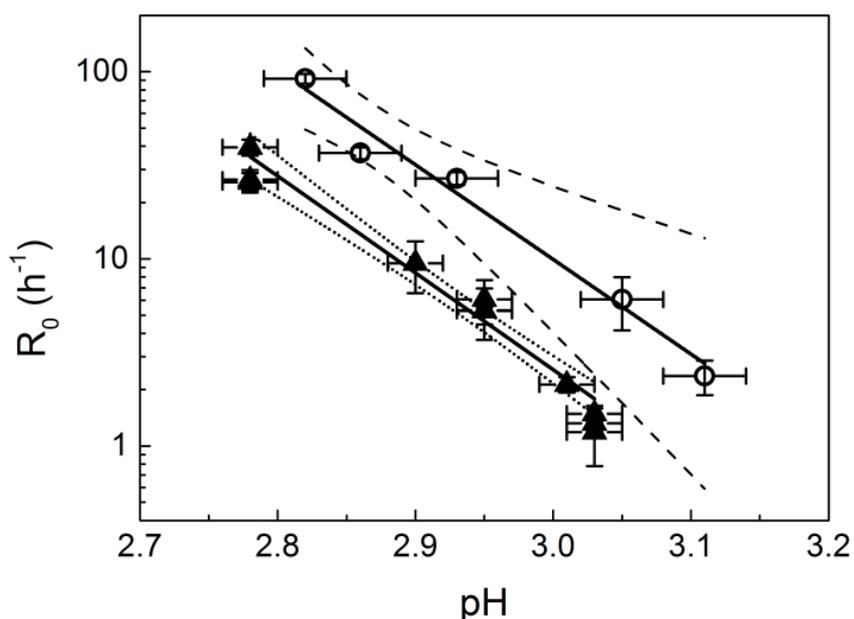


Figure 5.9 Semi-log plot of monomer decay rate, R_0 , against pH for solution-only experiments (filled triangles) and experiments including a protein A chromatography step (open circles). For column experiments, the typical concentration of the incubated elution fraction was 2.7mg/mL. Solution-only data seen in *Figure 5.5* was pooled due to overlapping 95% confidence intervals for the linear trends. For solution-only data points, x error bars represent the measurement error of the laboratory pH probe/meter, as reported by the manufacturer. For chromatography experiments, x error bars correspond to measured error based on three replicate measurements (with a micro pH probe). For both solution-only and column experiments, y error bars represent the standard error for the R_0 parameter (*equation 3*) obtained from least squares fitting. A linear trend was fitted to each data set, as shown by solid lines; dashed/dotted lines represent 95% confidence limits for each fit. Both fits were significant with $P < 0.01$ and adjusted r^2 of 0.97 and 0.90 for solution-only and column experiments, respectively.

It should be noted that the y axis in *Figure 5.9* is scaled logarithmically; what appears as a linear fit is actually exponential. This is of interest as it shows how the data collected is centred on a tipping point for aggregation behaviour of the IgG4; two orders of magnitude in aggregation rate are covered in approximately 0.3 pH units. Outside the range of conditions tested aggregation becomes either unacceptably rapid, or slow to the point of practical insignificance (in the context of an early/mid purification process). It is important to consider that the column step has an effect despite the narrow range in which it was possible to quantify effects using the methods described. In effect, the column step seems to push the tipping point in the direction of milder conditions. A related result was seen by Gagnon et al. (2015) for an IgG1. They found that adjusting to pH 3.0 after elution from protein A (at pH 3.5) resulted in considerable aggregate formation which was not observed in the absence of the chromatography step. It was concluded that elution from protein A increased the protein's "vulnerability" to a secondary stress, such as pH (Gagnon et al. 2015).

5.5 Conclusion

The data gathered captures the behaviour of an IgG4 within a fairly compact experimental space. The IgG4 demonstrates highly pH-dependent and apparently concentration-independent aggregation behaviour. Given these observations, it is hypothesised that the aggregation behaviour is governed predominantly by a pH-dependent unfolding/ re-folding equilibrium. The data shows that the chromatography step causes aggregation rates to increase, while other aspects of the aggregation kinetics appear largely un-affected. One possible molecular mechanism is that the conformational change, as well as changes in the hydration layer around the protein, required for protein adsorption and subsequent desorption, alter the structure of the IgG, possibly exposing regions of the protein involved in unfolding transitions. Being in this condition, immediate exposure to the acidic environment could increase the possibility of the protein unfolding and losing native structure to some extent before progression to aggregate formation during the pH hold (Filipe et al. 2012, Gagnon et al. 2015, Ueberbacher et al. 2010).

It should also be considered that only the tail end of the protein A elution peak was assessed in this work, as all prior effluent was at a pH too high to induce aggregation. The peak tail may represent a monomer sub-population with a higher affinity for the ligand. It is possible that these monomers bind more-tightly than early-eluting species, undergo increased structural perturbation during adsorption and desorption, and are thus more likely to aggregate, increasing the rate. This type of phenomenon, whereby a strongly bound sub-population was generated, has been described previously and investigated in detail for hydrophobic interaction chromatography (Haimer et al. 2007, Jungbauer et al. 2005). Heterogeneity in recombinant protein products is commonly documented, with species ranging from distinct charge variants to product molecules with a single erroneous amino acid substitution (Harris and Kilby 2014).

6 Using Neutron Reflectivity to Examine the Structure of IgG4 Adsorbed to a Model Surface

6.1 Route to reflectivity

At this point in the work, long-standing methods of protein analysis were used to characterise aggregates that formed after elution from protein A and after low pH incubation alone. HMW species were separated by SEC and collected in fractions. These were analysed using reducing and non-reducing SDS-PAGE and Native-PAGE. The results suggested that the various HMW species, including dimers, trimers and larger oligomers, were stable at low concentrations, and that post-column aggregates did not differ significantly from solution-only aggregates.

Having found that the aggregates themselves did not easily reveal any unique column-induced features, a decision was made to move the emphasis of investigation back to the cause, and try to advance understanding of the structure or behaviour of antibodies *while* adsorbed. This led to identification of neutron measurement techniques as an interesting and relatively little-used method for examining large structures at the molecular level.

It was first possible to gain access to neutron beam facilities through collaboration with researchers at the University of Delaware, using facilities based in Maryland, USA (the NIST Center for Neutron Research (NCNR), Gaithersburg). Here, the methods used were similar to those developed by the collaborating researchers, and the model system represented a simplification of that which we ultimately wished to examine. After successful completion of experiments at the NCNR, a proposal was submitted for work involving a more complex model system; this was to be carried out at neutron beam facilities based in Oxfordshire, UK (Isis Neutron and Muon Source, Harwell Science and Innovation Campus) – the proposal was well-received and resulted in implementation of the experiments. The results of both sets of

experiments using the neutron measurement techniques are discussed in the following two chapters of this work.

6.2 Neutron reflection with IgG at the NCNR

In this section a neutron reflection technique is used to probe the structure and orientation of IgG adsorbed to the model surface silica. Neutron methods are not commonplace in biologics and bioprocess development research, largely because facilities that produce the high intensity beams needed for measurements are large and expensive, and only a handful exist worldwide. However, these methods possess unique qualities useful for the study of biological structures and complexes.

Perevozchikova et al. (2015) produced neutron reflectivity data characterising an IgG1 adsorbed to a silicon wafer in citrate buffer at pH 4.5 and pH 6.2. IgG1 was applied to the silica surface and neutron reflectivity data was collected during equilibrium adsorption over 6 hours. The adsorption surface was then rinsed with the protein-free adsorption buffer. The data indicated that an IgG1 monolayer formed on the silica surface, with IgG-molecules apparently in a flat-on orientation (i.e. potentially the Fc region and both Fab arms contacting the surface). Prior to the buffer rinse, features of the data suggest a second protein layer distal to the surface and of a much lower volume fraction. It is suggested that the second layer constitutes some additional weakly adsorbed material, possibly a portion of material in the wafer-proximal layer that is not oriented flat-on. After rinsing, only the flat-on monolayer persisted and was not significantly altered in depth or volume fraction (Perevozchikova et al. 2015).

6.3 Chapter objectives

In this chapter the first objective is to determine whether our model IgG4 behaves in the same way as IgG1 under the conditions investigated by Perevozchikova et al (2015), as described above. In gathering this data we hope to provide insight into the adsorption behaviour of mAbs

at solid-liquid interfaces encountered in bioprocessing, particularly in the event that the two IgG molecules display similar behaviour. Perevozchikova et al. (2015) used orthogonal methods to detect and characterise aggregated IgG1 species in the rinse fractions (described above), i.e. aggregates present in solution after desorption from silica. They found that IgG1 (i) exposed to silica and (ii) desorbed from silica at pH 4.5 contained a significant proportions of aggregated material compared to the control stock solution, with the greatest increase in aggregated mass seen in the desorbed sample (ii). This data is important in demonstrating the link between adsorption events and aggregate formation.

As well as investigating IgG4 adsorption under buffer conditions used by Perevozchikova *et al.*, we will extend the study to determine adsorbed IgG4 conformation under conditions relevant to affinity chromatography.

The second objective of this chapter is to establish whether neutron reflectivity is a suitable method for characterising adsorbed IgG structures, with a view to moving from a simple model system in which IgG adsorbs directly at the solid-liquid interface, to a more complex system in which an affinity ligand is introduced, namely protein A.

6.4 Neutron measurements for biological materials

Neutrons are non-charged particles with a negligible electric dipole moment. Consequently, they do not interact with the electron charge cloud of atoms like X-rays do, but instead with other neutrons present in the nuclei of atoms (de Bergevin 2009, Pynn 2009). Interactions or *scattering* is short-range and, as the nucleus of an individual atom is typically 10^5 times smaller than its total volume, solid materials seem of low density to neutrons travelling through them (Pynn 2009). Thus, neutrons possess high penetrating power, meaning they can pass long distances (meters) through many materials without being scattered or absorbed (Pynn 2009).

Neutron beams are relatively low intensity; the flux of a typical neutron beam is around 10^{14} times lower than the photon flux provided by X-ray instruments at synchrotron sources (Pynn

2009). Consequently, neutrons beams do not cause structural damage to biological samples, making them useful for studying the structure of biological molecules and complexes. The low intensity of the neutron beam and the high proportion of neutron transmission by samples (compared to scattering) mean that signal intensity is weak. However, with the correct instrumentation and experimental design the problem of signal weakness is overcome and outweighed by its benefits (Clifton et al. 2013, Pynn 2009).

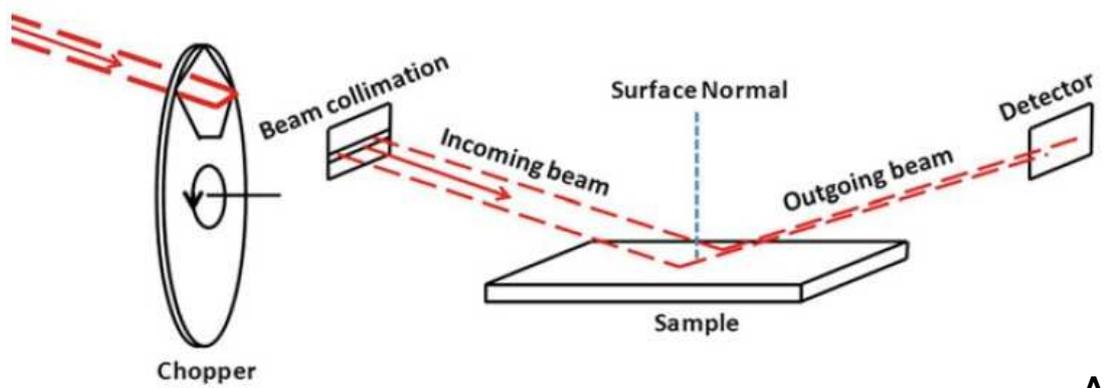
When used to study biological materials, another crucial property of neutrons is that their interactions with each of the hydrogen isotopes ^1H (protium) and ^2H (deuterium) are highly distinct. The property known as *scattering length*, which has units \AA at a scale around 10^{-5} , describes a material's ability to scatter neutrons; deuterium has a large scattering length while the scattering length of protium is negative (negative scattering length meaning that the neutron is subjected to an attractive potential as it approaches the nucleus) (Clifton et al. 2013). When measuring bulk materials or large structures like proteins, scattering lengths of all the atoms in a given molecule are summed and divided by the volume of the molecule to give the quantity *scattering length density* (SLD). It follows that the SLD of H_2O is negative, while that of D_2O is positive and large compared to that of many substances. The difference in SLD between two bulk phases or molecules is often referred to as *contrast* (Pynn 2009).

The abundance of hydrogen in nature makes it relatively easy to exploit this property (ability to create contrast) in two major ways: the material of interest may be isotopically labelled, or the sample may be suspended in a D_2O -based mobile phase. Using these methods can create high contrast between molecules of interest and their suspension medium, or between different subunits within complex molecules. Further, the D_2O content of the solvent can be adjusted such that the solvent SLD matches that of a particular component of the system being measured. Thus, scattering from the chosen component is no different to the solvent background, rendering it "neutron invisible" (Clifton et al. 2013). These techniques are known broadly as *contrast variation* and make it possible to use neutrons to characterise a range of structures that occur in nature.

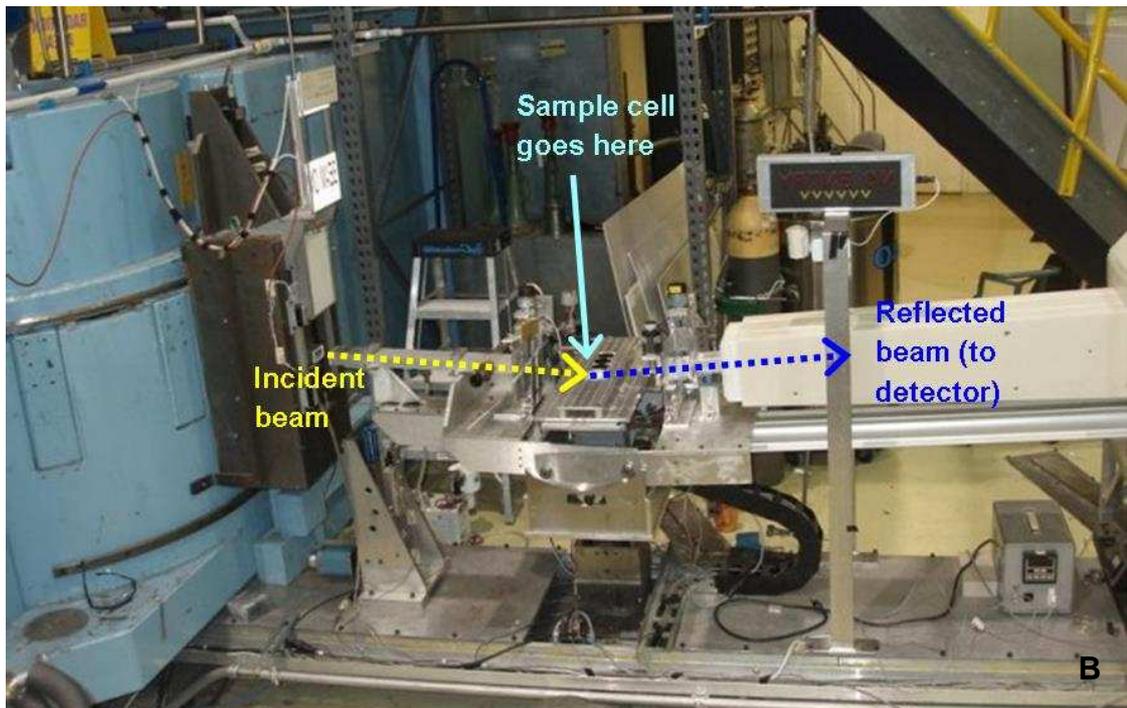
6.5 Neutron Reflectivity

One quality of neutrons as probes is that they offer the possibility to study structures over a broad range of length scales and under an enormous variety of physical conditions; pressure and temperature do not affect neutron waves, for example. Possible length scales range from atomic scale for crystalline materials, to hundreds of nanometres for large biological complexes and viruses. However, a given length scale and sample configuration may only be probed by using an appropriate instrument (Clifton et al. 2013, Fermon et al. 2009, Pynn 2009). One such class of instruments is the neutron reflectometer. Though many subclasses of reflectometer exist, in general using neutron reflectometry/reflectivity (terms used interchangeably) allows layered systems, or layers of material adsorbed parallel to a planar surface, to be studied in detail.

Figure 6.1A is a simplified depiction of a horizontal neutron reflectivity rig. An incoming neutron beam first passes through a series of *choppers* that select an appropriate wavelength band. The beam is then collimated as it passes through slits of defined dimensions. The beam enters the sample from above at a very low *grazing angle* which results in total external reflection of the beam; this is known as *specular reflection* as the angle of incidence is equal to the angle of reflection (Clifton et al. 2013, Pynn 2009). It may seem counter-intuitive that total neutron reflection can occur in this way, since one of the benefits of neutrons is their penetrating power. However, below a critical angle specular reflection occurs for smooth, flat surfaces, due to the difference in SLD between the two phases. Many of the laws of optical refraction are applicable to neutron reflectivity and, indeed, SLD is often termed the *neutron refractive index* (Fermon et al. 2009). Below the critical angle of incidence not all neutrons are reflected, some are instead transmitted. The presence of material adsorbed parallel to the surface affects the intensity of neutron reflection (Pynn 2009). The data obtained can be used to generate a model of the layered structure in the surface normal direction (*Fig. 6.1A*) and determine the bulk composition of the adsorbed material. This will be discussed in more detail in *section 6.6*.



A



B

Figure 6.1 (a) Schematic representation of a neutron reflectivity instrument showing a pulsed collimated beam of neutrons undergoing specular reflection from a sample surface. (b) Photograph of the NG7 reflectivity rig at the NIST Center for Neutron Research. The image illustrates the very low angles at which specular reflection occurs – dotted lines indicate the path of the neutron beam.

Figure 6.1 A is from Clifton et al. (2013), reprinted with permission. Copyright © 2013, Springer Science+Business Media New York.

Figure 6.1B shows the *NG7 cold neutron horizontal reflectometer* at the NCNR which was used for experiments described in this chapter. The same reflectivity instrument was used by Perevozchikova et al. (2015) to characterise various adsorbed proteins including IgG1. The incident beam travels through slits towards the sample stage, onto which the sample holder is fixed. In this case the sample holder was a solid-liquid flow cell (SLFC), as depicted schematically in *Figure 6.2*. This sample holder will be referred to as the *SLFC* or the *reflectivity cell*. The reflectivity cell is assembled manually and contains a flat silicon wafer with highly a polished surface which constitutes the adsorption medium. A Viton spacer creates a sealed chamber with a 0.1mm gap between the wafer and the bottom slab (Perevozchikova et al. 2015). Tubing fitted with simple valves is connected to inlet and outlet cores in the bottom slab and a high precision syringe pump controls flow, allowing continuous flow of buffer solution thorough the SLFC chamber at a chosen flow rate. When protein solution is introduced into the SLFC, protein that adsorbs to the silicon wafer is characterised by neutron reflectivity. In this work an SLFC was used to examine the structure of material adsorbed at the solid-liquid interface. It should be noted that other sample configurations are also common, for example air-liquid or liquid-liquid interfaces.

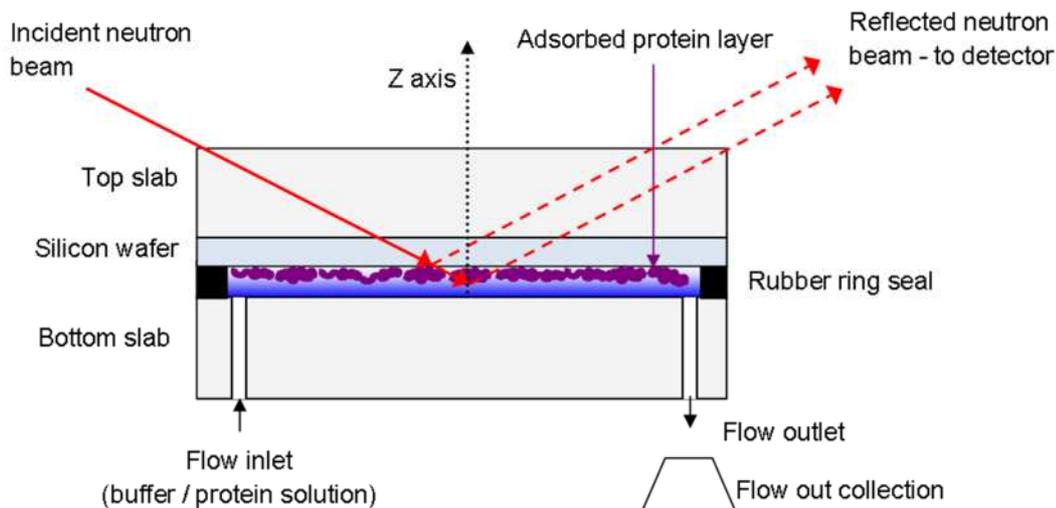


Figure 6.2 Schematic representation of the flow cell used for reflectivity experiments. The flow-through cavity has a volume of approximately 1.5 mL and a height of 0.1 mm. Flow is controlled by a high-precision syringe pump; during experiments all space inside the cell is filled with aqueous solution at all times. For a bare silicon surface the incident angle of the collimated neutron beam is equal to the reflected angle. Neutrons reflect at interfaces present inside the reflectivity cell, such as the SiO₂ –protein interface and the protein-water interface. The scattering pattern generated by adsorbed material depends on five major factors: its chemical composition, the number of layers present across the z axis, their respective depths and volume fractions, and the resolution or smoothness of transition between layers.

6.6 Neutron Reflectivity Theory and Data Analysis

6.6.1 Basic reflectivity theory

The neutron wave can be described in terms of a wavevector \vec{k} , having a magnitude and a direction. The wavevector \vec{k} points along the neutron's trajectory and has a magnitude

$$|\vec{k}| = 2\pi/\lambda$$

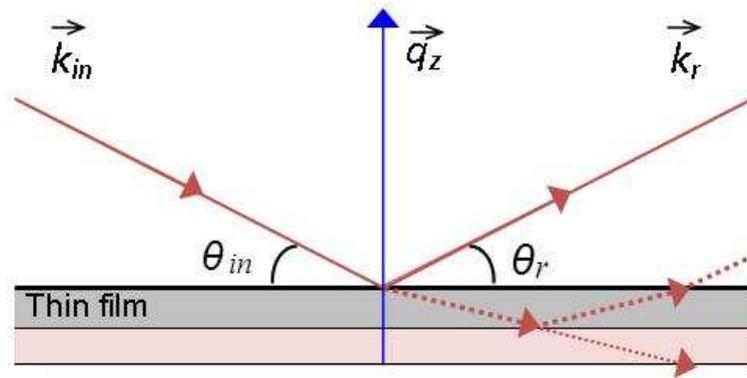
where λ is the neutron wavelength.

When the incident wave is reflected from the surface momentum transfer occurs; the wavevector transfer in the z direction, perpendicular to the sample surface, is denoted Q_z . For specular reflectivity $\vec{k}_r = \vec{k}_{in}$. Therefore

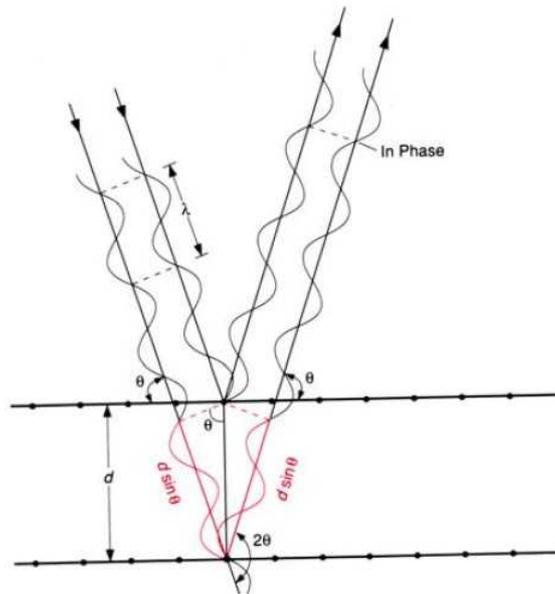
$$Q_z = k_{in,z} - k_{r,z} = \frac{2\pi}{\lambda} (\sin \theta_{in} + \sin \theta_r) = \frac{4\pi \sin \theta_{in}}{\lambda}$$

where $k_{in,z}$ and $k_{r,z}$ are the z components of the incident and reflected neutron wavevectors, respectively; θ_{in} and θ_r are the angles of incidence and reflection (which are equal for specular reflection), respectively (Pynn 2009). This relationship is depicted schematically in *Figure 6.3A*.

Q_z has units of \AA^{-1} and is often abbreviated to simply Q .



A



B

Figure 6.3 (a) Momentum transfer in neutron reflectivity. Momentum transfer in the z direction is denoted q_z . For specular reflection at a smooth interface the incident wavevector, k_{in} is equal to the reflected wavevector, k_r (solid red line). As the angle of incidence increases above a critical angle, the neutron wave is partially reflected and partially transmitted (dotted line). If a thin film is present at the interface, transmission through the film and reflection at the distal interface is dependent on the *scattering length density* of the film. (b) When neutron waves are reflected by parallel interfaces interference patterns can appear. Here, *constructive* interference is depicted for a dual interface system. Interference is constructive when waves reflected at different interfaces remain in phase, which occurs when the travelled distance between interfaces is an integral multiple of the wavelength. Conversely, if the reflected waves become out of phase the interference is destructive, causing a reduction in reflected intensity.

Figure 6.3 b is from Pynn (2009). Copyright © 2009 Springer Science+Business Media LLC.

Figure 6.4A shows simulated reflectivity data for a silicon-D₂O interface. Reflectivity data is initially presented as reflected neutron intensity (relative to the incident beam) as a function of Q . Data is collected across a Q range spanning around 1.5 orders of magnitude; this can be achieved by varying the neutron wavelength, the angle of incidence, or a combination of both. Looking at the top two plots in *Fig. 6.4*, a flat edge at the low end of the Q range can be seen; this is known as the *critical edge* and represents a region of total external reflection, giving a relative reflective intensity of 1. For an interface between a flat silicon surface and a D₂O liquid phase total external reflection occurs below Q values of 0.014 \AA^{-1} . Above the critical angle, which equates to a critical Q , reflected intensity decreases as the fourth power of the angle of incidence (Pynn 2009). Above a Q value of approximately 0.3 \AA^{-1} signal cannot be resolved from background; these factors define the chosen Q range.

Figure 6.3A shows how neutrons can be reflected at the upper and lower interfaces of a thin film, while in *Fig. 6.3B* the principles of constructive and destructive interference are illustrated. In *Figure 6.4B* we can see the effect of a thin film at the solid-liquid interface on the reflectivity profile. Fringes in the profile are caused by interference between neutrons reflected at the silicon-film and film-D₂O interfaces (Clifton et al. 2013). The distance in Q_z is related to the thickness of the film by

$$Q_{z2} - Q_{z1} = \frac{2\pi}{d}$$

where $Q_{z2} - Q_{z1}$ is the distance in Q between the fringes (\AA^{-1}) and d is the thickness of the interfacial film (\AA) (Clifton et al. 2013, Pynn 2009).

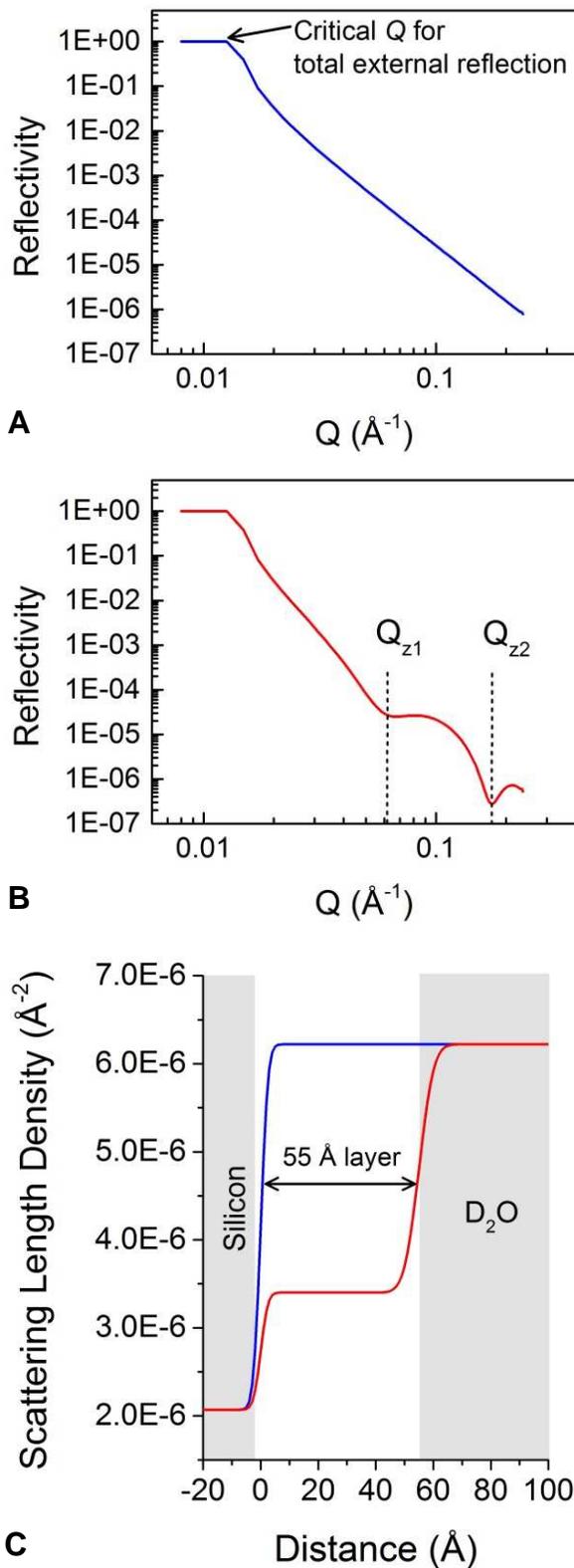


Figure 6.4 Simulated reflectivity data for a silicon-D₂O interface with no material present at the solid-liquid interface (a) and with a 55 \AA homogenous film of material (b). Below the critical Q for D₂O total reflection occurs, giving a relative reflective intensity of 1. The presence of the film of material alters the reflectivity profile; fringes at Q_{z1} and Q_{z2} are present due to interference between neutrons reflected from the silicon-film and film-D₂O interfaces. The distance in Q_z between the fringes is related to the thickness of the film. The strength and direction of the fringes depends on differences in SLD between the two bulk phases and the film. Reflectivity software is used to convert the data into the profile of scattering length density against z-direction distance that it describes (c). Before the interface the SLD corresponds to the substrate, silicon; in the absence of the film the SLD value transitions directly to that of the bulk mobile phase, D₂O (blue line). When the thin film is present there is an additional step in the transition (red line), representing the SLD of the thin film.

Figure 6.4 was adapted from Clifton *et al.* (2013).

As noted previously the *scattering length density* of a material describes its strength of interaction with neutrons – the higher the SLD, the stronger the interaction, the greater the intensity of scattered neutrons. Fringes as seen in *Fig. 6.4B* are typically not pronounced for films with depth less than 35Å. However, the SLD of the film relative to the other phases significantly impacts the shape of the reflectivity profile (Clifton et al. 2013). Recall the concept of *contrast variation* described in *section 6.4*. Implementing a contrast variation approach to experimental design allows complex information to be obtained even for very thin films and sparse layers of material.

6.6.2 Experiment design

When studying proteins, using a buffer based in 100% D₂O is effective as this gives the maximum possible contrast with natural protein (that is, protein that is not perdeuterated). In order to generate an accurate model fit of the data, at least one other *contrast* should be used. 100% H₂O gives good contrast against protein, but because H₂O is a weak scatterer detail obtained on the interfacial structure is less defined than with D₂O.

The scattering length densities of materials that will be present during neutron measurements must be estimated when planning the experiment. For pure elements and small molecules a reference table may be used, or the calculation can be done manually using

$$\rho = \frac{\sum b}{V}$$

where ρ is the scattering length density (Å⁻²), $\sum b$ is the sum of the scattering lengths (Å) and V is the molecular volume (Å³).

Scattering lengths of elements found commonly in biological molecules are shown in *Table 6.1*. Silicon is also shown as this is the substrate for adsorption in the reflectivity cell. The scattering length densities of the bulk phases in the reflectivity cell (*Fig 6.2*) are shown in the left-hand columns, note the positive value for D₂O and the negative value for H₂O. For large molecules like proteins an online SLD calculator may be used. In all cases, the level of expected hydrogen-deuterium exchange must be accounted for.

Table 6.1 Scattering lengths of elements commonly found in biological molecules, silicon is included as this is common to reflectivity methods. SLD values for H₂O, D₂O, silicon and silicon (di)oxide are also shown. SLD values were calculated using the NCNR Neutron Activation and Scattering Calculator at www.ncnr.nist.gov/resources/activation/.

Element	Scattering length (x10 ⁻⁵ Å)	Molecule	Scattering length density (x10 ⁻⁶ Å ⁻²)
H	-3.74	D ₂ O	6.33
D	6.67	H ₂ O	-0.56
C	6.65	Si	2.07
N	9.36	SiO	3.6
O	5.80	SiO ₂	4.19
S	2.85		
P	5.13		
Si	4.12		

The level of exchange significantly influences SLD because of the large difference in scattering length between protium and deuterium. Exchange rates depend critically on the number of labile hydrogens in the molecule and the percentage of D₂O in the solution. Solution pH and folded protein structure are also important factors (Clifton et al. 2013, Houde et al. 2009, Wales and Engen 2006). For IgG in 100% D₂O at pH 4.0 for more than seven days, exchange is assumed to be between 70% and 90%. *Table 6.2* contains manually calculated SLD values for a range of amino acids. The SLD of the dipeptide diphenylalanine in which the amide and hydroxyl hydrogens are exchanged to deuterium, and the SLD of diphenylalanine in H₂O (no exchange) are also shown. In this case 5% of the total hydrogenous volume is exchanged to deuterium while the SLD has increased by 12.6%, illustrating the impact of hydrogen-deuterium exchange on SLD at peptide scale. The SLD of IgG4-κ at a range of exchange levels was determined using the online tool Biomolecular Scattering Length Density Calculator (available at <http://psldc.isis.rl.ac.uk/Psldc/>). The amino acid sequence for the IgG, the percentage D₂O in the liquor and the chosen exchange level was input into the calculation tool; the output is shown in *Table 6.3*. The SLD values shown here, as well as those of the bulk phases D₂O, H₂O and silicon, were used to constrain the model when fitting the reflectivity data. The values also show

that there is sufficient contrast in SLD between the hydrogenous protein and buffers based in either 100% D₂O or 100% H₂O. The silicon wafer is washed with piranha solution before assembly as part of the SLFC; exposure to air between washing and assembly results in a layer of silicon oxide on the wafer surface. SiO₂ has an SLD of 3.4 x10⁻⁶ Å⁻², and the layer depth is typically between 5Å and 15Å, this provides a consistent starting point for data analysis.

Table 6.2 Molecular volumes, sum of scattering lengths ($\sum b$) and scattering length densities for a range of amino acids and for a dipeptide in H₂O and in D₂O where the amide and hydroxyl hydrogens are exchanged to deuterium. Molecular volumes were calculated using the Molinspiration Property Calculation Service at www.molinspiration.com.

Molecule	Formula	Molecular Volume (Å ³)	$\sum b$ (x10 ⁻⁵ Å)	SLD (x10 ⁻⁶ Å ⁻²)
Lysine	C ₆ H ₁₄ O ₂ N ₂	146.3	17.84	1.22
Cysteine	C ₃ H ₇ O ₂ NS	102.2	17.57	1.72
Glutamine	C ₅ H ₁₀ O ₃ N ₂	131.6	31.96	2.43
Phenylalanine	C ₉ H ₁₁ O ₂ N	156.0	39.64	2.54
Valine	C ₅ H ₁₁ O ₂ N	117.7	13.06	1.11
Diphenylalanine in H ₂ O	C ₁₈ H ₂₀ O ₃ N ₂	292.4	80.96	2.77
Diphenylalanine in D ₂ O	C ₁₈ H ₁₈ O ₃ N ₂ D ₂	292.4	101.78	3.48

Table 6.3 Scattering length density for IgG4-κ in H₂O and in D₂O at a range of exchange levels. The numbers of residues, number of exchangeable hydrogens, molecular volumes and sums of scattering lengths are also shown. All SLD values were calculated using the Biomolecular Scattering Length Density Calculator available at <http://psldc.isis.rl.ac.uk/Psldc/>.

D ₂ O (%)	Exchange (%)	Number of residues	Exchangeable H at pH 7	Molecular Volume (Å ³)	$\sum b$ (Å)	SLD (x10 ⁻⁶ Å ⁻²)
0	0	1336	2393	177756.4	0.345	1.94
100	50	1336	2393	177756.4	0.469	2.64
100	70	1336	2393	177756.4	0.518	2.92
100	90	1336	2393	177756.4	0.568	3.20

6.6.3 Data analysis methodology

The process of analysing neutron reflectivity data can be broken down into four major stages.

6.6.3.1 Data reduction

The first stage is *data reduction* which involves stitching together data points from the low and high ends of the Q range (two different incident angles are used to cover the whole Q-range, and it takes longer to reach sufficient neutron counts at higher Q), taking averages from repeat measurements, and identifying any oddities in the data. For example in an SLFC, if the transition from H₂O-based buffer to D₂O-based buffer is incomplete, the shape and position of the critical edge will be affected due to the presence of H₂O. Depending on the software used, background measurements may be subtracted at this stage (see *section 4.3.5.4*). Data reduction is a fairly straightforward process and only needs to be done once.

6.6.3.2 Building a constrained model

The next stage is *model fitting* which forms the bulk of the work. In this work, data was analysed using in-house software RasCal (A. Hughes, ISIS Spallation Neutron Source, Rutherford Appleton Laboratory) (Hughes 2013). This uses an optical matrix formalism to fit Abeles layer models to the data (Clifton et al. 2012). In this stage a model is set up and fitted to the data iteratively by a software algorithm. Data is converted from a plot of reflectivity against Q to a profile of SLD against z-direction distance (*Fig. 6.4*). The model is constructed in terms of slab layers. Input parameters include the number of layers, the thickness and SLD of each layer, and the roughness, or transition gradient between layers. For each parameter, limits are set to facilitate the fitting and improve the likelihood of generating a model that is a true representation of the sample. To begin the process, the known properties of the sample material should be evaluated. For proteins, the hydrodynamic radius, R_h , can be assumed to be the maximum depth for a monolayer of adsorbed protein. Measured by dynamic light scattering, the hydrodynamic radius of a protein is defined as the radius of an equivalent hard sphere diffusing at the same rate as the protein. When refining the model, the radius of gyration, R_g may be taken into consideration – the R_g of a protein can be defined as the average distance between the

centre of the molecule and each extremity of its mass. Realistically, R_g defines an ellipsoid rather than a sphere and is a measure of *elongation* (Rayner et al. 2014). The ratio between R_h and R_g reveals information about a molecule's shape: for roughly spherical molecules such as globular proteins the ratio may be less than or close to 1; when molecules are elongated R_g becomes larger than R_h .

The SLD of a layer is determined not only by the type of material present but also the level of coverage. The resultant value is an average SLD that lies between that of the solvent and that of the pure adsorbed material (i.e. protein). Therefore, when constructing the model, SLD is allowed to vary within this range. The fitted value may be used subsequently to estimate the coverage level for each layer using the following equation:

$$\phi_l = (\rho_s - \rho_l) / (\rho_s - \rho_a)$$

Equation 4

where ϕ_l is the volume fraction of adsorbed material, ρ_s is the SLD of the solvent, ρ_l is the fitted SLD of the layer and ρ_a is the calculated SLD of the (pure) adsorbed material (Clifton et al. 2011). It is important to note that the value for volume fraction obtained in this way does not account for tertiary structure of the adsorbed material. This is significant for proteins because the tertiary structure may incorporate solvent-filled cavities and highly protected hydrogen bonds (Fernandez et al. 2003).

IgG is an asymmetric molecule. This means that hypothetical layers can be incorporated into the model. For example, if the IgG adsorbs so that both Fab arms are in contact with the surface while the Fc region orients away from the surface, pointing out into the mobile phase, two hypothetical layers would likely be observed. A more protein-dense layer proximal to the substrate could represent the Fab arms, and a less dense layer preceding the bulk solvent could represent the Fc region (Perevozchikova et al. 2015, Zhao et al. 2009).

When building a model to fit to reflectivity data, it is prudent to begin with the simplest possible structure. For IgG a single layer with a maximum thickness of double the molecule's

hydrodynamic radius is a good place to start. Notwithstanding, a single layer model may not produce a good fit to the data, as gleaned from the chi-squared value as well as by visual inspection. Hence, the model structure can be split into two layers in order to achieve a better fit; each layer being allowed to vary within the original SLD and thickness ranges. The process of breaking the structure down into hypothetical layers may be continued until a satisfactory chi-squared value, around 2.0, is obtained. The limits for the roughness parameter depend on the size and character of the molecules in the layer being modelled. For small molecules like SiO₂ a suitable roughness range is between 1Å and 5Å; for IgG, a large, globular, flexible molecule, roughness may reasonably be assumed as great as 20Å.

Where data has been collected at more than one contrast for the same sample configuration and buffer constituents, the two (or more) contrasts can be fitted simultaneously. The two contrasts are fitted to a single layer-depth profile, while the SLD for each contrast is allowed to vary between that of the pure molecule and that of the solvent. For H₂O and D₂O contrasts simultaneous fitting can result in SLD-distance profiles that mirror one another across the x-axis. When data has been gathered for a single structure at a range of solution and/or sample contrasts, fitting these multiple contrasts to a single structural model greatly diminishes the chance of fitting an incorrect model and thus reduces ambiguity in interpretation of the data (Lu et al. 2007, Su et al. 1998, Xu et al. 2006b).

6.6.3.3 Evaluating the model

Once a satisfactory chi-squared value has been reached, the fitting error for each parameter in the model can be evaluated, e.g. the thickness of a layer, or its SLD in D₂O or H₂O. This allows the robustness of fitted parameters to be assessed, helpful for drawing conclusions about the data. Using the RasCal software (Hughes 2013), model-to-fit error estimates are obtained by running the *bootstrap error estimate* function for each fit (Clifton et al. 2012, Clifton et al. 2011). A number of “bootstrap” fits are run, e.g. 200 runs, using randomly selected sub-sets of data points, producing new values for each fitted parameter. At the end of the runs the frequency distribution of the bootstrap fitted values is displayed for each parameter, along with a table of

estimated fit error values. An example of the output from running the error analysis is shown in *Figure 6.5*. Here, the bootstrap fitted profiles in the top plot indicate a fairly robust fit – all the profiles take the same general shape. The frequency distribution plots allow a more detailed look at individual parameters. In *Figure 6.5* the frequency distribution for the SLD of the outermost layer in the H₂O contrast is shown. The distribution is even, narrow and centred on the best-fit value; this indicates a robust fit. The frequency distributions are of particular use when assessing SLD values for H₂O contrasts. This is because the SLDs of biological materials solvated in H₂O often fall close to zero, making numerical error estimates disproportionately large.

6.6.3.4 Adjusting the model

After error analysis, if bootstrap value distributions are broad or irregular for model parameters, it may be necessary to go back and adjust the model. Some parameters hold more value than others. For example scattering length densities should be robust for a reliable fit, as well as layer thicknesses. The roughness parameter is not typically robust; yet, large changes in roughness values may have little effect on the *model's* goodness of fit – in general the parameter is less important.

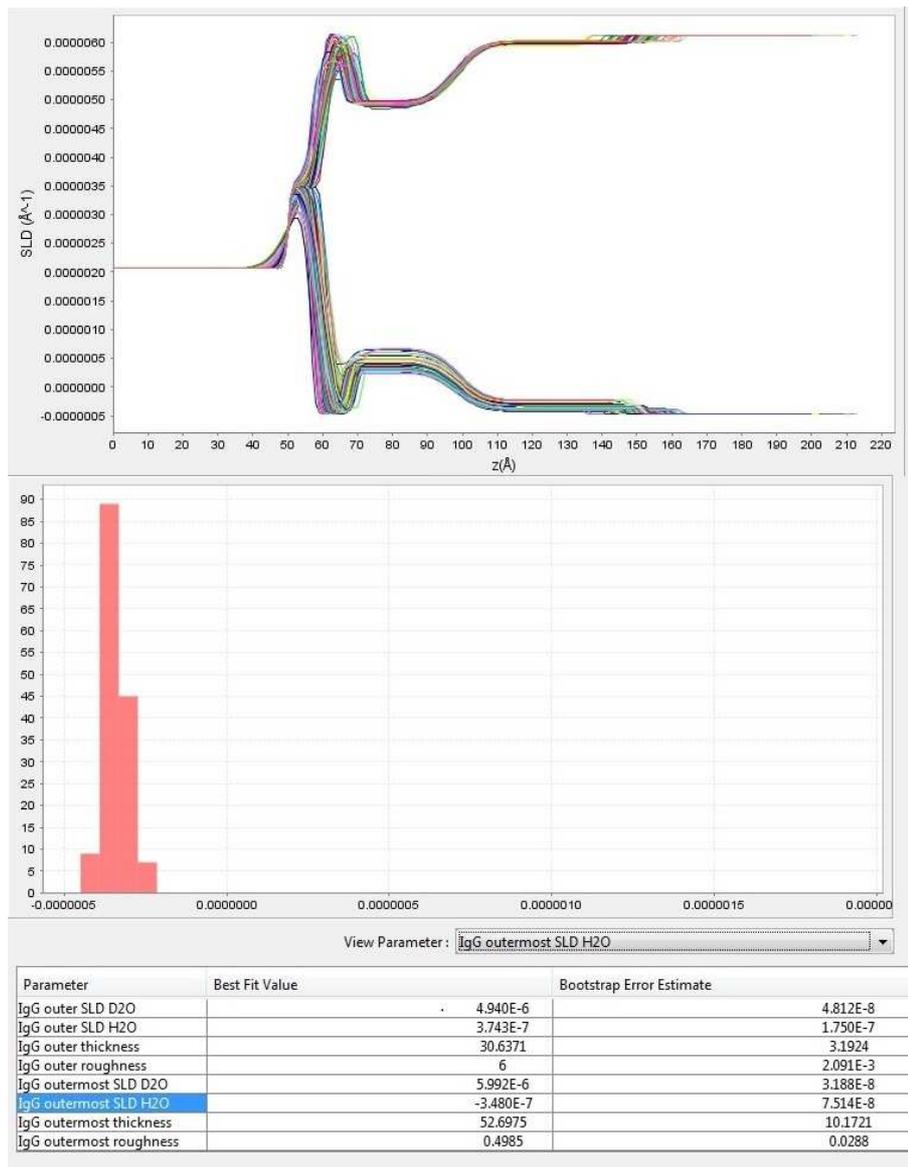


Figure 6.5 Example of output from bootstrap error analysis using RasCal software. Fitted SLD against Qz profiles for all bootstrap runs are shown in a single plot (top); selection of a parameter from the drop-down menu displays the frequency distribution for that parameter (middle); error estimates are displayed in a table (bottom).

6.7 Examining the conformation of IgG4 adsorbed to silica

6.7.1 Sample preparation and stability assessment

In order to produce data comparable to that of Perevozchikova *et al.* (2015), IgG4 was prepared in citrate buffer. The stability of the IgG in citrate buffer was assessed before preparation of the samples for reflectivity experiments. A secondary aim of stability experiments was to determine the lowest pH at which the IgG could be held for prolonged periods without aggregating. It is hypothesised that the structure of adsorbed material may change in response to a reduction in pH, prior to any aggregation events. Low pH is of particular interest due to the “tipping point” mechanism of pH/column-induced IgG4 aggregation documented in *Section 5*. Complete unfolding and aggregation of the material was expected to confound the reflectivity data, so the intention was to identify pH conditions that were not so low as to cause this. The long duration needed for data acquisition on the NG7 horizontal reflectometer, several hours, was also considered. This particular instrument, NG7, was used so that the results would be directly comparable to those of Perevozchikova *et al.* (2015) who used the same instrument. The NG7 reflectometer counts neutrons at a single wavelength so considerable time is required to provide sufficient counting statistics to resolve signal over background counts across the necessary Q range.

The IgG was dialysed into sodium citrate at pH conditions ranging from pH 3.2 to pH 6.3; dialysed samples were then incubated for 5 days either at room temperature or at 4°C. *Figure 6.6* shows the change in monomer concentration over time. At pH 3.2 the monomer fraction was reduced by approximately 15% during dialysis. As buffer exchange was achieved rapidly using multiple centrifugal concentration and re-dilution steps, it is possible that aggregate formation was expedited by periods at elevated IgG concentration. After 5 days at room temperature reduction in monomer was modest at pH 3.8, around 9%, and minimal at pH 4.3, <2%. At 4°C monomer concentration did not change during the incubation period at pH 3.8 and above, and at pH 3.2 only a 2% reduction was observed. It was concluded that the IgG was stable at 4°C, and that minimal aggregation could be expected at pH 3.8 over moderate durations (less than 10

hours) at room temperature. Between pH 3.8 and pH 3.2 a drop off in stability was observed. Subsequently, samples for reflectivity experiments were prepared in D₂O and H₂O-based sodium citrate at pH 4.1 and pH 6.2. Size exclusion chromatography confirmed that rapid buffer exchange in D₂O did not cause the IgG to aggregate or otherwise degrade (data not shown).

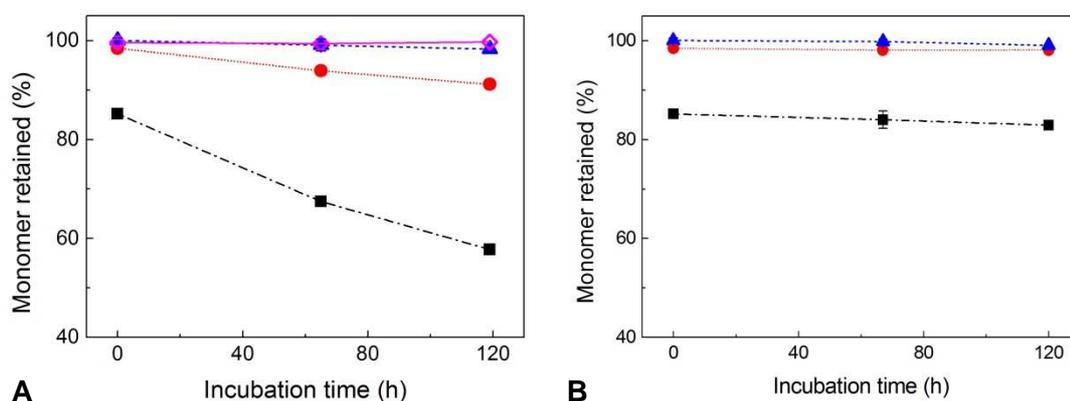


Figure 6.6 Monomer reduction over time at room temperature (a) and at 4°C (b) after dialysis into 30mM sodium citrate. Different symbols represent different pH conditions: open diamonds, pH 6.3; triangles, pH 4.3; circles, pH 3.8; and squares; pH 3.2. Error bars represent the standard deviation for duplicate measurements by SEC.

6.7.2 Conformation of IgG4 adsorbed at pH 4.1

IgG4 at 5.6 mg/mL in D₂O-based citrate buffer at pH 4.1 was flowed into the SLFC (Fig.6.2), mounted on the NG7 reflectometer, over a period of 30 minutes. Neutron counts were collected across the necessary Q range, taking approximately 6 hours; data collection began at the start of sample flow and continued thereafter while the sample remained in the SLFC. Gentle rinsing (low shear) with adsorption buffer was implemented and further neutron measurements taken. A re-adsorption step and a further two rinses were carried out, with neutron measurements taken at each stage. The data was assessed at-line and indicated that rinse and re-adsorption steps did not cause significant changes in the profile of reflectivity against Q, when comparing adsorption and re-adsorption data. Subsequently, a *contrast series* was completed by switching to an H₂O-

based buffer system and implementing adsorption and rinse steps in H₂O. It should be noted that IgG that adsorbed to silica in D₂O and remained after rinsing was not eluted from the cell before the adsorption step in H₂O. The key factor that differentiates *adsorption* and *re-adsorption* configurations from *rinse* configurations is the presence of IgG solution in the SLFC during neutron measurements.

The data was reduced and analysed using the methods described in *section 6.6.3*. Critically, synonymous D₂O and H₂O data-sets were fitted simultaneously, and the bootstrap error analysis method was used to evaluate each fitted parameter.

6.7.2.1 Reflectivity plots

The resultant plots of reflectivity against Q are shown in *Figure 6.7*. The top left-hand plot (*a*) shows IgG adsorption at the silicon-D₂O interface over the sequential adsorption and rinse stages; the lower left-hand plot (*b*) shows adsorption and rinse stages in H₂O-based buffer. Black data points are for the bare silicon wafer immersed in buffer. Here the typical smooth drop in reflectivity as the fourth power of the angle of incidence can be seen. Though not very obvious on visual inspection, some deviation from simulated data (e.g. that shown in *Fig.6.4*) exists due to a layer of silicon dioxide on the surface of the silicon wafer (this forms when the wafer is exposed to air before assembly). For the bare wafer-D₂O interface, the critical edge occurs at a slightly lower Q value than for the subsequent data sets. This was attributed to the presence of a small quantity of hydrogen in the buffer. Over time, hydrogen in the air exchanges with deuterium in D₂O. The buffer used for the initial bare wafer measurements was not freshly prepared, so some exchange had probably occurred. Looking at the fitted curves for the various IgG adsorption and rinse stages in *Fig 6.7.a*, significant deviation from the silicon-water only interface can be seen. A broad fringe of reduced reflective intensity begins at a Q value of around 0.03 and reaches its minima between 0.075 and 0.085Å⁻¹, before an upturn of increased reflective intensity. A second fringe was fitted at the top of the Q range, with a minima around 0.2Å⁻¹. It should be noted that the resolution between signal and background begins to deteriorate in this Q region. H₂O data in *Fig.6.7.b* follows a similar pattern. As well as lacking a

critical edge within the given Q range, the intensity of reflection at the silicon-H₂O interface is around an order of magnitude lower than that for a D₂O liquid phase. Due to the hydrogenous nature of the IgG protein, the presence of protein at the interface causes more limited alterations in the reflectivity profile than for the D₂O-based system. Still, the result of simultaneous fitting of D₂O and H₂O data gives greater insight than the sum of its parts.

Inspecting the data in this relatively raw form, we can see that there is not a great deal of difference across the series of IgG adsorption and rinse steps. Given that no alterations in buffer composition were made across the series (aside from D₂O/H₂O), this was considered a positive outcome. The result indicates that once the initial adsorption event had taken place, a relatively stable layer of material remained on the surface and was not affected by laminar flow of buffer or the presence of 9 mg/mL IgG in the reflectivity cell. This result is consistent with the findings of Perevozchikova *et al.* (2015).

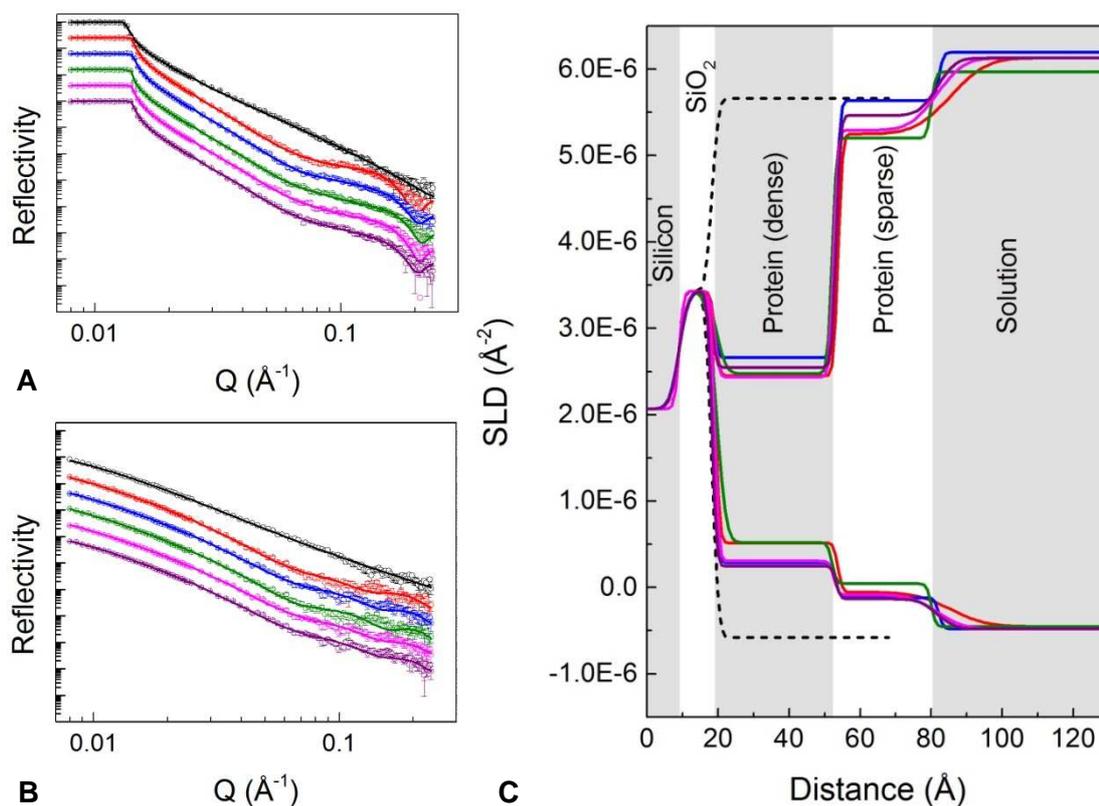


Figure 6.7 Left-hand plots *a* and *b*: reflectivity (relative intensity) against momentum transfer vector, Q (\AA^{-1}), for sequential IgG adsorption and wash steps carried out in D_2O based (*a*) and H_2O based (*b*) sodium citrate buffer. In all cases H_2O buffer was at pH 4.00 and D_2O buffer was at an apparent pH of 4.10, which approximates to pD 4.5. Data sets are offset down the y-axis for clarity; in D_2O the critical edge always occurs when reflectivity is equal to unity. Different colours represent different experimental conditions/stages. From top to bottom: black is the bare silicon wafer; red is IgG adsorption at 5.6 mg/mL; blue is adsorbed IgG after a 6.0 mL rinse; green is IgG re-adsorption at 5.6 mg/mL; pink is re-adsorbed IgG after a 6.0 mL rinse; and purple is re-adsorbed IgG after a second 6.0 mL rinse. Solid lines show the fitted models. Plot *c* shows the corresponding profiles of scattering length density, SLD (\AA^{-2}), against distance from the silicon surface, each generated by simultaneous fitting of D_2O and H_2O reflectivity data sets to a single layer depth profile. Curves that terminate at the upper and lower ends of the SLD scale represent D_2O and H_2O data, respectively.

6.7.2.2 Scattering length density-distance profiles: part 1

Figure 6.7.c shows the SLD against z-direction distance profiles that correspond to the fitted curves in Fig.6.7a and b. For all data sets, a layer of silicon oxide was included in the model; this layer homogeneously coats the silicon substrate, so the SLD value is constrained closely to the predicted value of $3.43 \times 10^{-6} \text{ \AA}^{-2}$ for both D₂O and H₂O data. The thickness of the layer was allowed to vary and typically fell close to 10Å. Looking at the wafer-only curve in Fig.6.7.c, we can see that the SLD of the bulk solvent is lower than that of the other curves; this corresponds to the shifted critical edge (in the direction of lower Q) in Fig6.7.a, and is consistent with the supposition of reduced D₂O content.

For all the other data sets, the lowest chi-squared values were obtained when the model included two distinct protein layers. Note that these can be considered hypothetical layers and could represent either sub-structures of a monolayer, or a bilayer of some description. Focussing on the red line for the D₂O-based system, which represents the first IgG adsorption stage, the following characteristics are observed: a 35Å thick wafer-proximal protein layer with a characteristically low SLD, and a second 32Å thick protein layer with an SLD significantly closer to that of the bulk solvent, D₂O-based buffer. The transitions between SiO₂ and the first protein layer, and between the first and second protein layers, are quite sharp, i.e. the value of *roughness* is low. Conversely, roughness at the interface between the second protein layer and the bulk solvent has a greater value, giving a broader transition. As can be seen in Fig.6.7.c, the remaining data sets have similar characteristics. The thickness of each protein layer varies little between data sets: the first layer ranges from 32Å to 35Å thickness; the second layer from 28Å to 32Å, with 29Å being the most common value. On average, the estimated fitting error for the thickness parameter was less than 5%. SLD values are a little more varied across the data sets, as can be seen in Fig.6.7.c. However, when adjusted for differences in SLD of the bulk solvent, i.e. when the volume fraction of adsorbed material is calculated using *equation 4*, this variation also becomes minimal. Notable differences in layer roughness are observed. However, using the bootstrap error analysis method (see *section 6.6.3.3 Evaluating the model*), we find that fitted values for the roughness parameter are unreliable. The roughness parameter allows construction

of a more realistic model, that is, it provides an alternative to blunt transitions between layers – this would not be realistic for proteins, which are naturally irregular in shape as well as flexible. However, fitting of the roughness parameter is the subject of some contention in the reflectivity community and, as the error analysis shows, it is not a robust parameter. Thus, roughness values may be taken as a guide, but are not helpful for drawing conclusions about the data.

Overall, we can conclude that the different data sets in *Fig.6.7.c* do not differ significantly; hence these curves will be discussed as a single unit of data.

Looking at the first protein layer, the SLD in D₂O indicates a very densely packed layer of protein. The second protein layer is much sparser. *Table 7.4* shows the estimated volume fraction of each protein layer for the first IgG adsorption and rinse stages; these were calculated using *equation 4*. For D₂O data, the volume fractions were estimated at three levels of hydrogen-deuterium exchange in IgG4, 50%, 70% and 90% (percentages of *labile* hydrogens). For H₂O data, volume fractions were estimated at 0% exchange, i.e. 100% H₂O. The volume fraction approximations represent maximum values for the midpoint of each layer; they do not account for any gradient in SLD between layers (roughness). The volume fraction values for the inner layer of protein highlight an unexpected result of the data fitting – they indicate protein volume fractions greater than 1. This seems physically impossible. Thus, the data was further probed for an explanation. Focussing on the data for initial IgG adsorption, when a plot of volume fraction against percentage H-D exchange is extrapolated, we find that for a volume fraction of 0.95, only 30% exchange can occur if the fitted SLD value of $2.45 \times 10^{-6} \text{ \AA}^{-2}$ is true (*Table 6.4*). These seem extreme values for both volume fraction and H-D exchange level. Yet, remarkably similar results for a mAb adsorbing to SiO₂ were obtained by Couston et al. (2012), though their data describes a protein layer averaging around 3 times the thickness.

Table 6.4 Theoretical SLD values for IgG at different H-D exchange levels; scattering length densities of the mobile phase and inner and outer protein layers from fitted models for IgG adsorption data, and estimated maximum protein volume fractions for each layer.

		SLD ($\times 10^{-6} \text{ \AA}^{-2}$)						Volume fraction				
Protein Layer		Mobile phase		Protein H-D exchange				Fitted layer	D ₂ O			H ₂ O
		D ₂ O	H ₂ O	90%	70%	50%	0%		90% ex.	70% ex.	50% ex.	0% ex.
pH 4.1 IgG adsorption	Inner							2.45	1.26	1.15	1.05	0.41
	Outer	6.13	-0.45					5.25	0.30	0.27	0.25	0.17
pH 4.1 rinse 1	Inner							2.66	1.18	1.08	0.99	0.32
	Outer	6.20	-0.48	3.20	2.92	2.64	1.94	5.64	0.19	0.17	0.16	0.15

The estimates of volume fraction suggest an adsorbed layer with potentially un-realistic characteristics, that is, very densely packed protein and a very low level of H-D exchange. However, a number of possible explanations present. First, hydrogen-deuterium exchange is strongly influenced by pH. For amide hydrogens, exchange is at its slowest around pH 3.0 (Raschke and Marqusee 1998), while the lowest exchange rates for hydroxyl side chains lie between neutral and moderately acidic pH (Liepinsh and Otting 1996). In studies examining exchange rates for proteins, experiments are often carried out at neutral pH, so at pH 4.1 the exchange rate may be lower than previously reported. Further, the degree of protection against exchange for a given region of a protein depends on its folded structure, including tertiary as well as secondary structure, and can change as a result of external factors such as solution chemistry and temperature (Liepinsh and Otting 1996, Raschke and Marqusee 1998). According to Raschke and Marqusee (1998), highly protected amide hydrogens may require months to years to exchange. The IgG4 used in this work was prepared in D₂O-based buffer approximately one week before reflectivity experiments commenced. Buffer exchange was carried out at 4°C and samples were stored and transported at 4°C thereafter. Thus, it is possible that the level of exchange was much lower than expected. Another possibility is that some hydrogenous material became “trapped” at the silicon-protein interface. In similar reflectivity experiments, Couston et al. (2012) report that improved chi-squared values were obtained when a thin layer with an SLD $\sim 0.5 \times 10^{-6} \text{ \AA}^{-2}$ was added to the model at the SiO₂-protein interface. Couston et al. suggest that

this represents a sparse hydrogenous coating, possibly residual after wafer-cleaning which is typically done using a Piranha etch solution (see *section 4.3.5.2 Reflectivity cell*). If such a layer existed in our own reflectivity work it is possible that, if not placed in the model as a defined layer, its presence at the interface could bring down the fitted SLD of the first (protein) layer located adjacent to the SiO₂ surface.

The volume fraction calculated for the inner protein layer from H₂O data averaged 37%, less than half that from D₂O data, while volume fractions for the outer layer were similar for both contrasts.

As well as apparent total saturation of the surface with protein, a striking feature of the data is the very thinness of the layer(s) of adsorbed material. The total depth of adsorbed material had a maximum of less than 65 Å; roughly the upper half of this depth had a much smaller volume fraction, about one quarter the density of the lower layer. These values can be interpreted as an IgG orientation in which maximum contact with the surface has occurred. It is possible that what we have is detailed structural information on a highly homogenous monolayer of IgG4.

Neutrons structural characterisation of proteins in D₂O yields *unhydrated* protein dimensions. The un-hydrated volume of wild-type IgG4 has been calculated as 188.7 nm³ (Rayner et al. 2014). If the volume is assumed spherical, the radius equals 3.6 nm (36 Å). This value is very close to that observed for the lower layer fitted for adsorbed IgG.

A comprehensive study of the solution structure of IgG4 was undertaken by Rayner et al. (2014). Ultracentrifugation, x-ray and neutron scattering, and atomistic modelling methods were used to characterise the structure of wild-type IgG4 in solution. The two scattering methods were used to determine the radius of gyration, R_g , for IgG. R_g is a measure of “structural elongation” so, instead of assuming the protein spherical we consider an ellipsoid, which is a more accurate representation of antibody shape. An ellipsoid is elongated and can therefore be described in terms of cross-section and length. The mean radius of gyration of the two shorter axes is denoted R_{xs} , while R_g gives the average about all three axes. For immunoglobulins, R_{xs} is split into two components, R_{xs-1} and R_{xs-2} . The average cross-sectional radius is described by R_{xs} .

z , and R_{xs-1} is a measure of Fab and Fc separation (Rayner et al. 2014). Rayner et al. (2014) characterised IgG4 at different protein concentrations and temperatures and in buffers of various molarities; the conditions included ones similar to those investigated in *this* work. X-ray data from Rayner et al. (2014) gave values averaging 2.53nm and 1.40nm for R_{xs-1} and R_{xs-2} , respectively. Their neutron data determined R_{xs-1} and R_{xs-2} values of 2.51nm and 1.19nm, respectively; the smaller values from neutrons was attributed to the “invisibility” of the hydration shell (Rayner et al. 2014), i.e. the hydration shell is matched to the bulk solvent. These values are clearly echoed in our own data. For neutrons, doubling R_{xs-2} gives a 24Å cross-sectional diameter, and doubling R_{xs-1} gives 50Å.

With supporting evidence from the data obtained by Rayner *et al.* (2014), as well as some agreement with work by Perevozchikova et al. (2015) our fitted SLD against distance profiles strongly suggest an IgG4 monolayer that is orientated largely “flat on”, i.e. with R_{xs-2} perpendicular to the surface. We suggest that the sparser outer IgG4 layer (*Fig 6.7.c*) represents single Fab arms that are extended into the solution phase. The much lower volume fraction could be because not all IgG molecules have a Fab arm extended in this way, or because the extended arms are orientated perpendicular to the surface, so occupy a smaller area in the x-y axis. Similar explanations are given by Perevozchikova *et al.* (2015); in both cases, there is not enough data to make a firm conclusion in this regard.

We have explored a range of potential explanations for the more improbable aspects of the fitted reflectivity profiles for IgG adsorption at pH 4.1, namely, the apparent volume fraction of 1.0 for the first protein layer. In many cases we look to published work for answers. However, regarding the data presented in this work, considerable effort was made to select the most robust fits out of a large number of trials. One such attempt resulted in considerably lower chi-squared values than those for the fit described above. The second gauge of goodness of fit was the outcome of the *bootstrap error analysis* (see *section 6.6.3*).

6.7.2.3 Scattering length density-distance profiles: part 2

Exploring different fits for the data resulted in an improved fit. The results of this are shown in *Figure 6.8*. Only the first IgG adsorption and rinse stages are shown, as subsequent re-adsorption and rinse stages were not found to impart any structural re-arrangements, as discussed in the previous section. Looking carefully at the plots in *Figure 6.8.a* and *6.8.b*, the major observable difference in the curve fit (compared to *Fig. 6.7.a-b*) is in the high Q region. The newly fitted curve takes a straight route through the middle of the points on its downturn from the fringe; in the previous fit the downturn was steeper and was followed by an upturn at the very highest end of the Q range. The new fit is interpreted as a better fit and indeed the chi-squared values were reduced from 2.66 and 2.78 to 1.81 and 1.82 for the adsorption and rinse stages, respectively.

The improved fit was achieved by adding an additional layer to the model, in a similar initiative to that described by Couston et al. (2012). As can be seen in *Figure 6.8.c*, the layer was placed proximal to silicon dioxide. The layer was constrained to a minimum thickness of 0Å and a maximum of 150Å. For each of the two contrasts, the SLD was constrained between that for a pure protein layer and that for a pure solvent layer. In this way, the layer was given considerable freedom. Results were similar for the two data sets (IgG adsorption and rinse stages). The new layer had a thickness of 7-8Å and an SLD approaching that of the bulk solvent for each contrast. For the clarity's sake, this layer will be designated "wafer-proximal", and the other layers will keep the same designations as before – *IgG inner* and *IgG outer*.

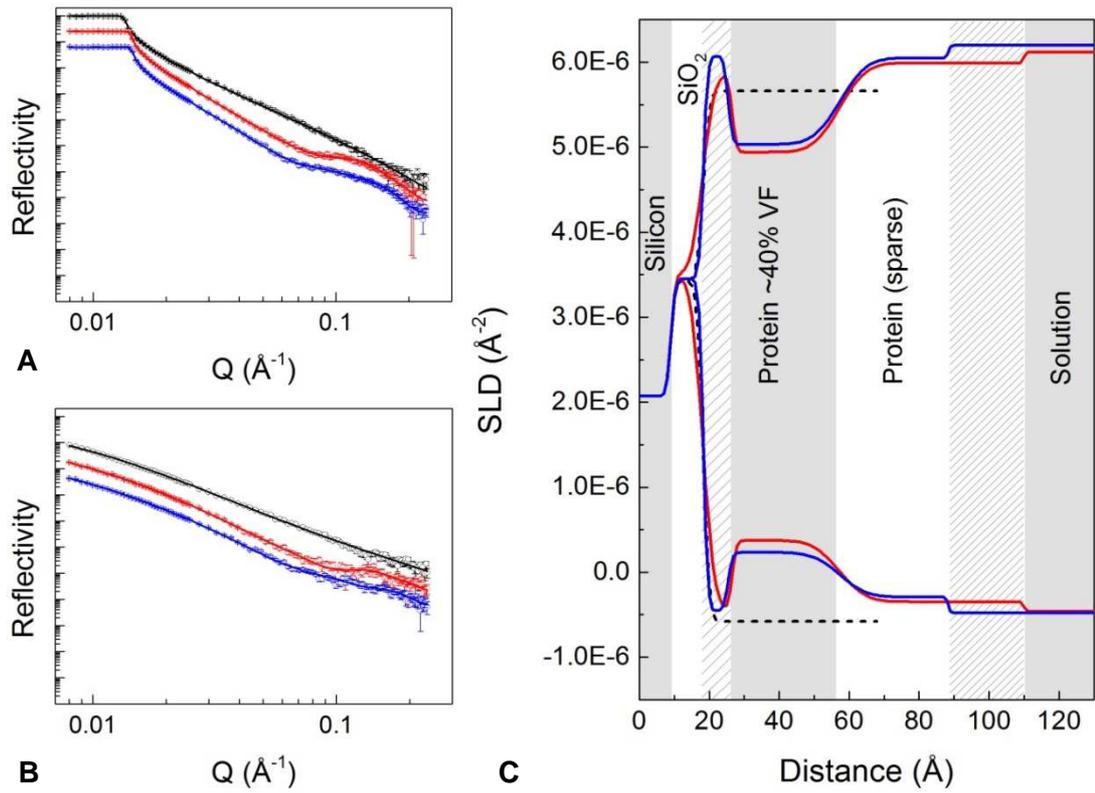


Figure 6.8 The data in *Figure 6.7* for IgG adsorbed at 5.6 mg/mL (red), and adsorbed IgG after a 6.0 mL rinse (blue) was re-fitted giving improved chi-squared values and good physical feasibility. An additional layer located adjacent to the SiO_2 layer was added to the model.

For the inner and outer IgG layers in the model, the most significant changes were in the SLD values for the D₂O contrast. It is noted that the SLD values for the D₂O and H₂O contrasts mirror one another; this is an indication of a good fit (in the absence of isotope effects within the adsorbed material) (Su et al. 1998). For the IgG adsorption stage, the SLDs suggest a maximum of 40% surface coverage for the inner IgG layer and 4% for the outer IgG layer. After surface rinsing the inner layer had 39% coverage and the outer layer 5% coverage. These volume fraction values assume 90% of labile hydrogens were exchanged to deuterium. The values seem more realistic than those from the first fit. If 50% exchange is assumed, the coverage estimates for the dense layer decrease by 6%, and the sparse layer is barely altered.

The inner layer had a thickness of 31-32Å before and after rinsing, these values are similar to those from the first fit. A clear difference between the SLD-distance profiles in *Fig. 6.8* before and after rinsing is in the thickness of the outer IgG layer. Before rinsing the outer layer had a thickness of 53Å and after rinsing the thickness was reduced to 31Å.

The fitting error for each parameter was assessed using the bootstrap error estimate method described in *section 6.6.3*. The results of this for the main parameters of the model of adsorbed IgG before rinsing are displayed in *Figure 6.9*. We can see that the SLDs for the inner and outer protein layers were robust (*Fig 6.9.a-b*). The SLDs for the wafer-proximal layer (*Fig. 6.9.c-d*) were not as robust but still gave reasonable distributions, with the majority of bootstrap values falling fairly close to the actual fitted values. Though the SLDs of the wafer-proximal layer were somewhat variable, its thickness was well-fitted (*Fig 6.9.e*). This supports inclusion of the wafer-proximal layer in the model. The inner protein layer was highly robust and the outer protein layer showed an even but slightly broader distribution (*Fig. 6.9.e*). This could suggest a more dynamic behaviour in the outer protein layer. *Figure 6.9.f* shows the bootstrap value distribution for the silicon dioxide layer roughness. This is an example of a poorly fitted parameter. Though, the parameter may appear more variable because it is constrained within a small length range. As mentioned previously, the roughness parameter is often not robust, but changes in roughness do not have large knock-on effects on other parameters.

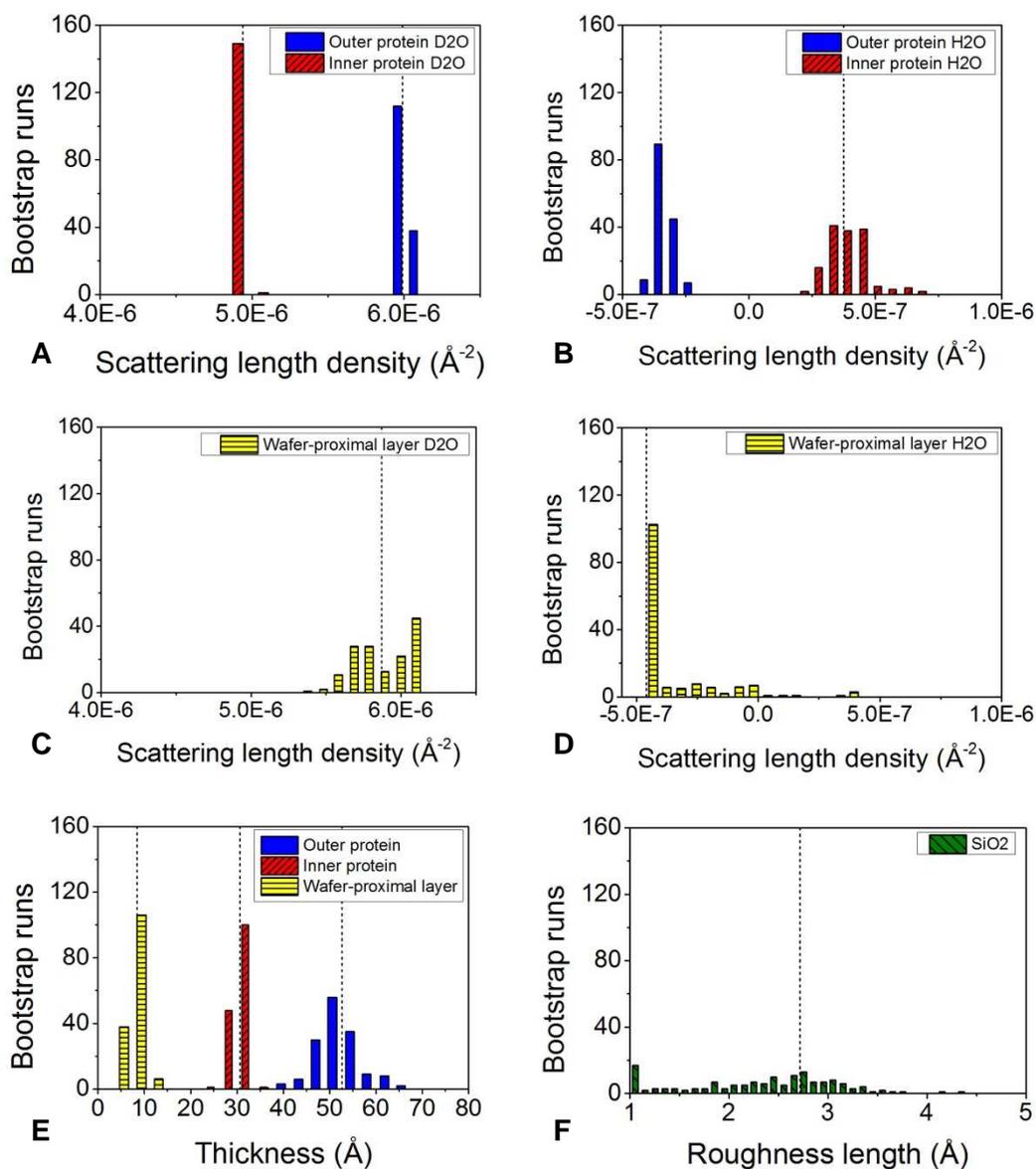


Figure 6.9 Graphical output of bootstrap error analysis for selected parameters from the re-fitted SLD profile of IgG adsorbed at pH4.1, seen in *Figure 6.8* (red curve, before rinsing) Actual fitted parameter values are shown by vertical dashed lines. The total number of bootstrap runs for each parameter was 150. Bootstrap distributions are shown for inner and outer protein layer SLDs in D₂O (a) and H₂O (b), the wafer-proximal layer SLD in D₂O (c) and in H₂O (d), layer thicknesses (e) and the roughness of the SiO₂ layer (f). Robust fitted parameters are indicated by a narrow and normal distribution across bootstrap runs, e.g. (a) – inner/outer protein SLD in D₂O; less robust parameters result in broad and irregular distributions, e.g. (f) – SiO₂ roughness.

It is possible that the wafer-proximal layer represents a hydration layer at the solid-liquid interface. This layer could comprise water molecules interacting with silicon dioxide, as well as water associated with the IgG. We also note that the volume fraction of the wafer-proximal layer is the same as that of the outer IgG layer, though the SLDs of the wafer-proximal layer were less robust. This could be due to charge interactions between hydrated residues and SiO₂ at the surface. Another possibility is that the IgG is lying face down in such a way that it makes only point contacts with the surface and encloses a solvent-filled cavity, resulting in an apparently mostly water-filled layer close to the surface.

One interesting aspect of the data is the change in thickness in the outer protein layer after rinsing. Rayner et al. (2014) estimated IgG4 to have an R_g of 48Å at 20°C. The data presented in *Figure 6.8* gives a total IgG4 thickness of 86-92Å (depending on whether we include the wafer-proximal layer) before rinsing. Halving this gives 42-46Å, values not much smaller than the R_g reported by Rayner et al. (2014). This quite strongly supports the hypothesis that a fraction of IgG molecules were weakly associated with the surface while the IgG solution was inside the SLFC. Weak association would allow molecules more rotational freedom, resulting in depth values closer to double the R_g , rather than the R_{xs-1} or R_{xs-2} (Rayner et al. 2014) are associated with a more static flat-on orientation. It is difficult to manually estimate the proportion of IgG molecules that make up the weakly associated population, because the volume fraction contribution from IgG in the surface normal direction depends on its orientation. The data suggests reorganisation of the weakly associated fraction on rinsing with buffer. Some molecules may be removed while others may re-adsorb in the flat-on orientation, overall resulting in little change to the layer volume fractions (*Fig.6.8*).

When adsorbed proteins demonstrate a flat-on orientation, the calculated volume fraction values can be used to estimate the quantity *surface excess*, that is, the mass of adsorbed material per unit area, which may in turn be used to estimate the average area occupied by each adsorbed protein molecule. This lends a strong insight into surface packing density (Zhao et al. 2011). The following equation is used to estimate surface excess, Γ

$$\Gamma = \phi_l \cdot \tau \cdot \rho'_p$$

Equation 5

where ϕ_l is the volume fraction (see *equation 4, section 6.6.3 Data analysis methodology*), τ is the layer thickness (cm) and ρ'_p is the protein mass density ($\text{g}\cdot\text{cm}^{-3}$) (Xu et al. 2006b).

The protein mass density of the IgG was assumed $1.42 \text{ g}\cdot\text{cm}^{-3}$ based on an estimate for an antibody described by Xu et al. (2006b). If the inner layer of IgG adsorbed at pH 4.1 is assumed to represent a monolayer of IgG in a flat-on orientation, the surface excess can be calculated: the resulting value is approximately $1.7 \text{ mg}\cdot\text{m}^{-2}$. Given the molecular weight of the IgG4 is 165kDa, this equates to an area of 16120 \AA^2 per IgG molecule. The minimum area required for an IgG4 to adsorb flat-on is approximately 13600 \AA^2 (Rayner et al. 2014, Xu et al. 2006b). Based on these values, the data suggests a closely packed layer of adsorbed proteins with each IgG molecule occupying around 1.2 times its minimum required area.

6.7.3 Conformation of adsorbed IgG4 under chromatographic conditions

Having generated data for IgG4 adsorbed to silica under conditions similar to those studied by Perevozchikova et al. (2015), the next step was to assess adsorption behaviour in buffers more relevant to affinity chromatography. That is not to say that adsorption at pH 4 is not relevant, because as was found in *chapter 5* of this work, elution from protein A using standard buffers for a bioprocess involved a rather long transition from the pH of the loading buffer to the pH of the elution buffer, with large portions of the elution pool exiting the column at around pH 4.0.

A new SLFC was prepared and neutron measurements taken for the bare silicon surface in D_2O -based and H_2O based 0.03M sodium citrate, pH 6.2. The cell was re-equilibrated in the D_2O contrast before injection of IgG4. The IgG solution remained in the reflectivity cell during neutron measurements. The IgG solution was then rinsed away with the same buffer and the surface re-analysed. Finally, the surface was washed with citrate buffer at pH 3.7. The adsorption and rinse stages were repeated in the H_2O contrast.

The reflectivity data that was collected is displayed in *Figure 6.10*. The fitting yielded suitably low chi-squared values of 1.76, 2.21 and 2.72 for adsorption at pH 6.2, rinsing at pH 6.2, and rinsing at pH 3.7, respectively. The bootstrap error estimates showed that the fits were not as robust as those for pH 4 adsorption. The SLD against distance profiles in *Figure 6.10* show larger layer roughness values than in the pH 4.1 data. In this case, the roughness may provide a useful insight. Compared to the profiles for pH 4.1, the wafer-proximal layer is thicker, rougher, and has SLD values that suggest a lower solvent fraction. This strongly suggests that all three layers in the pH 6.2 profiles in *Fig 6.10.c* represent elements of the IgG structure. When IgG4 was in the SLFC the total thickness of the adsorbed layer was 70Å. The middle layer had a maximum volume fraction of 0.45 and was flanked by layers each with volume fractions of 0.16. This is suggestive of a side-on orientation of IgG, where one Fab arm may partially contact the surface while the other extends into the solution; the Fc would be parallel to the surface, giving a higher volume fraction.

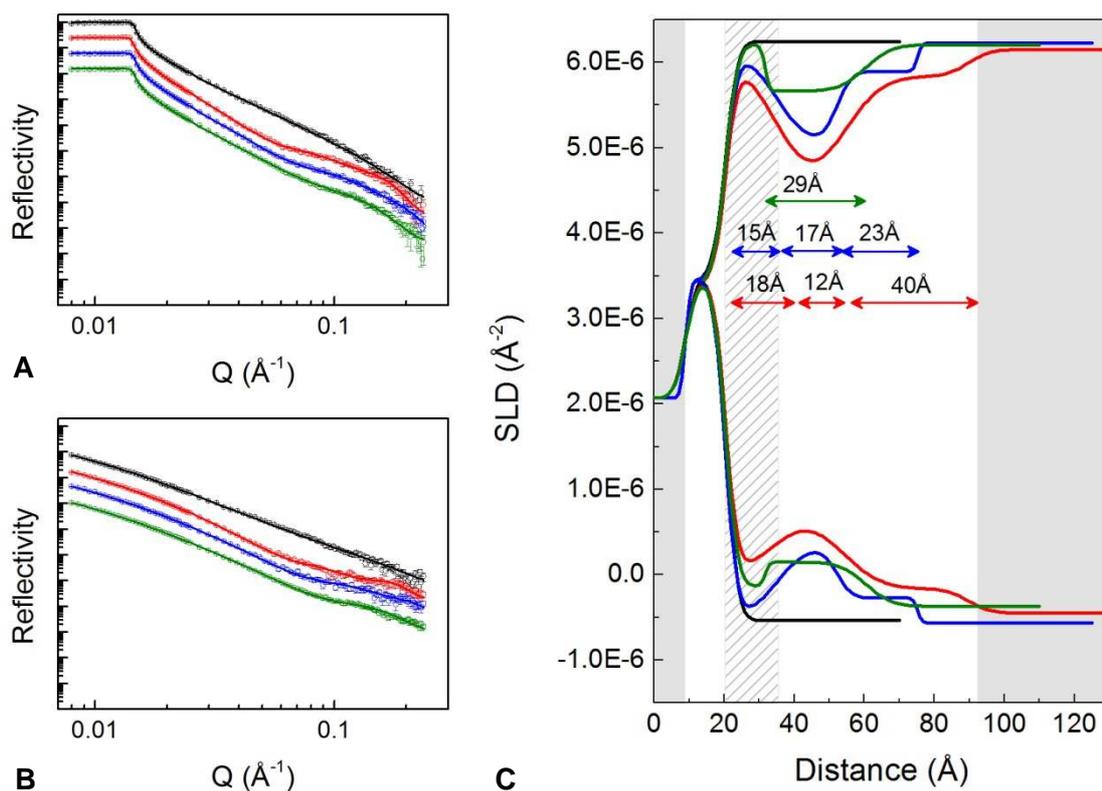


Figure 6.10 Left-hand plots *a* and *b*: Reflectivity against momentum transfer vector, Q (\AA^{-1}), for sequential IgG adsorption and wash steps carried out in D_2O based (*a*) and H_2O based (*b*) sodium citrate buffer. Data sets are offset down the y-axis for clarity; in D_2O the critical edge always occurs when reflectivity is equal to unity. Different colours represent different experimental conditions/stages. From top to bottom: black is the bare silicon wafer; red is 5.6 mg/mL IgG adsorption at pH 6.2; blue is adsorbed IgG after a 6.0 mL rinse with pH 6.2 buffer; green is after a subsequent 6.0 mL rinse with pH 3.7 buffer. Solid lines show the fitted models. Plot *c* shows the corresponding profiles of scattering length density, SLD (\AA^{-2}), against distance from the silicon surface, each generated by simultaneous fitting of D_2O and H_2O reflectivity data sets to a single layer depth profile. Curves that terminate at the upper and lower ends of the SLD scale represent D_2O and H_2O data, respectively. Layer thicknesses in each SLD-distance profile are shown by double-ended arrows in colours that correspond to curve colours.

After pH 6.2 rinsing, the SLD-distance profile retained roughly the same shape but the surface distal layer became considerably shorter. Overall the pH 6.2 data suggests a less flat orientation with a dense body of protein further away from the surface than at pH 4.1. The larger volume fraction for the densest protein layer suggests closer packing of IgG molecules. This agrees with the expected behaviour of a protein as the pH moves closer to its isoelectric point. That is, charge repulsion between the molecules is reduced. Barnett et al. (2015) found this effect to be pronounced for an IgG1 molecule in citrate buffer compared to acetate. They hypothesised that accumulation of citrate ions on the protein surface had a charge screening effect (Barnett et al. 2015). Xu et al. (2006b) carried out a similar study, characterising a mouse IgG1 adsorbed at the silica-water interface. The IgG was allowed to adsorb at a range of concentrations at pH 7.0 – this pH was higher than, but quite close to the pI of the IgG1, meaning that the IgG carried a modest negative charge (Xu et al. 2006b). Despite the potential for repulsion between the negatively charged IgG and the silica surface, at the lowest IgG concentration (2 mg/L) the data was best fitted to a model comprising a single 40Å thick protein layer with a low volume fraction of 0.15 – this is indicative of a flat on IgG orientation. When the IgG concentration was increased to 10 and 50 mg/L the data revealed significant changes in orientation and packing density of the adsorbed protein molecules. Three-layer models were found to fit the data best. At 50 mg/L layer depths were 14Å, 35Å and 35Å and volume fractions were 0.3, 0.44 and 0.9 for inner, middle and outer model layers, respectively. The overall configuration of the model is highly similar to that generated in this work for IgG4 adsorbed at pH 6.2. Xu et al. (2006) suggest that their model represents antibody that has moved from a flat on to a partially tilted orientation. The authors note that electrostatic repulsion between the silica surface and the IgG may help to maintain the antibody's globular structure and prevent adsorption-related structural deformation. Conversely, in this work IgG4 carries a moderate positive charge at pH 6.2 and the IgG sublayers are somewhat thinner than those seen by Xu et al. (2006b). This could indicate a degree of structural deformation.

Washing at pH 3.7 caused significant changes in the SLD-distance profile (*Fig. 6.10.c*). The shape of the profile is very similar to that of IgG adsorbed at pH 4, but lacks the outermost

layer. This suggests that the IgG was strongly adsorbed to the surface in a completely flat on orientation. It is interesting to note that a significant portion of the IgG, though apparently weakly adsorbed at pH 6.2, appeared to re-orient and become strongly adsorbed when the pH was lowered. It is hypothesised that the volume fraction at pH 3.7 was lower because the surface association of proteins at pH 6.2 was weaker, meaning that some molecules were simply washed away and or/charge repelled when the buffer was changed. Inter-layer protein-protein charge repulsion has been shown to influence the packing density of proteins adsorbed at solid-liquid interfaces (Lu et al. 2007, Su et al. 1998). A study by Xu et al. (2006a) found that the lateral packing density of antibody molecules adsorbed to a silica surface decreased as the pH moved away from the antibody's pI. In the case of the work presented here, it is important to note that the flow rate for rinsing was very slow – 0.15mL/min. This means that the pH transition may have been more gradual than sharp due to mixing at the slow-moving buffer front. This may have encouraged proteins to associate closely with the surface before protein-protein charge repulsion took effect.

6.8 Conclusions

In this chapter neutron reflectivity was used to examine adsorption behaviour of IgG4 on a silica surface. The objectives were to first produce data comparable to that of Perevozchikova et al. (2015) by examining IgG4 adsorption under the moderately acidic condition pH 4.1. The data acquired agreed with that of Perevozchikova et al. (2015) in terms of the predicted orientation of the IgG, that is, largely flat on. The first fitting approach suggested that the structure of the adsorbed layer was not affected by rinsing at the same pH, and no further IgG adsorption took place after the first IgG application. This was very much in agreement with findings of Perevozchikova et al. (2015). From the first fitting approach the data suggested a volume fraction of 1.0 for the layer of material closest to the surface. This finding was very similar to that of Couston et al. (2012) who observed a very dense layer of adsorbed IgG and suggested 50% retention of labile hydrogens to make their fitted SLD values physically feasible. For the

data presented in the work here using the first model, 70% retention of labile hydrogens would be required to give a feasible volume fraction. Consequently different model fits were explored. An excellent fit with good feasibility was achieved by fitting an additional layer proximal to the silicon surface. At pH 4.1 the wafer-proximal layer was suggested to represent surface and protein hydration and had a thickness of 8Å. The new fits suggested a 31Å thick inner IgG4 layer with 40% surface coverage and an outer layer with 5% coverage. Before surface rinsing, when IgG solution was still in the SLFC, the depth of the outer layer was 53Å and after rinsing it dropped to 31Å. It is hypothesised that this represented a reorganisation of a fraction of the adsorbed material from a weakly adsorbed orientation to a more strongly adsorbed orientation. A similar result was reported by Perevozchikova et al. (2015). An intriguing characteristic of the adsorbed material was that the layers were rather thin – approximately 30Å each after rinsing. In this way, the data differed from that of Perevozchikova et al. (2015), who observed thicker layers. The layer thicknesses were compared to values of IgG4 of cross-sectional diameter, R_{xs-2} , obtained by Rayner et al. (2014) using small angle neutron scattering. These values were highly comparable to the layer thickness fitted here. IgG layer thicknesses were also comparable to those obtained by Xu et al. (2006a) and Xu et al. (2006b).

A second objective was to assess IgG4 adsorption behaviour under pH conditions closer to those which would be used in affinity chromatography. Experiments were done in a new SLFC. IgG was allowed to adsorb to the surface at pH 6.2, and then rinsed at the same pH and then at pH 3.7. The SLD-distance profiles for pH 6.2 adsorption indicated a different orientation to that observed at pH 4.1; a tilted, rather than completely flat-on, orientation was proposed. The data also indicated that IgG4 adsorption may have been weaker at pH 6.2, and that antibody packing was slightly denser – these are expected findings for an increase in pH towards the protein's pI (Barnett et al. 2015, Su et al. 1998). Work carried out by Su et al. (1998) used neutron reflectivity to study the structure of bovine serum albumin (BSA) adsorbed at the silica-water interface under various pH conditions. Their findings suggested that below the pI of BSA, electrostatic attraction between the negatively charged silica surface and positively charged BSA molecules was the driving force for adsorption. Less BSA adsorbed when the pH was

adjusted further from the pI; it was suggested that this was due to lateral repulsion between adsorbing BSA molecules (Su et al. 1998, Su et al. 1999). In the work presented here all pH conditions tested were below the pI of the IgG, meaning that IgG was positively charged to varying degrees for the various experiments. The structures obtained for adsorbed IgG at pH 4.1 and pH 6.2 support the theory that though electrostatic attraction is a strong driving force for adsorption, intermolecular repulsion limits packing density as adsorbing molecules become more charged. Reducing the buffer pH from 6.2 to 3.7 inside the SLFC caused a change in the interfacial structure. The SLD-distance profile suggested re-orientation of the IgG to a strongly adsorbed, completely flat-on configuration with a lower volume fraction. It was interesting to observe the distinct shift in configuration towards increased surface contact as the pH was lowered. The data highlights the importance of protein-protein charge repulsion in determining the packing density of interfacially adsorbed proteins, supporting the findings of various other studies (Su et al. 1998, Xu et al. 2006a, Xu et al. 2006b, Zhao et al. 2012). Many of these studies have also found individual adsorbed proteins to apparently occupy surface areas smaller than their predicted molecular footprints, despite relatively short layer depths. Such findings are indicative of partially tilted/flat, closely packed adsorbed proteins. Typically this orientation has occurred under pH conditions close to the protein's pI. These observations are in agreement with the findings presented in this work.

7 Characterisation of IgG4 Bound to Immobilised Protein A

7.1 Introduction

In this section neutron reflectivity is used to probe the structure of a model system more representative of a protein A resin than that explored in *Chapter 6*. Staphylococcal protein A was covalently attached to a flat silicon wafer. In preliminary experiments the immobilisation process was monitored at each stage using attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR). Subsequently, neutron reflectivity was used to examine protein structure on the surface at various stages from protein A immobilisation through IgG adsorption, and after washing the surface at low pH. Experiments were carried out at the ISIS Neutron and Muon Source, located at the Rutherford Appleton Laboratory, Didcot, UK. ISIS is a *spallation* neutron source, not a reactor (which is the source at the NCNR where the NG7 reflectometer is located). At spallation sources a “time of flight” method is used to determine the wavelength of each neutron that hits the detector from a white neutron beam containing multiple wavelengths (Clifton et al. 2013). Since Q is proportional to incident angle and wavelength (see *section 6.6.1*), simultaneous exploitation of multiple wavelengths in the white beam allows a broader Q range to be covered in a shorter time than at reactor sources. The upshot is that a greater number of experiments are possible within a given time on the instrument. This increased capacity allowed generation of a large data set, assessing multiple sample configurations using three solution contrasts and across three different sample cells.

7.2 Chapter objectives

In the previous chapter neutron reflectivity was used to examine the structure and orientation of IgG4 adsorbed directly to a silica surface at pH 4.1, pH 6.2 and pH 3.7. The latter buffer conditions were chosen for their relevance to protein A chromatography. The data generated seemed robust and shared a number of features with similar types of data produced by other authors. Notably the dimensions of adsorbed IgG4 ascertained here were in good agreement with those obtained for another wild-type IgG4 molecule by Rayner et al. (2014) using small angle neutron scattering (SANS). They also agreed largely with those of IgG adsorbed to silica obtained by Xu et al. (2006a), Xu et al. (2006b) and Zhao et al. (2012) using neutron reflectivity. This promoted confidence that the neutron reflectivity method could be used to gather detailed information on the structure of adsorbed IgG4 without perdeuteration of the protein.

This chapter aims to structurally characterise IgG4 adsorbed to immobilised protein A. Of particular interest are the orientation of the IgG and its proximity to the solid support surface. IgG4 will be adsorbed and characterised under the same buffer conditions as those used in the chromatography experiments described in *Chapter 5*.

7.3 Generating the model system

The model system was built “from scratch” starting with a standard polished silicon wafer designed for the reflectivity rig.

7.3.1 Cross-linker chemistry

The cleavable cross-linker sulfosuccinimidyl 6-(3'-(2-pyridyldithio)propionamido)hexanoate (sulfo-LC-SPDP) was used to attach protein A to the silica surface via a thiol linkage. Sulfo-LC-SPDP is one of three commercially available variants of this cross-linker; LC-SPDP and SPDP also exist. LC-SPDP has a longer chain than standard SPDP, and the sulfonate group

(sulfo-) confers solubility in aqueous solutions (Hermanson 2013b). Henceforth, *sulfo-LC-SPDP* will also be referred to simply as *SPDP*.

SPDP can be used to couple cysteine-containing proteins and cysteine-free proteins to surfaces via a cleavable thiol linkage. For cysteine-free proteins, an additional protein modification stage is necessary.

SPDP contains two reactive groups. One is the N-hydroxysuccinimide ester (NHS ester) group which reacts with primary amines to form amide bonds (*Fig. 7.1b*). SPDP reacts readily with proteins at N-terminal amines and side-chains of accessible lysine residues. It has also been shown to react with arginine side-chains under conditions where the side-chain is not protonated, i.e. highly basic conditions (McGee et al. 2012). The third basic amino acid, histidine, can react with the NHS ester but the acyl imidazole group formed is rapidly hydrolysed and so the cross-linker lost (Hermanson 2013a).

The reactive group at the other end of the chain in SPDP is 2-pyridyldithiol, which reacts with sulphhydryl groups to form disulphide bonds and a non-reactive pyridine 2-thione by-product (*Fig. 7.1b*). The reaction is termed thiol-disulphide exchange and is one-directional (Hermanson 2013a). If the protein to be conjugated contains cysteine, the disulphide exchange reaction can be achieved after mild reduction of the protein solution. If the SPDP linker is first coupled to the surface via the NHS ester group, protein conjugation can be achieved in a second step targeting cysteine residues only.

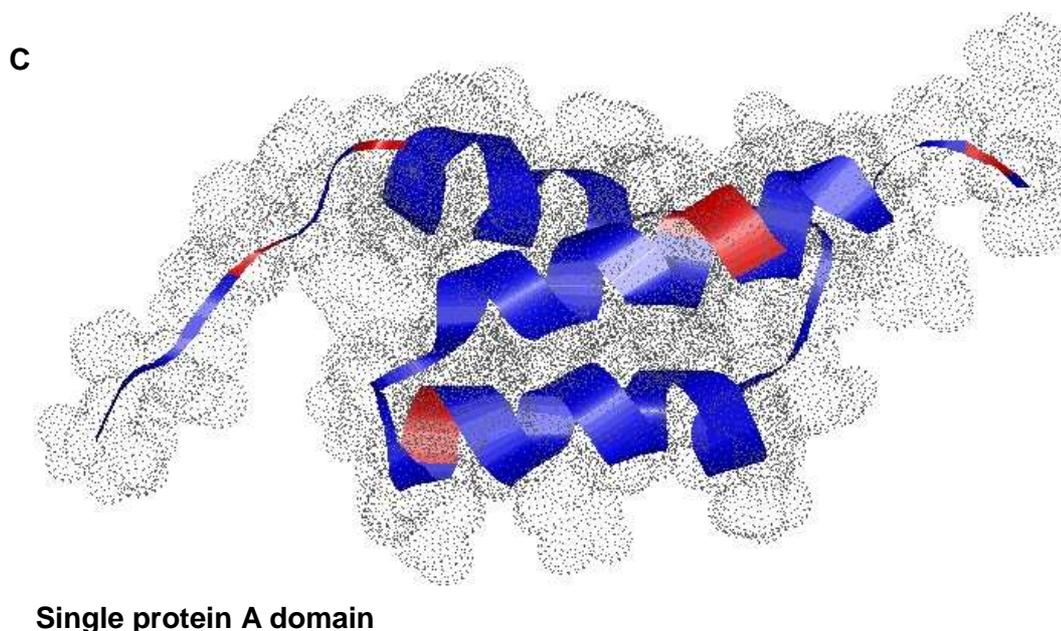
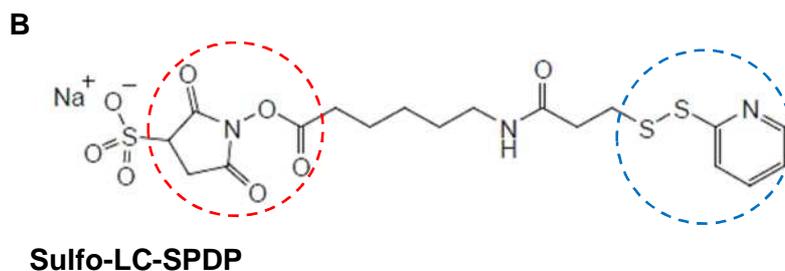
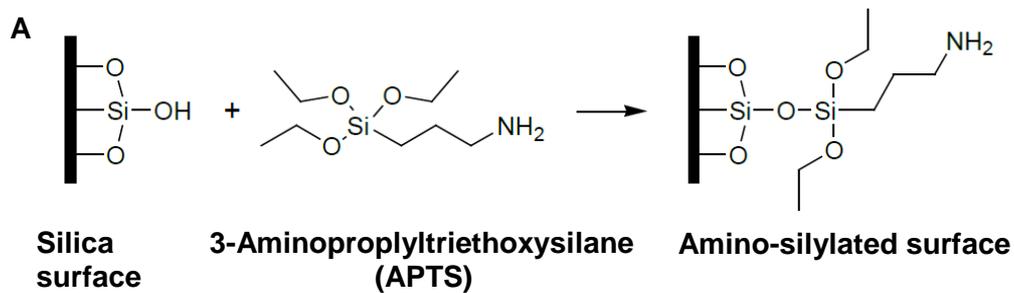


Figure 7.1 Components used to immobilise protein A on the silicon wafer. The surface was amino-silylated surface by reaction of APTS with silicon oxide on the surface of the silicon wafer (a). The SPDP cross-linker (b) was reacted with amine groups on the silylated silicon wafer; the amine-reactive NHS ester group is circled in red. 2-pyridyldithiol groups, circled in green, were then reduced, resulting in sulfhydryl-activated cross-linkers attached to the surface. SPDP was reacted with protein A via its amine side chains; (c) shows an NMR structure of the b domain of protein A; positions of lysine residues are highlighted in red. NMR structure RCSB Protein Data Bank ID: 1BDD.

The nuclear magnetic resonance (NMR) structure of the B domain of protein A shown in *Figure 7.1* was obtained by Gouda et al. (1992). The visual representation of the structure was edited using RasMol software (Bernstein 2000, Sayle and Milner-White 1995).

7.3.2 Cross-linking staphylococcal protein A to silica

Staphylococcal protein A does not contain cysteine residues. Therefore a three-stage process was used to couple the protein A to silica. An additional step was also required to modify the silica surface before coupling. The process was as follows:

- A. SPDP was reacted with protein A via its primary amines and the resulting protein-linker conjugates were purified from reaction by-products to produce *modified protein A*.
- B. The silica surface was modified with the silylation reagent 3-aminopropyltriethoxysilane (APTS), coating it with primary amine groups (*Fig. 7.1a*).
- C. SPDP was reacted with the amine-coated surface. Surface-conjugated cross-linkers were then reduced using DTT.
- D. *Modified protein A* was then applied to the surface, allowing the 2-pyridyldithiol groups on *modified protein A* to react with the sulfhydryl groups of the reduced surface cross-linkers.

In the reaction scheme described, SPDP cross-linkers are attached to protein A via amine groups, which could include the N-terminal amine as well as those of lysine residues. The pH of the reaction was too low (between pH 7 and 8) for the NHS ester to react with arginine residues (McGee et al. 2012) so these will be assumed non-reacting, as will histidine residues. *Figure 7.1c* shows the structure of the B domain of staphylococcal protein A (Gouda et al. 1992) with lysine residues highlighted in red. In an ideal situation, *reaction A* (described above) would yield *modified protein A* with only one linker per protein molecule. This would reduce the likelihood of excess cross-linkers interfering with the protein A-IgG association, the structure of which we wish to characterise. Another consideration is that excess of SPDP can cause proteins to associate via the linker hydrophobic chains (particularly for the long chain version, LC-SPDP), resulting in protein precipitation (Hermanson 2013b). Attempts to control the molar

ratio of SPDP to protein can be made by adjusting the reagent quantities (Carlsson et al. 1978, Hermanson 2013c), but the location of SPDP attachment cannot be controlled.

The molar ratio of SPDP to protein can be determined by diluting modified protein solution and adding a reducing agent, releasing the pyridine 2-thione group from the cross-linkers. The quantity of pyridine 2-thione released is determined by measuring absorption at 343nm (Carlsson et al. 1978). Using this method the molar ratio of SPDP to protein A was found to be approximately 2:1 (data not shown). Each protein A molecule contains 5 IgG-binding domains (Capp et al. 2014, Jansson et al. 1998, Lund et al. 2011, Moks et al. 1986, Peyser 2010), so two linkers per protein is not expected to have a drastic effect on IgG binding. For this work, patented recombinant protein A from Repligen (MA, USA) was used. This particular protein A is encoded by a nucleotide sequence derived from *Staphylococcus aureus* strain 8325-4, and is expressed in *E. coli* (Peyser 2010). Henceforth, this patented form of protein A will be referred to as “rSPA”. Repligen’s rSPA contains 5 IgG-binding domains – E, D, A, B and C – and a truncated form of the X domain, which, in its native form, is involved in cell wall attachment. Schematic depictions of the domain structures of staphylococcal protein A and recombinant protein A can be seen in *Figure 3.4, section 3.3.2 Protein A*. Repligen’s rSPA has the same domain structure as staphylococcal protein A, but the X domain is smaller. Unlike the recombinant protein A used in GE Healthcare’s MabSelect Xtra, there is no C terminal cysteine. The amino acid sequence of Repligen’s rSPA is shown in *Figure 7.2*. According to Peyser (2010), the truncated X domain may provide a useful site for attachment of rSPA to a solid support due to the abundance of lysine residues in this region of the protein (Peyser 2010). Nonetheless, the location of the linkers is not known, so if they reside near the Fc-binding sites there could be more impact than anticipated. The residues on the C domain of protein A involved in Fc binding can be seen in *Figure 3.5, section 3.3.2 Protein A*; most of them reside on helix I and half as many are on helix II. There is one lysine which is involved in the protein A-Fc interaction according to Deis et al. (2015). This resides on helix II and, as can be seen in *Figure 7.2*, is present in all five IgG-binding domains. Thus, interaction of the cross-linker with Lys35 on helix II could potentially affect the protein A-IgG binding affinity. However, Lys35 is

not part of the specific molecular recognition pocket crucial to high affinity Fc binding, though it does reside close-by (Deis et al. 2015). Further, as noted above, far more lysine residues reside on the truncated X domain of rSPA, and the ratio of SPDP to protein A was found to be low at 2:1.

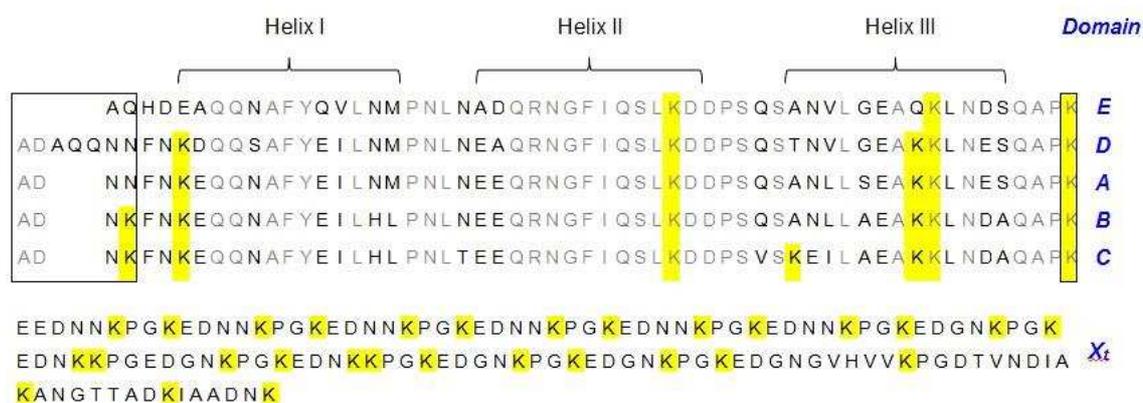


Figure 7.2 Amino acid sequence of recombinant staphylococcal protein A (rSPA) from Repligen. E, D, A, B and C domains are aligned to show sequence homogeneity. Residues that are identical across domains are shown in grey, those that vary across domains are in black and lysine residues are highlighted in yellow. Unstructured regions that link each three-helix bundle are within black boxes. The truncated X domain is denoted X_t and is shown as a continuous string of residues in black.

Figure layout adapted from Capp et al. (2014) using the sequence for rSPA from Peyser (2010) .

Various methods, as well as the pyridine-2-thione assay, were used to confirm that the linker was attached to protein A. After reaction with SPDP, the modified protein was separated from reaction by-products by gel filtration. The gel filtration fractions were analysed spectrophotometrically by absorption between 200nm and 400nm. Wave-scans of the fractions were compared with that of pure SPDP and unmodified rSPA. SPDP showed a major peak at 283 nm while the major peak for rSPA was at 276nm. Attachment of the linker resulted in a shift of the rSPA peak maximum to 278-279nm. Gel filtration fractions that followed showed a clear peak at 268 nm, indicative of the liberated NHS group (see *Fig. 7.1*). The suspected protein-containing fractions were analysed by SDS-PAGE and the resulting bands at 48kDa confirmed the presence of rSPA (data not shown).

7.3.3 Blocking agents

When protein A was cross-linked silica, the level of surface coverage was not known. Therefore, it was deemed necessary to apply a blocking agent after rSPA cross-linking in order to limit non-specific binding of IgG to the potentially reactive silica surface. Two different blocking agents were investigated: bovine serum albumin (BSA) and polyethylene glycol (PEG) 6000. Regions lacking the immobilised protein were predicted to contain a mixture of hydrophilic groups, due to un-reacted silanols; polar amine groups, due to un-reacted amino-silanes; and hydrophobic alkyl groups, part of the APTS molecule (*Fig. 7.1a*). It was not known how many of the three methyl groups on APTS reacted with silanols during silylation of silica, meaning free alkyl/methyl groups may have been present on the modified surface. Unmodified PEG has been shown to act as an effective blocking molecule, possibly through hydrophobic interaction, and has properties that tend to repel or disperse proteins (Le Brun et al. 2011, Liu et al. 2013, Nakanishi et al. 2008, Samuel and Harris 1997). BSA is a well-characterised protein block for hydrophilic and hydrophobic surfaces (Białopiotrowicz and Jańczuk 2001, Jeyachandran et al. 2010, Nakanishi et al. 2008). Jeyachandran et al. (2010) found that incubation of a hydrophilic surface with 10 mg/mL BSA solution for 30 minutes produced a BSA layer with 63% surface coverage. This was found to effectively block 68% and 100% non-specific binding of IgG and staphylococcal protein A, respectively. Similar blocking of a

hydrophobic surface with BSA had 95% blocking efficiency against IgG and 100% against protein A (Jeyachandran et al. 2010). Hence, a similar high concentration / short time-period blocking strategy was used in the work described here.

7.4 Monitoring surface modifications with ATR-FTIR

ATR-FTIR was used to monitor the covalent attachment of recombinant staphylococcal protein A (rSPA) to an amino-silylated silicon crystal, blocking of un-reacted sites with BSA, and IgG4 binding and elution. Non-specific-binding of rSPA and IgG to the amino-silylated surface was also investigated. An advantage of using ATR-FTIR was that the sample configuration was almost identical to that used in reflectivity experiments – a polished, flat silicon crystal (/wafer) assembled inside a low-volume flow-through cell with a small chamber height. FTIR allowed close monitoring of the chemical environment and provided information on the secondary structure of proteins that were immobilised or adsorbed. Neutron reflectivity gave information on three-dimensional form and apparent density at the solid-liquid interface. Thus, the two techniques were complementary.

FTIR is a vibrational spectroscopic method that measures the wavelength and intensity of absorption of infrared radiation by a sample. FTIR spectra are displayed as relative absorbance against the reciprocal of wavelength, the *wavenumber*, which has units of cm^{-1} . Vibrational frequency, proportional to wavenumber, is related to the strength and polarity of chemical bonds. Specific structural bonded units, i.e. chemical groups like CH_2 or COOH , give rise to characteristic absorption bands. Bands are influenced by intramolecular structure because intramolecular bonds, namely hydrogen bonds, affect the electron density of a given bond; similarly the chemical or solution environment can influence absorption bands (Arrondo et al. 1993, Barth 2007, Kong and Yu 2007, Seshadri et al. 1999). For proteins, the most useful and prominent bands are the *amide I* and *amide II* bands, at $1700\text{-}1600\text{ cm}^{-1}$ and $1600\text{-}1500\text{ cm}^{-1}$, respectively. These bands arise due to stretching and bending vibrations of the protein backbone amide groups. The amide I band is predominantly due to carbonyl stretching and the amide I due to mostly N-H bending but also C-N stretching (Barth 2007, Seshadri et al. 1999). Amide

bands are broad and consist of a number of overlapping peaks. Specific sub-peaks are characteristic of different types of secondary structure, such as β -sheets, β -turns, α -helices and random coils. Amide I is the more prominent of the two bands, and is also the most useful for identifying secondary structure features (Barth 2007).

ATR-FTIR is commonly used to analyse thin protein films on surfaces (Barth 2007, Goormaghtigh et al. 1990, Seshadri et al. 1999); in the case of this work, the techniques was used to analyse protein closely associated with a surface, though not necessarily a homogenous “film”. The sample configuration in ATR-FTIR is shown schematically in *Figure 7.3*. The IR beam enters the ATR crystal at an angle which results in its total reflection at the solid-liquid interface; a small number of reflections occur before the beam exits to the detector. At the point of reflection, the beam penetrates the sample layer and absorbance due to sample vibrational frequencies is eventually transmitted to the detector (Barth 2007).

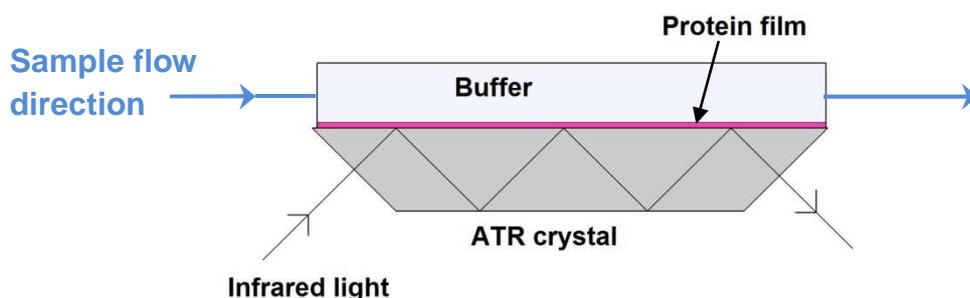


Figure 7.3 Simplified depiction of an ATR-FTIR sample and infrared beam configuration, with the sample cell in flow-through mode.

All FTIR analyses were carried out in D₂O-based buffers. The reasons for this were twofold: (i) H₂O absorbs strongly at 1645 cm⁻¹ which is in the amide I band region (Arrondo et al. 1993); (ii) it was complimentary to reflectivity experiments in the D₂O solution contrast. In the reflectivity experiments described in this work, crucial information was to be gained from the D₂O contrast (see *section 6.6.2 Experiment design*).

Proteins are known to interact with surfaces like silica, and the results described in *Chapter 6* show that IgG4 can adsorb directly to silica under mild to moderately acidic conditions (the highest pH was pH 6.2). However, for the experiments described in this chapter it was important to ensure that such “non-specific” binding was absent or minimal as high levels of non-specific binding, particularly of IgG, could confound results. *Figure 7.4* shows one sequence of FTIR experiments that was carried out in order to assess non-specific binding of both protein A and IgG to the model surface. The method allowed concomitant monitoring of the protein A cross-linking process and subsequent IgG adsorption. Amino-silylation of the silicon crystal was carried out “on the bench”, as the solvents required were not suitable for use in the FTIR instrument. All stages of surface modification thereafter were carried out in flow-through mode with the silicon crystal in-place in the FTIR instrument. Initial background readings were taken with the amino-silylated crystal in place.

The FTIR spectra displayed in *Figure 7.5.a-e* and *Figure 7.6* correspond to the relevant stages in *Fig. 7.4*. For example *Fig. 7.5.a* corresponds to illustration (i) in *Fig. 7.4*, which represents interaction of *un-modified* rSPA with the amino-silylated surface. The red spectrum in *Fig. 7.5.a* suggests a relatively small amount of non-specific interaction with the surface; a small peak is present in the amide I / II region after the rSPA solution has been washed out of the cell. *Figure 7.5.b* shows that the presence of the cross-linker attached to rSPA does not impart additional attraction to the surface, since the size of the amide peak barely increases after incubation with modified rSPA. Next, the cross-linking stages were implemented. The same FTIR crystal was kept in place and a new background taken. *Figure 7.5.c* shows spectra before, during and after reaction of SPDP with the surface. No major changes were observed during cross-linker attachment. It is possible that the spike just below 1500 cm^{-1} is due to the pyridyldithiol group. Spectra acquired during cross-linking of modified rSPA to the surface are shown in *Figure 7.5.d*. Amide I and II bands are visible during and after cross-linking. The amide peak(s) for cross-linked rSPA are more pronounced than that for non-specifically bound rSPA. It should be noted that non-specifically bound rSPA was not removed from the surface before rSPA cross-linking (see *Fig. 7.4*).

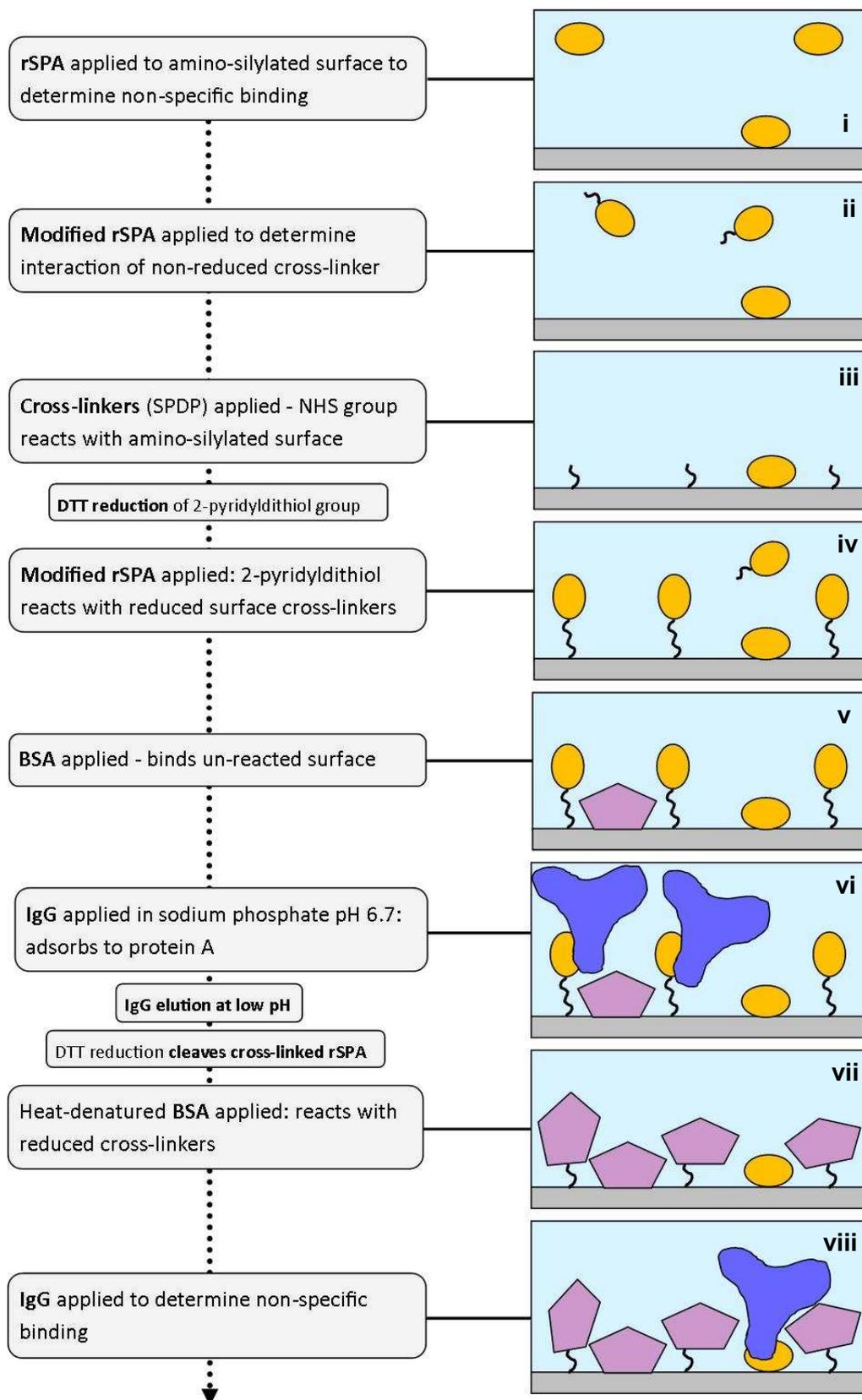


Figure 7.4 Flow chart and cartoon diagrams to illustrate a procedure used to assess rSPA non-specific binding to silica, rSPA cross-linking, IgG adsorption and IgG non-specific binding. The surface was analysed by ATR-FTIR.

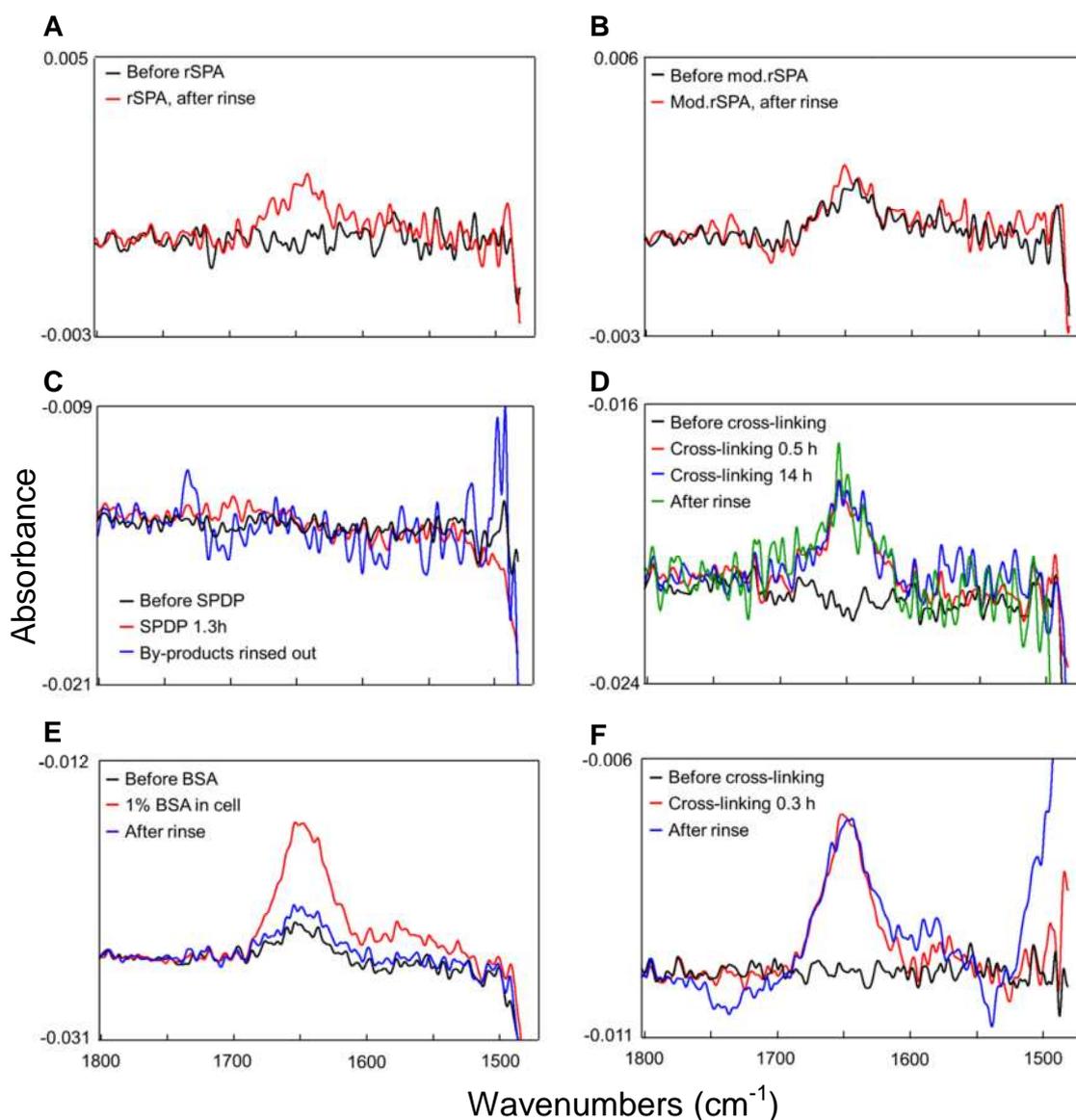


Figure 7.5 FTIR spectra for sequential stages of silica surface modification. Within each plot, the event sequence is analogous to the following colour sequence: black (initial) – red – blue – green (end). Non-specific binding of rSPA (a) and of linker-modified rSPA (b) to the amino-silylated silica surface was first assessed. Keeping the same silicon crystal in place, the instrument was re-blanked and SPDP cross-linker flowed into the sample cell (c). Modified rSPA was then introduced, allowing thiol bond formation, resulting in immobilised rSPA (d). Finally, BSA was used to block remaining unreacted or “sticky” sites (e). The cross-linking steps (c-d) were repeated on a fresh crystal, giving an improved signal to noise ratio for rSPA (f).

After rSPA cross-linking 1% BSA solution was incubated inside the sample cell. In *Figure 7.5.e* a large amide band can be seen in the red spectrum, this is due to the high concentration of protein in the solution (10 mg/mL). When the solution was rinsed away, a small quantity of protein remained on the surface, as indicated by the blue spectrum. It is possible that BSA adsorption was lessened because the surface had already been “blocked” to some extent by non-specifically bound rSPA (*Fig. 7.5.a*). However, when cross-linking methods were repeated on a fresh silicon crystal, without initial application of rSPA or mod.rSPA, a similarly modest quantity of BSA adsorption was observed (data not shown). Spectra acquired during and after rSPA cross-linking on the fresh crystal are shown in *Figure 7.5.f*. Here, the signal to noise ratio is improved, possibly because the spectrum was blanked fewer times before these measurements (compared to those in *Fig. 7.5.d*). Cross-linking stages were repeated at least twice more on different occasions using fresh silicon crystals, yielding similar results.

Continuing the steps illustrated in *Fig 7.4*, IgG was introduced into the cell (illustration *vi*). The buffer was changed from a “coupling buffer” (used for rSPA cross-linking) to the affinity adsorption buffer 0.02M sodium phosphate pH 6.7, and the instrument was blanked before IgG addition. Spectra for IgG adsorption are shown in *Figure 7.6.a*. Here, there is a very well-defined amide peak with amide I and amide II bands visible. The peak declines modestly when the IgG solution is rinsed out of the cell. What appears to be the incline of an additional peak can be seen to the right of the plot a little below 1500 cm^{-1} . This is likely caused by the band known as *amide II'*. Amide II' is in the region 1490-1460 cm^{-1} and arises when the amide II band loses the N-H bending contribution and is converted to a mostly C-N stretching vibration (Barth 2007) – this happens as a result of backbone hydrogen-deuterium exchange (note that all FTIR experiments were carried out in D_2O -based buffers).

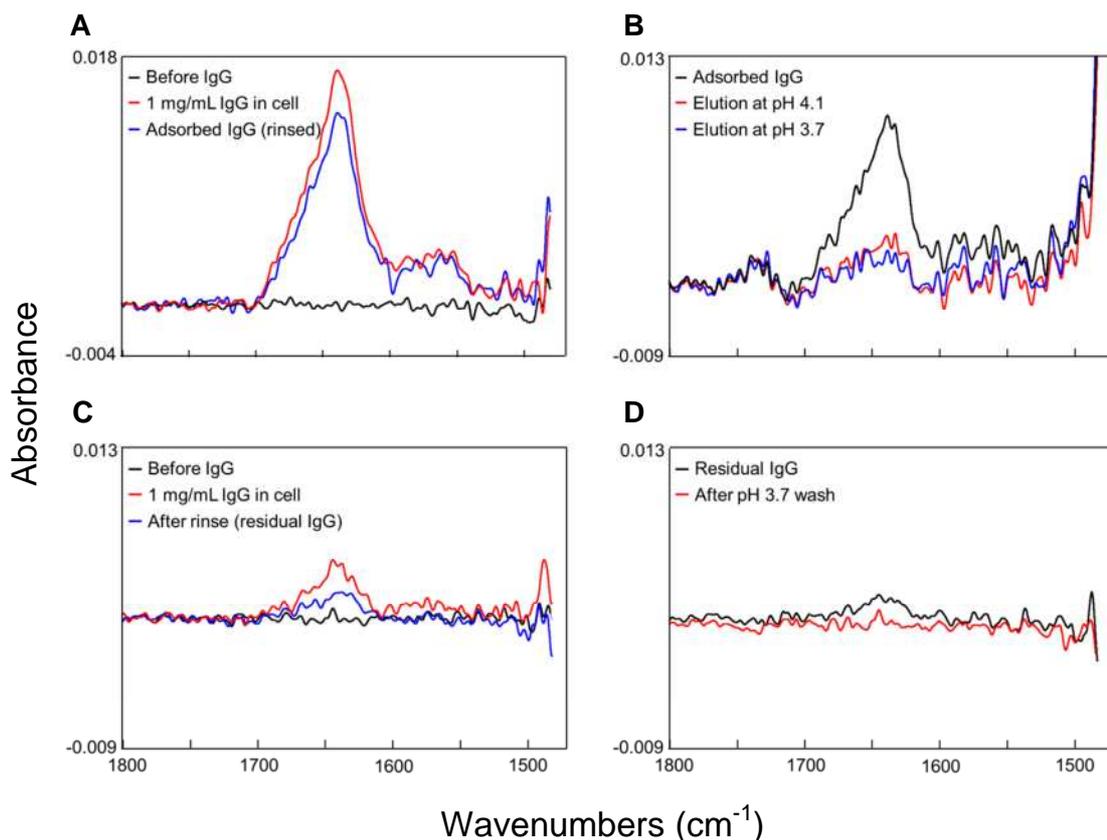


Figure 7.6 FTIR spectra showing stages of IgG adsorption to protein A, elution from protein A, and IgG non-specific binding and subsequent elution. Within each plot, the event sequence is analogous to the following colour sequence: black (initial) – red – blue (end). IgG was flowed into the sample cell after rSPA cross-linking (a). Elution of adsorbed IgG was effected using sodium acetate at pH 4.1 followed by glycine-HCl at pH 3.7 (b). Cross-linked rSPA was cleaved by DTT reduction and reactive cross-linkers were blocked with denatured BSA (see *Figure 7.4*). IgG was then introduced to check for non-specific binding (c). Residual IgG was eluted at pH 3.7 (d).

IgG elution was effected first with sodium acetate buffer at pH 4.1, then with glycine-HCl, pH 3.7. The spectra in *Figure 7.6.b* show that the majority of IgG4 was eluted at pH 4.1 and possibly a fraction more at pH 3.7. The baseline is not straight enough to say whether the amount of IgG remaining was significant. pH 4.1 seems a fairly high pH for elution of the IgG, typically the higher end of the range in bioprocesses is around pH 3.8. It is possible that the protein A-IgG interaction was weakened in D₂O-based buffer, compared to H₂O. In H₂O buffers the K_D of the protein A-Fc interaction is 10-30 nM (Deis et al. 2015). A D₂O environment is thought to influence protein structure because it affects the strength and length of hydrogen bonds (Kong and Yu 2007). Indeed, shifts in the positions of amide I sub peaks in D₂O systems have been documented in FTIR studies, suggesting changes in intra-molecular hydrogen bonds, though this is not a well-understood phenomenon (Arrondo et al. 1993, Kong and Yu 2007). Similarly, D₂O may affect protein-protein interactions. Delano et al. (2000) identified four hydrogen-bonding sites in the *consensus binding site (CBS)* of IgG's Fc, two of which are involved in the interaction with protein A. Shukla et al. stated that elution is brought about when hydrophobically stacked histidine residues in the *CBS* –protein A complex become positively charged at low pH and repel one another (DeLano 2000, Shukla et al. 2007a). A D₂O-based buffer of a given pH* (measured using a probe calibrated with H₂O-based buffers and adding a constant of 0.4 pH units) does not have identical protonating behaviour to that of the same H₂O-based buffer of equivalent pH (Krežel and Bal 2004). Thus, both the binding affinity and the buffer properties may be subject to shifts, resulting in a significant overall shift in elution conditions. However, given that the interaction still takes place, it is assumed that the overall binding mechanism, in terms of protein conformation and orientation, is not significantly affected.

The final stages depicted in *Figure 7.4* were implemented to assess non-specific binding of IgG4 to the modified surface (illustrations *vii* and *viii*). Cross-linked protein A was cleaved from the surface by reduction with DTT. It was then necessary to block the resultant reactive cross-linkers. A blocked surface, rather than a bare surface, was considered more relevant to the surface that would be presented for IgG adsorption to protein A. BSA was used as it contains a

number of di-sulphide bonds, as well as one free cysteine, (Barone et al. 1995) which would react with the reduced surface linkers, resulting in cross-linked BSA. The BSA solution was briefly heat-denatured prior to application in the sample cell in order to ensure exposure of cysteine residues. Indeed, BSA crosslinking resulted in an amide peak with an area directly proportional to that of cross-linked rSPA, considering its slightly larger size (data not shown). The system was then returned to IgG adsorption buffer and blanked before 1mg/mL IgG was introduced into the cell. *Figure 7.6.c.* shows the resultant spectra. There are two major observations to be made. First, the blue spectrum shows a very small peak in the amide I region for apparently non-specifically bound IgG. Secondly, if we compare *Fig. 7.6.c* and *a*, we can see that when IgG solution is inside the cell (red spectra) the amide band is much larger in *a* than that in *c*. This suggests that the red spectrum in *Fig. 7.6.c* solely represents IgG adsorbed to rSPA, some of which is weakly adsorbed and easily rinses away, perhaps due to steric hindrance at the flat surface. Conversely, in *Fig. 7.6.c* a far less IgG associates with the surface, even before the IgG solution is rinsed through.

Finally, the surface was rinsed with pH 3.7 buffer. This appeared to elute the residual IgG, as evidenced by the spectra in *Fig. 7.6.d*. The proposed mechanism for this is illustrated in *Fig. 7.4*, illustration *viii* – some non-specifically adsorbed rSPA from the first step (i) may have remained, thus allowing elution of the bound IgG in step *viii* at low pH.

FTIR spectra for cross-linked rSPA and adsorbed IgG4 were integrated within the amide I and amide II regions. These peak areas are compared in *Figure 7.7.a*. In this example the IgG peak area is 2.5 times larger than the rSPA area. This is suggestive of a 1:1 ratio of IgG: rSPA, as the IgG is approximately 3 times the molecular weight of rSPA. This was somewhat lower than expected, as rSPA has 5 IgG-binding domains and is known to bind at least two IgG molecules. The apparently low binding could have been caused by a number of things. It is possible that the presence of excess cross-linkers attached to protein A affected its interaction with IgG; even though the average number of linkers per rSPA molecule was 2, some may have contained more, or their position may have had more of an impact. Considering a different aspect, there

could have been increased steric hindrance because of the very flatness of the silicon crystal, or protein A ligands may have been very closely packed. Another possibility is that the BSA blocking molecules placed steric limitation on IgG binding. As was done for rSPA cross-linking, IgG adsorption and detection by FTIR was repeated at least three times, yielding quantitatively similar results between repeats.

The spectra in *Figure 7.7.b* and *7.7.c* are presented to highlight the differences between amide band shapes for IgG4 and rSPA. The differing positions of major sub-peaks are the result of different secondary structure contents of the two proteins. The spectra indicate that each protein's characteristic secondary structure content was maintained on immobilisation/adsorption. Peaks were identified automatically by the Omnic software (Thermo Scientific). Only amide II peaks detected are displayed (not amide I). The peak designations are valid for proteins in D₂O-based solutions (Arrondo et al. 1993, Barth 2007, Kong and Yu 2007). In *Figure 7.7.b* the peak at 1657 cm⁻¹ indicates alpha-helical structure. Peaks at 1646 cm⁻¹ and 1642 cm⁻¹ indicate random coil, though some sources have ascribed peaks at wavenumbers as low as 1645 cm⁻¹ to alpha helices (Goormaghtigh et al. 1990). In *Figure 7.7.c* the peak at 1632 cm⁻¹ suggests beta sheet (Barth 2007, Goormaghtigh et al. 1990, Kong and Yu 2007). These observations are mostly consistent with what would be expected for staphylococcal protein A. Looking at *Fig. 7.1.c* and *Fig. 7.2* we can see that as well as the alpha-helical regions, there are significant lengths of random coil between the helical regions, between the different domains, and presumably within the domain X_i. It is possible that some β sheet structure exists within the X_i domain. The identity of the steep incline to the right of plot, starting at about 1540 cm⁻¹, is not clear. It could be related to the amide II' peak, a reasonable possibility since the structure of rSPA suggests a high level of solvent exposure (Capp et al. 2014), which would result in fast H-D exchange and cause amide II' to arise. However, the peak seems disproportionately large. Though it may screen an aptly proportioned amide II' peak, we suggest that it is in fact an artefact that arose when the sample cell was adjusted to encourage the passage of trapped air.

In *Figure 7.7.c* a spectrum for adsorbed IgG is shown with key amide I peak designations. The most significant peak is at 1635 cm^{-1} signifying β -sheet structure. Peaks at 1688 cm^{-1} and 1613 cm^{-1} also suggest β sheet. The peak at 1661 cm^{-1} has been designated β -turn in accordance with values in the literature (Arrondo et al. 1993, Barth 2007, Goormaghtigh et al. 1990, Kong and Yu 2007). These peak assignments are consistent with the structure of the immunoglobulin fold, which comprises 7-9 anti-parallel β strands folded around a hydrophobic core (Barth 2007, Bork et al. 1994).

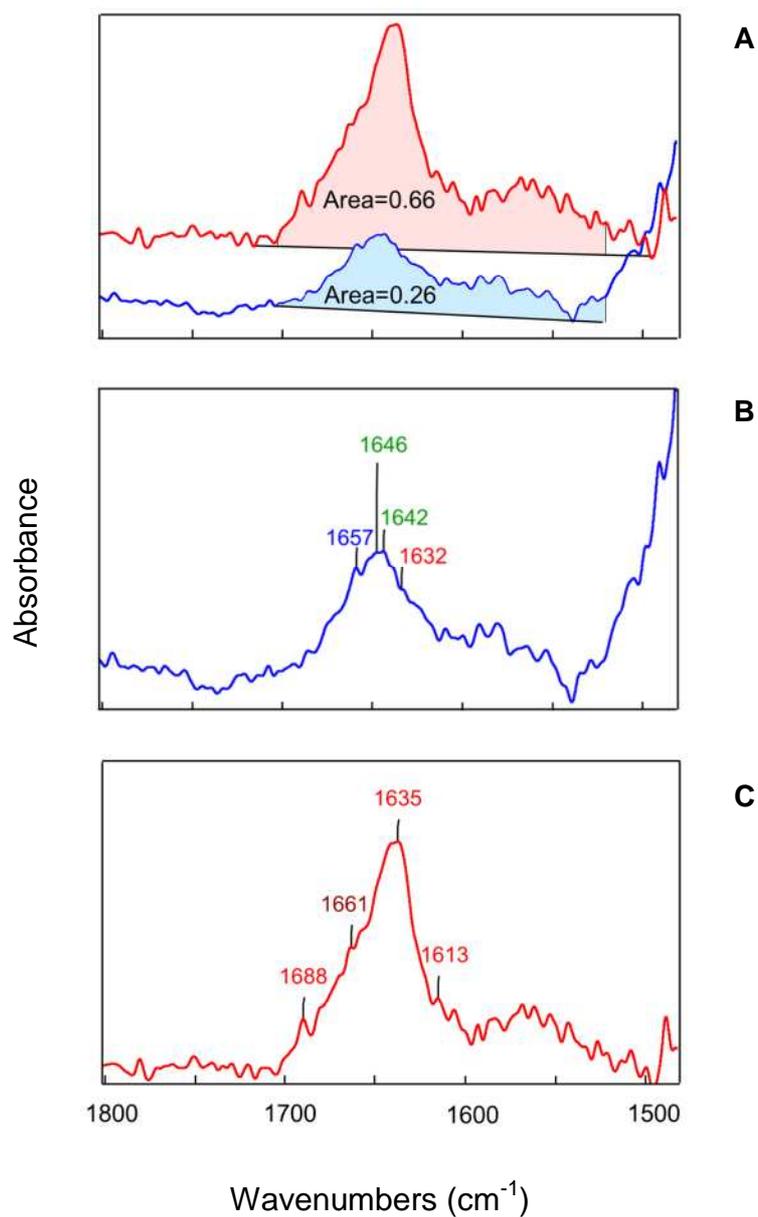


Figure 7.7 Annotated FTIR spectra for immobilised rSPA and adsorbed IgG showing amide peak areas (a) and sub-peaks identified for rSPA (b) and IgG4 (c). Blue spectra are rSPA and red spectra are IgG4. Peak wavenumbers are highlighted in different colours to indicate the type of secondary structure they correspond to: blue indicates α -helix; green, random coil; red, β -sheet; dark red, β -turn.

In summary, the results of the FTIR experiments show that rSPA was covalently attached to a silica surface using the cleavable cross-linker SPDP. Data indicated that there was not a large excess of cross-linkers attached to rSPA, meaning that a representative interaction between rSPA and IgG could be achieved. Immobilised rSPA retained its characteristic secondary structure during the immobilisation process, mostly alpha helix and random coil. An apparently modest quantity of BSA adsorbed to the surface, fulfilling the role of inhibiting non-specific binding of IgG to potentially reactive sites where protein A was not present. IgG adsorbed to protein A under typical chromatography loading conditions, with data suggesting a 1:1 ratio of protein A to IgG. The secondary structure of IgG was retained on adsorption, and the differences in amide band shape for protein A and IgG were clearly defined. IgG was eluted at pH 4.1, a little higher than the pH typically required for elution. This could be a result of changes in the strength of the interaction and of the acidic properties of the elution buffer due to the use of D₂O-based buffers. This was not expected to change the physical mechanism of the IgG-protein A interaction. Lowering the pH to 3.7 did not cause significant further elution. The rSPA ligand was cleaved and replaced with BSA. IgG did not interact to any notable extent with the fully BSA-blocked surface. SPDP linkers attached to protein A were not found to evoke interaction with the silica surface in the absence of free sulfhydryl groups. Some non-specific binding of rSPA occurred, but cross-linking occurred nonetheless and was thought to be preferential.

These findings demonstrate that generation of the model surface was successfully implemented in a reproducible manner. A strong basis is provided for investigation of the arrangement of the various molecules at the solid-liquid interface.

7.5 Neutron reflectivity: characterisation of IgG adsorbed to protein A at the solid-liquid interface

7.5.1 Reflectivity experiments and data modelling

For an introduction to neutrons and reflectivity methods for characterisation of biological systems see *sections 6.4-6.6*. Neutron reflectivity was used to characterise the structure of IgG adsorbed to immobilised protein A in three separate solid-liquid flow cells (SLFCs) – for a schematic illustration of an SLFC see *section 6.5*. In all cases, the silicon wafers were treated with APTS, producing an amino-silylated surface, before assembly into SLFCs and fixing to the reflectometer sample cell stage. All subsequent surface modifications were carried out in flow-through mode. The three SLFCs were utilised as follows:

Cell A: rSPA cross-linked to surface; surface incubated with 1% BSA blocking solution; IgG adsorption at pH 6.7; IgG eluted.

Cell B: rSPA cross-linked to surface; surface incubated with 1% PEG₆₀₀₀ blocking solution; IgG adsorption at pH 6.7; IgG eluted.

Cell C: rSPA cross-linked to surface; surface incubated with 1% PEG₆₀₀₀ blocking solution; IgG adsorption at pH 6.7; IgG eluted.

For each cell, neutron measurements were made at various stages of surface modification across a range of solution phase contrasts, that is, in D₂O-based, H₂O-based or *silicon-matched water*-based buffer. *Silicon-matched water* (SMW) contains 38% D₂O / 62% H₂O and has the same scattering length density (SLD) as silicon. Thus, neutron reflectivity data obtained with a SMW solution phase arises purely as a result of interfacial structure/material as there is no scattering contribution from the solution (Clifton et al. 2013). For the first SLFC, *cell A*, measurements were made after each stage of cross-linking. For SLFCs *B* and *C*, measurements were made at selected stages of interest. When protein was present, measurements were made at all three solution phase contrasts. However, in the absence of IgG, SMW data generally did not contain

sufficient information to for robust fitting. Thus, much of the SMW data has been omitted. Where deemed appropriate, data was acquired for the D₂O contrast only.

Table 7.1 summarises model inputs and outputs in terms of which solution phase contrasts were used for each sample model across the different SLFCs, and the resulting chi-squared values.

Table 7.1 Chi-squared values for fitted models for different sample configurations across the three SLFCs (A, B and C). Models were fitted using data sets from one, two or three solution phase contrasts, indicated by shading.

Sample Cell Solution Phase Contrast	A			B			C		
	D2O	H2O	SMW	D2O	H2O	SMW	D2O	H2O	SMW
Si (amino-silylated)	1.81						2.78		
Si + SPDP	2.03			1.22					
Si, SPDP, rSPA	2.54			4.1					
Si, SPDP, rSPA, BSA	1.81								
Si, SPDP, rSPA, BSA, IgG4	2.95								
Si, SPDP, rSPA, PEG				2.55			2.04		
Si, SPDP, rSPA, PEG, IgG4				3.93			2.53		

In terms of chi-squared values, the best model fits were generated from *cell A* and *cell C* data sets. This will be discussed further below. Bootstrap error analysis was used to assess fitting error for each parameter in a given model (*see section 6.6.3.3* for details).

For each data set, a constrained model was built using the methods described in *section 6.6.3*. In all cases, the simplest possible model, i.e. least number of layers that produced an adequate fit, was used. As was done for data in *chapter 6*, limits for constraining SLD parameters for the various layers were determined based on SLD estimates of the pure materials in D₂O and H₂O, either done manually or using online SLD calculators (*see section 6.6.2 Experiment design* for details). For SPDP, suitable limits were determined by manually estimating the SLD of SPDP bound to APTS. It was noted that SPDP bound to APTS has two exchangeable hydrogens, and that the linker length doubles on cross-linking rSPA. Estimates were also made using the NCNR Neutron Activation and Scattering Calculator (at www.ncnr.nist.gov/resources/activation/) for comparison. Estimates in D₂O ranged from $8.4 \times 10^{-7} \text{ \AA}^{-2}$ to $2.2 \times 10^{-6} \text{ \AA}^{-2}$ and in H₂O from 5.1×10^{-7}

to 1.5×10^{-6} . Thus, for D₂O data the minimum SLD (for a pure SPDP layer) was set at 5×10^{-7} , and for H₂O data the maximum was set at 2.5×10^{-6} . For rSPA, SLDs in D₂O and H₂O were estimated using the Biomolecular Scattering Length Density Calculator available at <http://psldc.isis.rl.ac.uk/Psldc/>; values were similar to those for IgG4-κ – see table 6.3, section 6.6.2 *Experiment design*. The same was assumed for BSA. For PEG₆₀₀₀, the SLD was estimated to be approximately $8.0 \times 10^{-7} \text{ \AA}^{-2}$. For stages after rSPA crosslinking, layers were expected to overlap in composition.

It should be noted that all protein samples were prepared, stored and applied to the surface in D₂O-based buffers. When solution phase contrasts were changed from D₂O to SMW or H₂O, there will have been some hydrogen-deuterium back-exchange, affecting solution-phase SLDs. The effect on solution phase SLDs was accounted for in the modelling process. If deuterium incorporated into protein amide groups did not exchange back to hydrogen when the solution phase contrast was changed, the SLD of the protein layer would have been affected. This was considered when analysing the data.

7.5.2 IgG adsorption to rSPA-modified surface with BSA blocking

Reflectivity data for the surface modification and blocking stages in *cell A* is shown in *Figure 7.8*. Inspecting plots *a* and *b*, it is difficult to visually identify changes across the sequential sample configurations. However, there are differences between the data-sets, and it should be noted that neutron methods are known to be “signal limited” (Clifton et al. 2013). As material was added to the surface, reflectivity in the Q region approximating $0.02\text{-}0.1 \text{ \AA}^{-1}$ shifted to slightly lower values. The biggest shift occurred on rSPA crosslinking (blue curve).

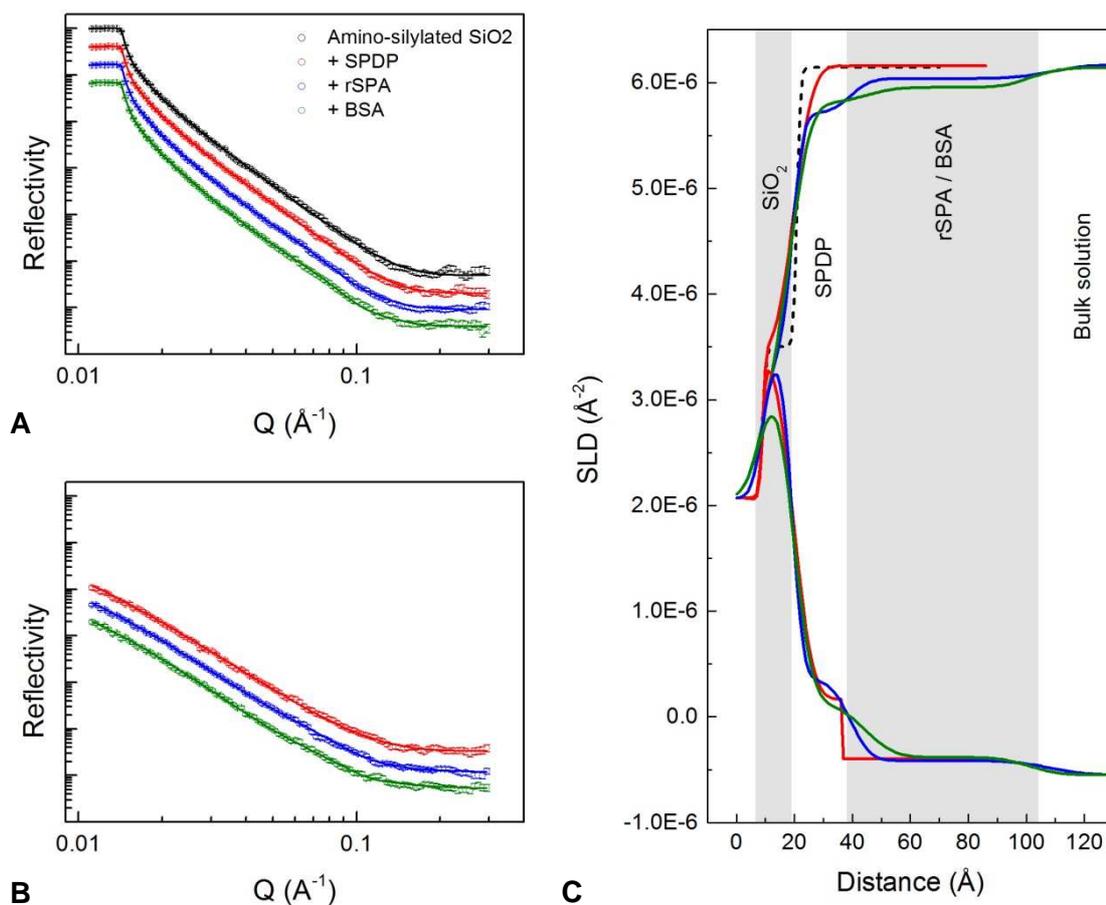


Figure 7.8 Reflectivity data for attachment of rSPA to an amino-silylated silica surface, and addition of the blocking agent BSA. Left-hand plots *a* and *b*: reflectivity (intensity) against Q (\AA^{-1}) in D_2O -based (*a*) and H_2O -based (*b*) buffer at various stages of surface modification. Data sets are offset down the y-axis for clarity; in D_2O the critical edge always occurs when reflectivity is equal to unity. Different colours represent different stages of surface modification. From top to bottom: black is the amino-silylated wafer; red is with the cross-linker attached, blue is after rSPA cross-linking and green is after blocking with BSA. Solid lines show the fitted models. Plot *c* shows the corresponding profiles of scattering length density, SLD (\AA^{-2}), against distance from the silicon surface. Each profile was generated by simultaneous fitting of D_2O and H_2O reflectivity data sets to a single layer depth profile. Curves that terminate at the upper and lower ends of the SLD scale represent D_2O and H_2O data, respectively. Shading is to aid layer visualisation and suggestions of layer identity are in rotated text.

The SLD against distance profiles arising from the model fits are shown in *Figure 7.8.c*. When only the cross-linker was on the surface, a 15Å thick SPDP layer gave the best fit. The SLD of the layer in D₂O was only fractionally lower than that of the bulk solution, but in H₂O the difference in SLD between the SPDP layer and the bulk solution was well-defined. The model for cross-linked rSPA identified a 20Å thick layer proximal to SiO₂ and 69Å thick layer beyond this. The 69Å layer was attributed to rSPA. The SLD values in D₂O and H₂O both suggest only about a 5% surface volume fraction. This seems rather low coverage. However, coverage can appear low if long thin proteins are orientated with their length fully extended in the surface normal direction (Perevozchikova et al. 2015). Nonetheless, 5% is a fairly low figure.

The conformation of full-length protein A has not been the subject of many studies; more commonly, individual domains are characterised or functional studies are implemented. However, Capp et al. (2014) investigated precisely this subject in some detail. They used small angle X-ray scattering (SAXS) and a polymer physics modelling approach to estimate the “statistical conformation” of the 5 IgG-binding domains of staphylococcal protein A, termed “*SpA-N*”. The IgG-binding domains of protein A are joined in series, making a highly flexible molecule in which the radius of gyration is not proportional to the number of domains. This allows models that are usually used to describe flexible polymers to be applied to protein A (Capp et al. 2014). The model found to best-describe a 5-domain string of protein A domains was the “excluded volume pearl necklace model” (EV-PNM). In this model, each domain is assumed a sphere with a given radius. The spheres are connected by freely-jointed links, but no two spheres are allowed to occupy the same volume (Capp et al. 2014). Capp et al. (2014) found the R_g of *SpA-N* to be 37Å, the R_g of a single domain to be 13Å, the modelled sphere diameter to be 21Å and the quantity *persistence length* to be 36Å. Persistence length is related to R_g and describes the length scale of chain flexibility in polymer models. Since the persistence length value is much smaller than the full length of the protein, rather, about the length of 1.5 individual domains, the protein is predicted to be very flexible (Capp et al. 2014).

The modelled depth of the rSPA layer is slightly less than double the persistence length for SpA-N determined by Capp et al. (2014). One crucial difference for the experiments described in this work is that we have used staphylococcal protein A which contains the X_t domain (Fig. 7.2). The X_t domain contains more than double the number of residues of a single IgG-binding domain. It is also in series with the other domains, so is expected to contribute significantly to the total length of the protein. Another potential contributing factor is the presence of excess SPDP ligands. Our data suggests that on average each rSPA molecule had two linkers attached. The excess linker may reduce the flexibility of protein A. It is also considered that protein A may be less flexible in the immobilised state, compared to in free solution. Thus, 69Å seems a reasonable value for the protein A layer thickness.

The SLD in D₂O for the first layer (originally SPDP only) decreased when protein A was added and the thickness increased marginally. An increase in thickness is expected because surface cross-linkers react with cross-linkers on protein A, doubling the length. We would not expect the increase in length to be proportional to the change in linker length because the linker is highly flexible. A decrease in SLD for the D₂O contrast is indicative of an increase in hydrogenous material at the interface. The change in SLD could be due to the flexibility of protein A, as well as the linker, which could both bend back towards the surface as well as extending outwards. We note that we are dealing with a very weak signal here, so we cannot make firm conclusions about the structures present. On the other hand, the model fitting process found a two-layer model to give the best fit, and the bootstrap error analysis gave fitting errors of around 1% for protein A and SPDP layer SLDs. The error estimates were larger for the thickness parameter, around 10% for the protein A layer and 15% for the SPDP layer, but the distributions of re-fitted values were normal. This further supports the concept of a highly mobile layer or material.

Using *Equation 5* (section 6.7.2.3), a volume fraction of 0.05, an rSPA layer thickness of 69Å and a protein mass density of 1.35 g.cm⁻³ (Fischer et al. 2004), the surface excess for the rSPA layer is estimated to be 0.47 mg.m⁻². This equates to approximately 17100Å per rSPA molecule

for an evenly dispersed layer. Assuming upright rSPA molecules with cross-sectional areas of $26 \times 26 \text{ \AA}$, each rSPA molecule is estimated to occupy an area large enough for 25 rSPA molecules. If the cross-sectional area is assumed larger, due to a tilted or flexing orientation, the excess surface space reduces to around 10 molecules' worth.

The addition of BSA caused small shifts in SLD values, suggesting that a small quantity of BSA was adsorbed in the blocking process. The volume fraction estimate for the protein layer increased by about 1.5% after BSA addition. This is in agreement with the FTIR data. The decrease in SLD in the D₂O contrast after BSA blocking extended across the full length of the protein A layer, suggesting that BSA retained a folded structure and extended out into the solution phase. The thickness of the SPDP layer shifted to a slightly higher value while the protein layer thickness was reduced by about 10 Å. This suggests that BSA was adsorbed at the surface, possibly limiting the movement of molecules closer to the surface.

Introduction of IgG4 into *Cell A* caused significant changes in the profile of reflectivity against Q , as can be seen in *Figure 7.9*. When IgG was adsorbed there was sufficient interfacial structure to fit SLD-distance profiles using all three solution-phase contrasts. Looking at the data for D₂O in *Fig. 7.9*, a broad fringe can be seen covering Q values 0.014-0.04 Å⁻¹; fringes in the low Q region arise due to large structures that extend away from the interface. Another clear fringe is around 0.08 Å⁻¹. Importantly, the shape of this reflectivity profile is very different to that seen in *sections 6.7.2-3*, where a large fringe can be seen at the high end of the Q range.

The best fit for the data was achieved using a 5 layer model (including the SiO₂ layer). The resulting SLD against distance profiles for the three contrasts are shown in *Figure 7.10*. Bootstrap error analysis found that the most robust fitted parameters were the SLDs in D₂O and SMW for the three major protein-containing layers (the three outermost ones); error estimates for these parameters ranged between 1% and 3% of the fitted values. Error estimates for layer thickness were much higher, almost 20% for the inner protein layers and 10% for the outermost layer. However, the distributions of bootstrap values were normal and centred around the fitted values (data not shown, see *section 6.6.3.3* for an example of a bootstrap distribution). Hence it

is assumed that the fitted values for layer thickness are relatively accurate for a potentially dynamic system. This assumption must be made if we are to use that data to try and estimate the arrangement of protein structures at the solid-liquid interface. Neutron and X-ray data on the structure of wild-type IgG4 acquired by Rayner et al. (2014) indicated that the IgG is not overly flexible about the hinge region, since R_g values were consistent across a range of sample conditions. At low concentrations (below 2mg/mL) 1-2nm movements of the Fab regions relative to the Fc region were observed (Rayner et al. 2014).

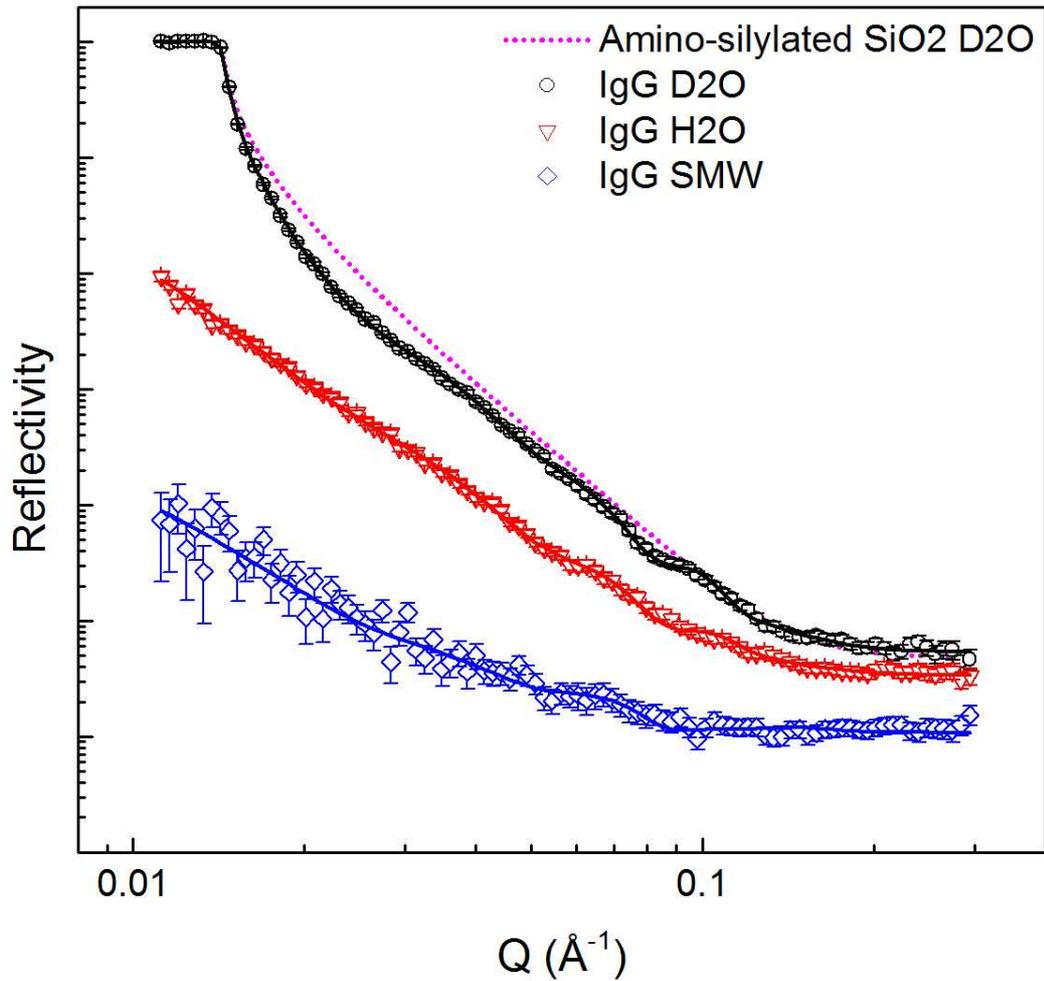


Figure 7.9 Reflectivity (intensity) against Q at three solution phase contrasts for IgG4 adsorbed at pH 6.7 to rSPA-modified silica with BSA blocking. Black data points represent D_2O solution phase; red, H_2O solution phase; blue, *silicon matched water* (SMW) solution phase (38% D_2O). H_2O and SMW data sets are offset down the y-axis to prevent overlap; in D_2O the critical edge always occurs when reflectivity is equal to unity. Solid lines represent fitted models. The pink dotted line represents the fitted model for the amino-silylated surface in D_2O (Fig. 7.8) and is shown for comparison.

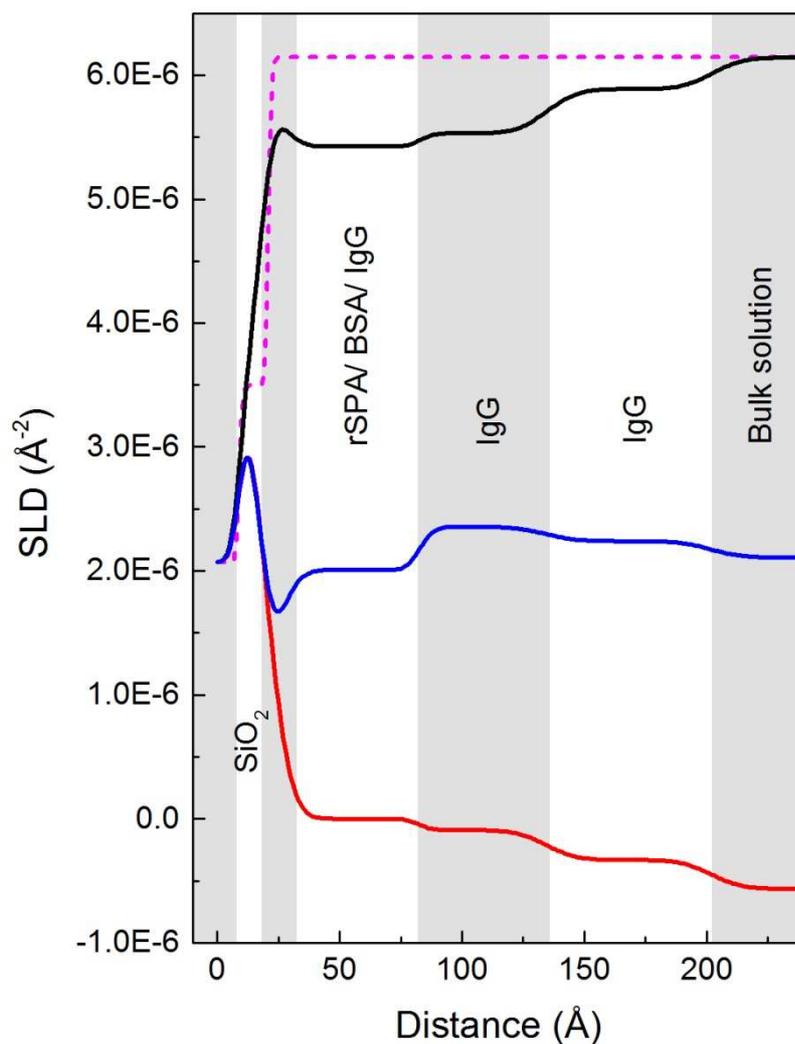


Figure 7.10 Profile of scattering length density against distance for IgG4 adsorbed at pH 6.7 to protein A-modified silica with BSA blocking. Data was fitted simultaneously for three solution phase contrasts: D₂O (black), H₂O (red) and *silicon matched water* (SMW) (blue). The pink dotted line represents the bare amino-silylated silicon wafer in D₂O. Shading is to aid layer visualisation. Layer identities are suggested in black text.

Volume fractions for the three protein layers were estimated using *Equation 4 (section 6.6.3)*. SLDs for pure rSPA and IgG4 at three different levels of H-D exchange were determined using the Biomolecular Scattering Length Density Calculator at <http://psldc.isis.rl.ac.uk/Psldc/>. The volume fraction estimates are displayed in *Table 7.2*.

Table 7.2 Volume fraction estimates for layers of different sample configurations in *Cell A*. Volume fraction estimates were made based on the calculated SLD valued for the pure proteins at different levels of H-D exchange. SLDs for pure rSPA and IgG at 90%, 70% and 50% H-D exchange were calculated using the Biomolecular Scattering Length Density calculator at <http://psldc.isis.rl.ac.uk/Psldc/>. Volume fraction estimates were calculated using *equation 4 (section 7.4.3)*.

Sample configuration	Layer	SLD ($\times 10^{-6} \text{ \AA}^{-2}$)					Volume fraction		
		Protein H-D exchange			Fitted values		Protein H-D exchange		
		90%	70%	50%	Layer	D ₂ O	90%	70%	50%
rSPA	rSPA	3.37	3.07	3.00	6.04	6.17	0.05	0.04	0.04
rSPA + BSA	rSPA/BSA	3.37	3.07	3.00	5.95	6.14	0.07	0.06	0.06
	rSPA/BSA	3.28	2.99	2.82	5.42	6.14	0.25	0.23	0.22
IgG adsorbed	IgG inner	3.20	2.92	2.64	5.53	6.14	0.21	0.19	0.17
	IgG outer				5.89	6.14	0.09	0.08	0.07

For simplicity's sake, the following discussion will assume that 90% of hydrogen was exchanged to deuterium in both rSPA and IgG4. Looking at the values in *Table 7.2* we can see that when IgG was adsorbed the volume fraction of the first protein layer increased by 0.18, while the thickness stayed approximately the same as before at 60Å. If we compare the volume fraction contributions by protein A and IgG, we find that the contribution from IgG is 3.6 times that from protein A.

For the middle protein layer the volume fraction was 3% lower than the total of the first layer. The outermost layer had a volume fraction of 0.09, half that of the middle layer. The middle and outer protein layers had thicknesses of 52Å and 67Å, respectively. Thus, the three protein layers had a total thickness of 180Å. Capp et al. (2014) found the R_g of a single protein A domain to be 13Å. Doubling this and multiplying by the total number of IgG-binding domains gives a length of 130Å for fully extended linear protein A. This value does not account for the X_r domain

which would presumably add to the total length (*Fig. 7.2, Fig. 3.4*). Rayner et al. (2014) suggested that the maximum length of IgG4 was 160Å and reported the R_g of IgG4 at 20°C to be 48Å. With these factors in mind, it becomes quite difficult to interpret the data in terms of the most likely arrangement of IgG molecules bound to protein A and their respective orientations relative to the solid surface. The various layer depths could be applied in a number of ways, drawing comparisons with various different structural dimensions previously reported. Rabe et al. (2011) noted that reflectivity data provides high resolution in the surface normal direction, but lacks resolution in the lateral dimension. In some ways, this is at odds with the data presented in this chapter, as SLD values were found to be robust and layer thicknesses much less so. Nonetheless, without a computational method, data interpretation for a system containing more than one protein can become quite speculative.

Returning to *Fig. 7.10*, another factor to consider is the depth of the cross-linker layer after IgG adsorption, which dropped from 20-25Å to 5Å. This strongly suggests that adsorbed proteins associate closely with the solid surface at the interface.

In one possible configuration, adsorbed IgG could support protein A such that it becomes extended with its length perpendicular to the surface, and another IgG molecule could be bound to a surface-distal domain of protein A. The volume fraction of the outermost protein layer is double that of rSPA alone before IgG adsorption (*Table 7.2*), so some contribution to the outermost layer must come from IgG, even if protein A is fully extended. In another possible configuration, IgG could be bound to the outermost domain of protein A, causing it to flex to one side such that the IgG sits close to the solid surface.

The thickest and also least protein-dense layer was the outermost layer. Since the total protein depth, 180Å, was larger than the maximum length of IgG, 160Å (Rayner et al. 2014), and the densest protein layer was apparently closest to the surface, it is hypothesised that the three-layer structure was a result of two IgG molecules bound to protein A in side-on or skewed orientations. The deeper outermost layer could be the result of a single Fab arm extended into the solution.

7.5.3 IgG adsorption to rSPA-modified surface with PEG blocking

In *Cell B*, PEG₆₀₀₀ was used instead of BSA to block potential reactive sites on the surface after rSPA cross-linking. All other aspects of the experiment were the same as for *Cell A*. Reflectivity data collected at the various stages of surface modification is shown in *Figure 7.11*. The reflectivity curves lack visually obvious features, as was the case in *Cell A*. Similarly, model fitting resulted in SLD-distance profiles with distinct layers representative of the cross-linker and of protein A. The SLD-distance profile for the H₂O contrast suggested a fairly dense cross-linker and protein A layer. The addition of PEG caused a reduction in the SLD of the innermost layer (after SiO₂), though data was collected for the D₂O contrast only. The bootstrap error analysis of this model found that the fitted values for thickness and SLD of the innermost layer were not robust, with the bootstrap fitted profiles showing broad variation for these parameters. This could be due to weak association of PEG with the surface. Error analysis found protein A thickness and SLD parameters to be well fitted. In the absence of PEG the SPDP and protein A layers had thickness of 30Å and 58Å, respectively. We note that the total depth of these two layers is near-identical to that of the same sample configuration in *Cell A*.

The reflectivity data obtained after IgG4 adsorption is shown in *Figure 7.12*. The reflectivity curve in *Fig. 7.12.a* shares major characteristics with the equivalent sample in *Cell A*, but is by no means identical. It should be noted that the fit to the data points is not perfect; since there is a fringe in the fitted curve, at a Q value just below 0.1Å⁻¹, that falls a little outside the data points. It was not possible to eliminate this discrepancy with the resources available.

Looking at the fitted SLD profiles for adsorbed IgG in *Fig. 7.12.b* we see a similar structure to the one in *Cell A*. Notable differences include a thicker protein layer closest to the silica surface, and overall higher protein density. The SMW is clearly different, resolving almost no structure at the solid-liquid interface. This not an error in the data, rather the result of very similar SLD values for pure protein and silicon.

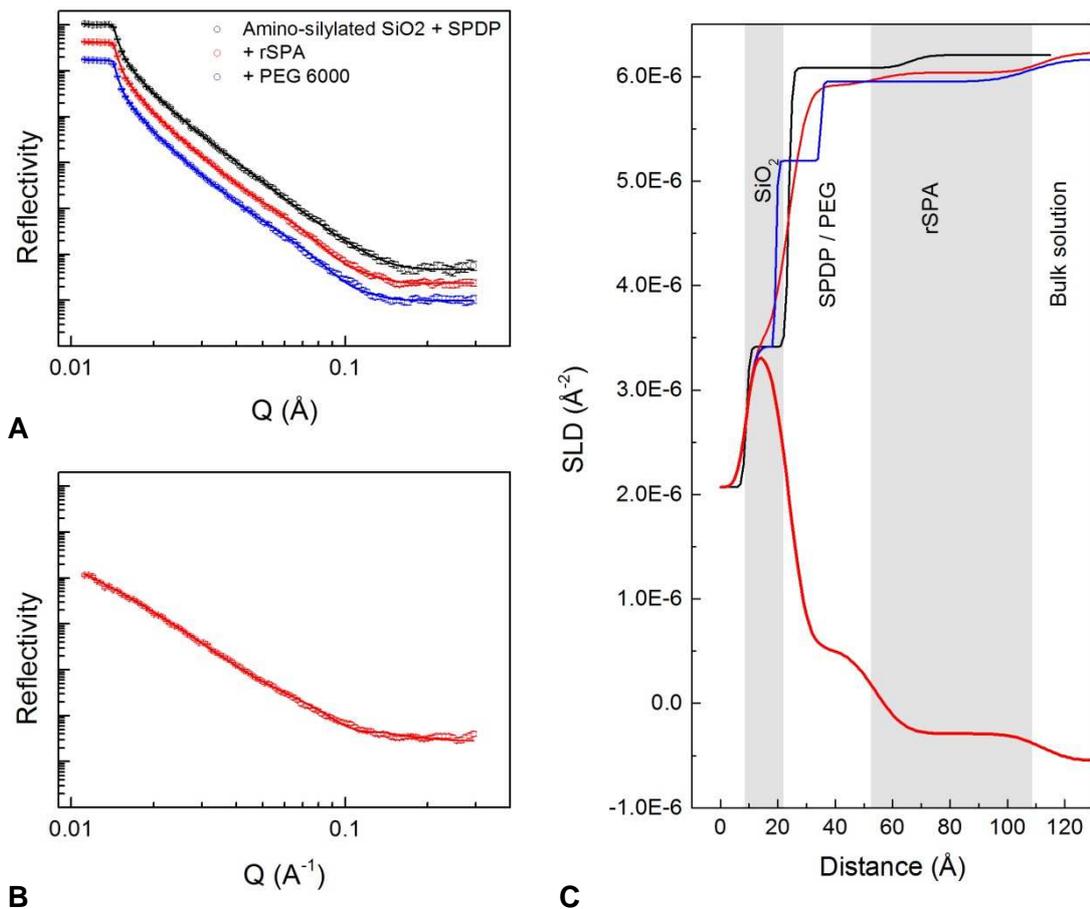


Figure 7.11 Reflectivity data for attachment of rSPA to an amino-silylated silica surface, and addition of the blocking agent PEG₆₀₀₀. Left-hand plots *a* and *b*: reflectivity (intensity) against Q (\AA^{-1}) in D_2O -based (*a*) and H_2O -based (*b*) buffer at various stages of surface modification. Data sets are offset down the y-axis for clarity; in D_2O the critical edge always occurs when reflectivity is equal to unity. Different colours represent different stages of surface modification. From top to bottom: black is the amino-silylated wafer with cross-linker attached; red is after rSPA cross-linking and blue is after blocking with PEG₆₀₀₀. Solid lines show the fitted models. Plot *c* shows the corresponding profiles of scattering length density, SLD (\AA^{-2}), against distance from the silicon surface. Where data was collected at both solution phase contrasts, the profile was generated by simultaneous fitting of D_2O and H_2O reflectivity data sets to a single layer depth profile. Curves that terminate at the upper and lower ends of the SLD scale represent D_2O and H_2O data, respectively. Shading is to aid layer visualisation and suggestions of layer identities are in rotated text.

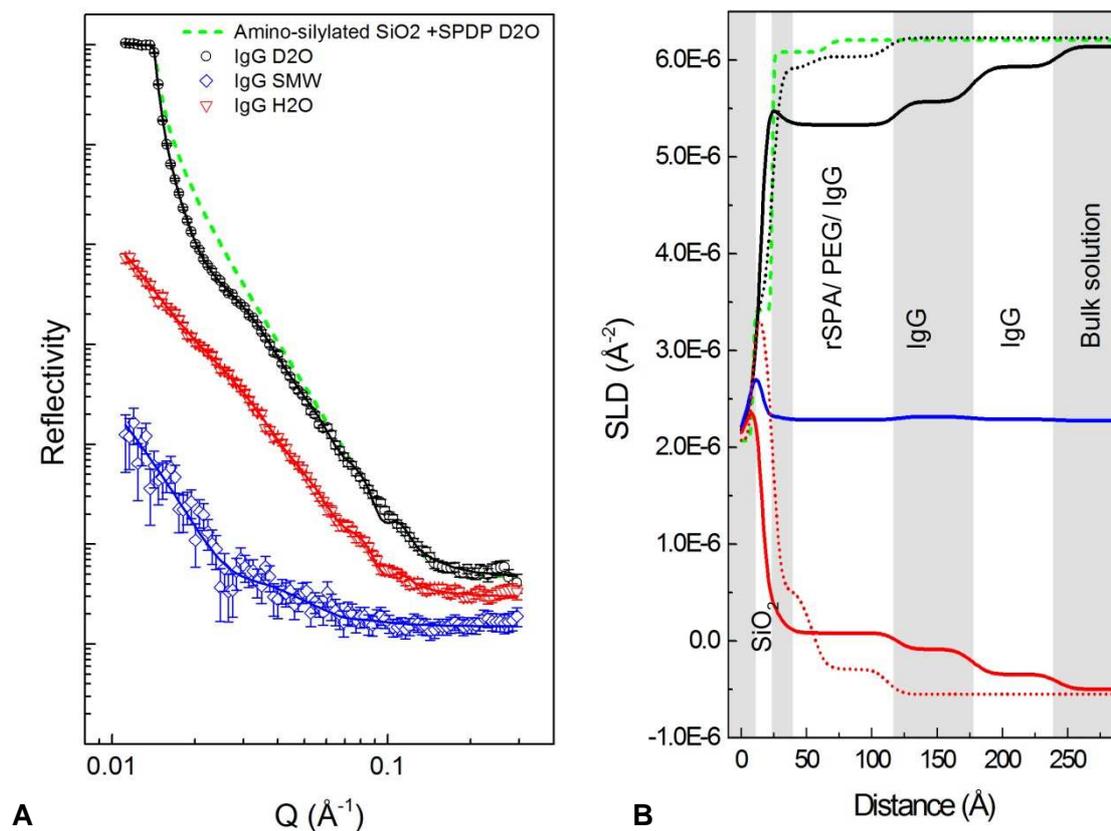


Figure 7.12 Reflectivity data for IgG4 adsorbed at pH 6.7 to rSPA-modified silica with PEG₆₀₀₀ blocking. Left-hand plot (a): reflectivity (intensity) against momentum transfer, Q (\AA^{-1}), at three solution phase contrasts: D_2O (black), H_2O (red) and *silicon matched water* (blue). H_2O and SMW data sets are offset down the y-axis to prevent overlap; in D_2O the critical edge always occurs when reflectivity is equal to unity. Solid lines represent fitted models. The fitted curve for the amino-silylated surface with cross-linker only is shown for comparison (*green dotted line*). Plot b shows the corresponding profiles of scattering length density (\AA^{-2}) against distance from the silicon surface. Data was fitted simultaneously for the three solution phase contrasts in a. Black, red and blue curves represent D_2O , H_2O and SMW data, respectively. SLD-distance profiles for the surface with cross-linked rSPA (dotted lines) and with cross-linker only (green dashed line) are shown for comparison. Shading is to aid layer visualisation and suggestions of layer identities are in rotated text.

The same surface modification strategy, including blocking with PEG₆₀₀₀, was implemented in the third SLFC, *Cell C*. The resulting data is displayed in *Figure 7.13*. The data produced in *Cells B* and *C* was very similar. This is a rather useful outcome as it shows PEG blocking consistently led to a different IgG configuration compared to BSA blocking. Another aspect was that it was more difficult to achieve a good fit for the data. The chi-squared value with IgG adsorbed was lower for *Cell C* than for *Cell A*, but only two contrasts were fitted to the *Cell C* data. The difficult aspect of the fitting related to the high Q region. This is associated with short-range structural features. It could be that PEG interactions at the surface produced a complex layer characteristic at the interface.

Volume fraction estimates were 0.28, 0.19 and 0.07 for the inner, middle and outer protein layers, respectively, in *Cell B* when IgG was adsorbed. Values for *Cell C* were very similar.

Reflectivity data for adsorbed IgG in the three sample cells is compared in *Figure 7.14*. Here, the innermost protein layer can be seen to almost double in length when PEG was used as a blocking agent instead of BSA. It is possible that PEG had repulsive effects that cause protein A to extend further into the solution phase. An interesting possibility is that IgG may have oriented with its length more perpendicular to the surface, in turn causing protein A to extend outwards and bind a second IgG molecule to its outermost domain, resulting in extension of IgG structures up to 250Å away from the surface.

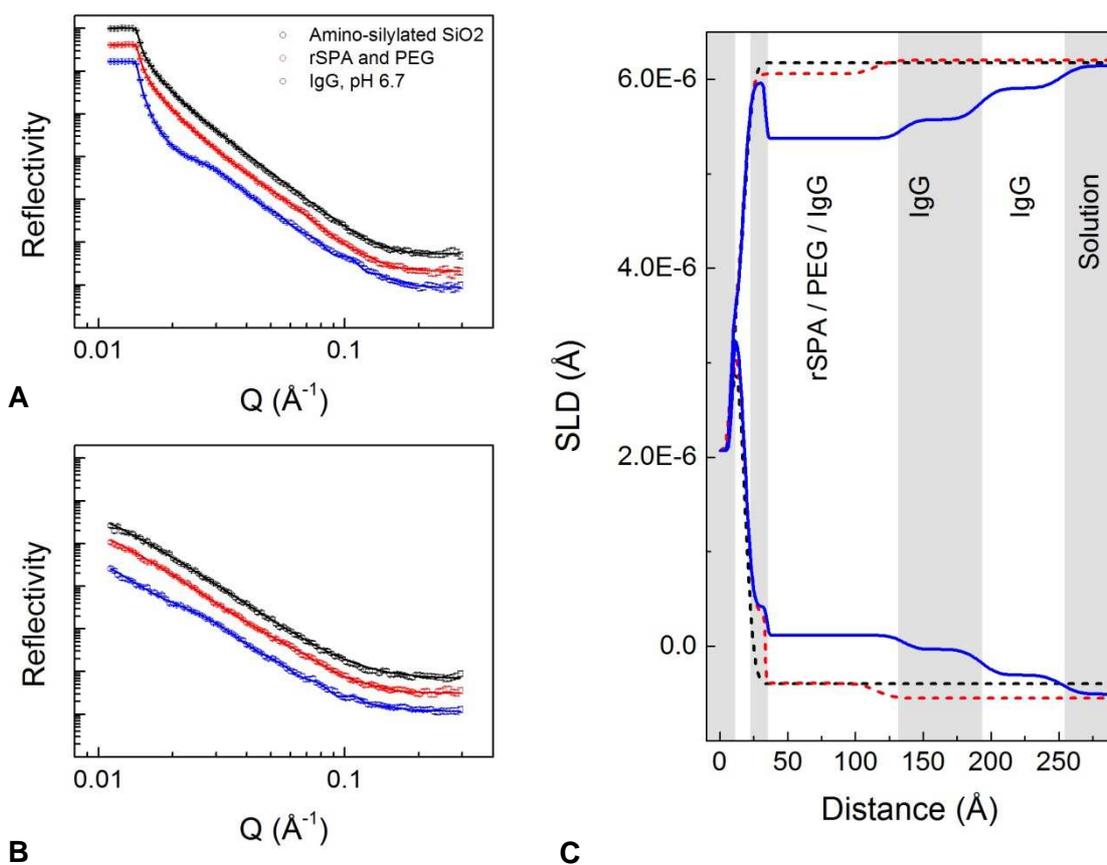


Figure 7.13 Reflectivity data for IgG4 adsorbed at pH 6.7 to rSPA-modified silica with PEG_{6000} blocking – Cell C. Left-hand plots a and b: reflectivity (intensity) against Q (\AA^{-1}) in D_2O -based (a) and H_2O -based (b) buffer at various stages of surface modification. Data sets are offset down the y-axis for clarity; in D_2O the critical edge always occurs when reflectivity is equal to unity. Different colours represent different stages of surface modification: black, amino-silylated surface; red, after rSPA cross-linking and PEG_{6000} blocking; blue, IgG adsorbed. Solid lines represent fitted models. Plot c shows the corresponding profiles of scattering length density against distance from the silicon surface. Each profile was generated by simultaneous fitting of D_2O and H_2O data sets to a single layer depth profile. Curves that terminate at the upper and lower ends of the SLD scale represent D_2O and H_2O data, respectively. Shading is to aid layer visualisation; layer identities are suggested for the final data set, i.e. IgG adsorbed.

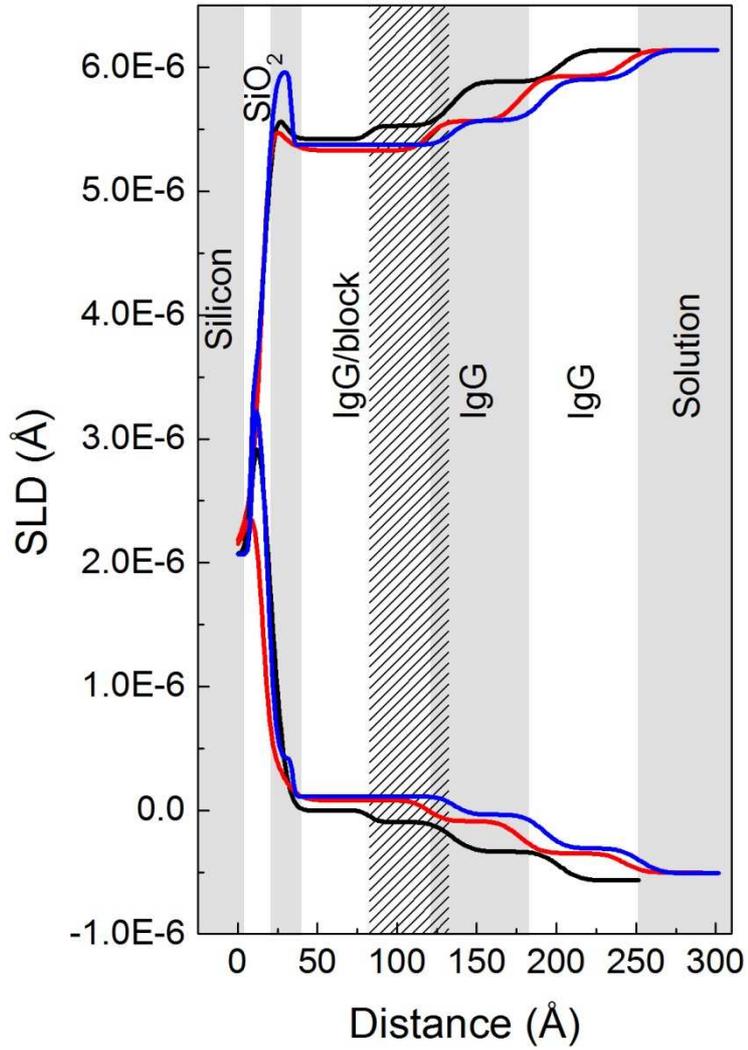


Figure 7.14 SLD-Distance profiles for IgG adsorbed to protein A with BSA surface blocking (black line) and PEG₆₀₀₀ surface blocking in *Cell B* (red line) and *Cell C* (blue line).

Compared to investigations of antibodies adsorbing directly to silica surfaces, as discussed in *section 6*, few studies have used neutron reflection to characterise the structure of complex interfacial assemblies involving several biological and/or organic components. However, one recent study (Le Brun et al. 2011) used polarised neutron reflection and surface plasmon resonance to characterise the structure and function of an oriented protein array in complex with an antibody which subsequently bound an antigen. The structures that were characterised had many similarities to those investigated in this work. The protein construct, termed *ZZctOmpA*, consisted of a beta-barrel domain with a C terminal cysteine and two protein A Z domains (see *Figure 3.4, section 3.3.2 Protein A*). An oriented layer of *ZZctOmpA* was immobilised on a gold surface via the protein construct's C terminal cysteine. Bare surface patches were then filled by addition of amphiphilic thioPEG, which binds to the surface and orients with hydrophilic head groups directed out into the solution phase (Le Brun et al. 2011). Neutron reflectivity data collected at the various stages of surface modification and after antibody addition had similarities with the data presented in this work. One important observation related to assembly of the *ZZctOmpA* array. Neutron measurements made after *ZZctOmpA* attachment showed a poorly-defined layer with high roughness and relatively low thickness. However, after thioPEG addition the layer became well-defined and had significantly increased depth. The increase in layer depth was deduced to be due to a more uniform and upright orientation of the protein component, because thioPEG alone (without protein) produced a much thinner layer with a much lower SLD. Thus thioPEG was found to yield the desired effect of "filling out" the surface layer and causing the protein construct to orient away from the surface. A similar result was apparent when PEG was used as a blocking molecule in the work presented here. Looking at *Fig. 7.11* we can see that before PEG addition the rSPA layer had a high level of roughness at the SPDP-rSPA transition; after PEG addition the rSPA layer became notably more defined. The SLD of the rSPA layer was also similar to that obtained by Le Brun et al. (2011) for the *ZZctOmpA* only layer. Further similarity can be seen for neutron measurements made after antibody addition, with a total layer(s) depth of around 200Å and an SLD of around $5.5 \times 10^{-6} \text{Å}^{-2}$ for the antibody/*ZZctOmpA* layer. This was interpreted as 1:1 ratio of *ZZctOmpA* to antibody,

with the antibody in an “upright” orientation (with Fab regions pointing out into the solution) (Le Brun et al. 2011). The data was modelled such that the antibody region was well-fitted as a single layer with an SLD of around $5.5 \times 10^{-6} \text{ \AA}^{-2}$. This differs to the data presented here, in which models after IgG addition showed three IgG-related layers, suggestive of skewed IgG orientation and a higher ratio of IgG to protein A.

Many aspects of the data generated by Le Brun et al. (2011) lend support to the data interpretation in this work. Particularly, we see agreement in the influence of a PEG-based blocker/spacer on the orientation and accessibility of flexible immobilised proteins, and subsequent changes in the layer depths of adsorbed antibody.

8 Conclusion

The overall objective of this work was to advance understanding of how adsorption-mediated processes might influence antibody aggregation phenomena in bioprocessing. Emphasis was placed on affinity chromatography, namely protein A, because it represents a very widespread purification step used in many bioprocesses; it has also been identified as a point of aggregate formation in such bioprocesses.

The first objective was to demonstrate how elution from protein A affected the aggregation rate of IgG4 that was incubated in the elution buffer. Incubation in the elution buffer was representative of low pH hold methods that are often used for virus inactivation. This objective was addressed in the first results chapter, *Chapter 5*. The results showed that the rate of IgG aggregation was approximately 5-fold higher after elution from protein A, compared to in solution only. The data gathered captured the behaviour of an IgG4 within a fairly compact experimental space. The IgG4 demonstrated highly pH-dependent and apparently concentration-independent aggregation behaviour. Given these observations, it was hypothesised that the aggregation behaviour was governed predominantly by a pH-dependent unfolding/ re-folding equilibrium. The chromatography step was shown to increase rates of aggregation, but other aspects of the aggregation kinetics appeared largely un-affected. One possible molecular mechanism is that the conformational change, as well as changes in the hydration layer around the protein, required for protein adsorption and subsequent desorption, alter the structure of the IgG, possibly exposing regions of the protein involved in unfolding transitions. Being in this condition, immediate exposure to the acidic environment could increase the possibility of the protein unfolding and losing native structure to some extent before progression to aggregate formation during the pH hold. It is believed that the work completed in *Chapter 5* fulfilled the first objective. The data presented uniquely demonstrates how the chromatography step may drive an unfolding/refolding equilibrium towards higher pH values. In this way, the findings are practically valuable to the bioprocess industry.

The second objective was to gain insight into the aggregation mechanisms at play. Straightforward methods of aggregate characterisation, such as SEC and Native-PAGE, were used to compare column-induced and solution (low pH) aggregates. The results did not yield any striking findings, showing the two aggregate populations to be much the same. Thus, a new approach was considered, in which the objective would be to characterise the structures of adsorbed IgG molecules before aggregation took place.

In *Chapter 6*, neutron reflectivity was used to characterise the structure of IgG adsorbed to the model surface silica under a range of buffer conditions. The findings provided a detailed insight into the behaviour of IgG 4 at the solid-liquid interface. When the IgG was adsorbed at pH 4.1 it demonstrated an apparently flat-on orientation, pertaining to maximum possible contact with the surface for the folded protein. This is assumed to be preferential due to the net positive charge on the IgG at pH 4 causing attraction to the silica surface. The surface coverage of the adsorbed IgG was estimated to be 40% which represents a closely packed layer in the flat-on orientation. While the IgG solution remained in the reflectivity cell a fraction of weakly adsorbed IgG was also observed. Data suggested partial re-orientation of this material into the flat-on configuration when the IgG solution was rinsed out of the cell. At pH 6.2, surface coverage was found to be slightly higher, and data was indicative of an IgG orientation less-closely and less strongly associated with the silica surface, possibly side-on or in tilted orientation. In the context of this work, an observation of particular interest was the distinct shift in configuration towards increased surface contact when the pH was lowered from 6.2 to 3.7. The data demonstrated that electrostatic attraction was a strong driving force for surface adsorption, and protein-protein repulsion influenced packing density of the interfacial layer. The data presented in *Chapter 6* provided a strong basis for using the reflectivity technique to examine IgG structures adsorbed to more complex model surfaces.

The work presented in *Chapter 7* first describes how a model surface suitable for reflectivity experiments was generated and examined using ATR-FTIR. The model surface included recombinant staphylococcal protein A, cross-linked to the surface, and one of two blocking

molecules, BSA or PEG₆₀₀₀. FTIR analysis demonstrated that the model surface was reproducible and the structure of protein A was retained such that it was able to bind IgG4. The model surface was subsequently characterised using neutron reflectivity. Analysis of the reflectivity data revealed layer-based model structures in which it was possible to identify distinct contributions from protein A and from IgG. It was clear that IgG was adsorbed to protein A, and not non-specifically adsorbed to the surface, partly because the data was strikingly different to that described in *Chapter 6*. Key structural features included three protein layers extending to a maximum of 250Å away from the solid surface into the solution. Layer volume fractions showed robust fits for all three sample cells used, suggesting that the IgG orientation was specific and consistent. The data found the different blocking molecules to influence the IgG layer structure, with the use of the PEG₆₀₀₀ as a block resulting in more extended structures and increased protein volume fraction. Though the data was reproducible and data fits largely robust, it was difficult to interpret the structures based on the data sets acquired. Nonetheless, the data gave a new insight into the arrangement of IgG adsorbed to immobilised protein A in terms of the distribution of protein layers in the surface normal direction, and their respective volume fractions. The influence of other surface molecules on the orientation of the IgG and its interaction with protein A was also highlighted.

9 Future Work

The data described in this thesis provides a broad scope for future investigations.

Leading on from the first results chapter, *Chapter 5*, investigation into the aggregation behaviour of the IgG across the whole peak is suggested. The leading side of the peak was not investigated here, partly because its pH was too high. We predict that much could be learned from inducing aggregation in the front peak fractions and assessing aggregation rate.

One aspect of the objectives that was not fulfilled was determining the influence of chromatography parameters on aggregation rates.. Some preliminary work was done suggesting that the protein concentration in the load material may influence aggregation rate. Interestingly, rates seemed to increase at lower protein loads. Other parameters of interest were the elution buffer molarity, and the step in molarity between load/wash phases and elution.

On the neutron measurement side, there are many possibilities. The work presented here feels somewhat a prelude to a much more comprehensive body of work. The key element missing from the neutron work we have presented was an ability to relate the findings to subsequent aggregation behaviour of de-adsorbed or eluted proteins. Part of the reason for this is that the sample configurations used involved adsorption of nanomolar quantities of material (since surfaces were completely flat), making quantification of subsequent aggregates difficult.

Some preliminary work was carried out at the NCNR in which silica beads from Millipore (the base beads of Prosep) were prepared in SANS cells at a range a solution phase contrasts. The match-point for the beads was determined at 60.5% D₂O. This means that the beads were “invisible” to the neutrons. Attachment of deuterated protein A to the beads and adsorption of IgG could allow adsorbed structures inside the beads to be characterised. Further, aggregate formation inside the beads could be monitored with SANS.

When the proposal for the work at the ISIS Neutron and Muon Source was first submitted, the planned investigation involved comparing IgG conformation and orientation at different protein A ligand densities. It was hypothesised that at higher ligand densities there would be more steric

hindrance and reduced rotational freedoms for adsorbed antibodies. Studies focussing on enthalpy and entropy changes on adsorption and desorption in this type of configuration have suggested less structural re-arrangement at higher packing densities. Structural information in this field would provide a very valid contribution to understanding.

10 References

- Andrews, J. M. and Roberts, C. J. (2007) Non-Native Aggregation of α -Chymotrypsinogen Occurs through Nucleation and Growth with Competing Nucleus Sizes and Negative Activation Energies†. *Biochemistry*, 46(25), pp. 7558-7571.
- Arakawa, T., Ejima, D., Li, T. and Philo, J. S. (2010) The critical role of mobile phase composition in size exclusion chromatography of protein pharmaceuticals. *J Pharm Sci*, 99(4), pp. 1674-92.
- Arrondo, J. L. R., Muga, A., Castresana, J. and Goñi, F. M. (1993) Quantitative studies of the structure of proteins in solution by fourier-transform infrared spectroscopy. *Progress in Biophysics and Molecular Biology*, 59(1), pp. 23-56.
- Barnett, G. V., Razinkov, V. I., Kerwin, B. A., Laue, T. M., Woodka, A. H., Butler, P. D., Perevozchikova, T. and Roberts, C. J. (2015) Specific-Ion Effects on the Aggregation Mechanisms and Protein-Protein Interactions for Anti-streptavidin Immunoglobulin Gamma-1. *The Journal of Physical Chemistry B*.
- Barone, G., Capasso, S., Del Vecchio, P., De Sena, C., Fessas, D., Giancola, C., Graziano, G. and Tramonti, P. (1995) Thermal denaturation of bovine serum albumin and its oligomers and derivatives pH dependence. *Journal of thermal analysis*, 45(6), pp. 1255-1264.
- Barth, A. (2007) Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1767(9), pp. 1073-1101.
- Bernstein, H. J. (2000) Recent changes to RasMol, recombining the variants. *Trends in Biochemical Sciences*, 25(9), pp. 453-455.
- Beynon, R. J., Beynon, P. R. and Easterby, J. (1996) *Buffer Solutions, The Basics (Garland Science)*, Taylor & Francis.
- Białopiotrowicz, T. and Jańczuk, B. (2001) Wettability and surface free energy of bovine serum albumin films. *Journal of Surfactants and Detergents*, 4(3), pp. 287-292.

- Biddlecombe, J. G., Smith, G., Uddin, S., Mulot, S., Spencer, D., Gee, C., Fish, B. C. and Bracewell, D. G. (2009) Factors influencing antibody stability at solid-liquid interfaces in a high shear environment. *Biotechnol Prog*, 25(5), pp. 1499-507.
- Bolli, R., Woodtli, K., Bärtschi, M., Höfferer, L. and Lerch, P. (2010) L-Proline reduces IgG dimer content and enhances the stability of intravenous immunoglobulin (IVIG) solutions. *Biologicals*, 38(1), pp. 150-157.
- Bork, P., Holm, L. and Sander, C. (1994) The Immunoglobulin Fold: Structural Classification, Sequence Patterns and Common Core. *Journal of Molecular Biology*, 242(4), pp. 309-320.
- Bottomley, S. P., Sutton, B. J. and Gore, M. G. (1995) Elution of human IgG from affinity columns containing immobilised variants of protein A. *J Immunol Methods*, 182(2), pp. 185-92.
- Busini, V., Moiani, D., Moscatelli, D., Zamolo, L. and Cavallotti, C. (2006) Investigation of the Influence of Spacer Arm on the Structural Evolution of Affinity Ligands Supported on Agarose. *The Journal of Physical Chemistry B*, 110(46), pp. 23564-23577.
- Capp, Jo A., Hagarman, A., Richardson, David C. and Oas, Terrence G. (2014) The Statistical Conformation of a Highly Flexible Protein: Small-Angle X-Ray Scattering of S. aureus Protein A. *Structure*, 22(8), pp. 1184-1195.
- Carlsson, J., Drevin, H. and Axén, R. (1978) Protein thiolation and reversible protein-protein conjugation. N-Succinimidyl 3-(2-pyridyldithio)propionate, a new heterobifunctional reagent. *Biochemical Journal*, 173(3), pp. 723-737.
- Chan, H. S. (2001) Amino Acid Side-chain Hydrophobicity. in *eLS*: John Wiley & Sons, Ltd.
- Chi, E. Y., Krishnan, S., Randolph, T. W. and Carpenter, J. F. (2003) Physical stability of proteins in aqueous solution: Mechanism and driving forces in nonnative protein aggregation. *Pharm Res*, 20(9), pp. 1325-1336.
- Clifton, L. A., Neylon, C. and Lakey, J. H. (2013) Examining protein-lipid complexes using neutron scattering. *Methods Mol Biol*, 974, pp. 119-50.

- Clifton, L. A., Sanders, M., Kinane, C., Arnold, T., Edler, K. J., Neylon, C., Green, R. J. and Frazier, R. A. (2012) The role of protein hydrophobicity in thionin-phospholipid interactions: a comparison of [small alpha]1 and [small alpha]2-purothionin adsorbed anionic phospholipid monolayers. *Physical Chemistry Chemical Physics*, *14*(39), pp. 13569-13579.
- Clifton, L. A., Sanders, M. R., Hughes, A. V., Neylon, C., Frazier, R. A. and Green, R. J. (2011) Lipid binding interactions of antimicrobial plant seed defence proteins: puroindoline-a and [small beta]-purothionin. *Physical Chemistry Chemical Physics*, *13*(38), pp. 17153-17162.
- Clifton, L. A., Skoda, M. W. A., Le Brun, A. P., Ciesielski, F., Kuzmenko, I., Holt, S. A. and Lakey, J. H. (2015) Effect of Divalent Cation Removal on the Structure of Gram-Negative Bacterial Outer Membrane Models. *Langmuir*, *31*(1), pp. 404-412.
- Couston, R. G., Skoda, M. W., Uddin, S. and van der Walle, C. F. (2012) Adsorption behavior of a human monoclonal antibody at hydrophilic and hydrophobic surfaces. *MAbs*, *5*(1), pp. 126-139.
- Cozzone, A. J. (2001) Proteins: Fundamental Chemical Properties. in *eLS*: John Wiley & Sons, Ltd.
- Cromwell, M. E., Hilario, E. and Jacobson, F. (2006) Protein aggregation and bioprocessing. *AAPS J*, *8*(3), pp. E572-9.
- Cuatrecasas, P. (1970) Protein Purification by Affinity Chromatography: DERIVATIZATIONS OF AGAROSE AND POLYACRYLAMIDE BEADS. *Journal of Biological Chemistry*, *245*(12), pp. 3059-3065.
- de Bergevin, F. (2009) The Interaction of X-Rays (and Neutrons) with Matter. in Daillant, J. and Gibaud, A., (eds.) *X-ray and Neutron Reflectivity*. Springer Berlin Heidelberg. pp. 1-57.
- Deis, L. N., Wu, Q., Wang, Y., Qi, Y., Daniels, K. G., Zhou, P. and Oas, T. G. (2015) Suppression of conformational heterogeneity at a protein-protein interface. *Proceedings of the National Academy of Sciences*, *112*(29), pp. 9028-9033.

- Deisenhofer, J. (1981) Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9- and 2.8-Å resolution. *Biochemistry*, *20*(9), pp. 2361-70.
- DeLano, W. L. (2000) Convergent Solutions to Binding at a Protein-Protein Interface. *Science*, *287*(5456), pp. 1279-1283.
- den Engelsman, J., Garidel, P., Smulders, R., Koll, H., Smith, B., Bassarab, S., Seidl, A., Hainzl, O. and Jiskoot, W. (2011) Strategies for the assessment of protein aggregates in pharmaceutical biotech product development. *Pharm Res*, *28*(4), pp. 920-33.
- Drake, A. (1994) Circular Dichroism. in Jones, C., Mulloy, B. and Thomas, A., (eds.) *Microscopy, Optical Spectroscopy, and Macroscopic Techniques*: Humana Press. pp. 219-244.
- Ejima, D., Tsumoto, K., Fukada, H., Yumioka, R., Nagase, K., Arakawa, T. and Philo, J. S. (2007) Effects of acid exposure on the conformation, stability, and aggregation of monoclonal antibodies. *Proteins*, *66*(4), pp. 954-62.
- Ejima, D., Yumioka, R., Arakawa, T. and Tsumoto, K. (2005) Arginine as an effective additive in gel permeation chromatography. *Journal of Chromatography A*, *1094*(1-2), pp. 49-55.
- Fahrner, R. L., Whitney, D. H., Vanderlaan, M. and Blank, G. S. (1999) Performance comparison of Protein A affinity-chromatography sorbents for purifying recombinant monoclonal antibodies. *Biotechnology and Applied Biochemistry*, *30*(2), pp. 121-128.
- Fekete, S., Beck, A., Veuthey, J. L. and Guilleme, D. (2014) Theory and practice of size exclusion chromatography for the analysis of protein aggregates. *J Pharm Biomed Anal*, *101C*, pp. 161-173.
- Fermon, C., Ott, F. and Menelle, A. (2009) Neutron Reflectometry. in Daillant, J. and Gibaud, A., (eds.) *X-ray and Neutron Reflectivity*: Springer Berlin Heidelberg. pp. 183-234.
- Fernandez, A., Kardos, J., Scott, L. R., Goto, Y. and Berry, R. S. (2003) Structural defects and the diagnosis of amyloidogenic propensity. *Proc Natl Acad Sci U S A*, *100*(11), pp. 6446-51.

- Filipe, V., Kukrer, B., Hawe, A. and Jiskoot, W. (2012) Transient molten globules and metastable aggregates induced by brief exposure of a monoclonal IgG to low pH. *J Pharm Sci*, 101(7), pp. 2327-39.
- Firer, M. A. (2001) Efficient elution of functional proteins in affinity chromatography. *J Biochem Biophys Methods*, 49(1-3), pp. 433-42.
- Fischer, H., Polikarpov, I. and Craievich, A. F. (2004) Average protein density is a molecular-weight-dependent function. *Protein Science : A Publication of the Protein Society*, 13(10), pp. 2825-2828.
- Gagnon, P., Nian, R., Leong, D. and Hoi, A. (2015) Transient conformational modification of immunoglobulin G during purification by protein A affinity chromatography. *Journal of Chromatography A*, 1395(0), pp. 136-142.
- Garber, E. and Demarest, S. J. (2007) A broad range of Fab stabilities within a host of therapeutic IgGs. *Biochem Biophys Res Commun*, 355(3), pp. 751-7.
- Goding, J. W. (1996) 8 - Production of Monoclonal Antibodies. in Goding, J. W., (ed.) *Monoclonal Antibodies (Third Edition)*, London: Academic Press. pp. 141-191.
- Goormaghtigh, E., Cabiaux, V. and Ruyschaert, J.-M. (1990) Secondary structure and dosage of soluble and membrane proteins by attenuated total reflection Fourier-transform infrared spectroscopy on hydrated films. *European Journal of Biochemistry*, 193(2), pp. 409-420.
- Gouda, H., Torigoe, H., Saito, A., Sato, M., Arata, Y. and Shimada, I. (1992) Three-dimensional solution structure of the B domain of staphylococcal protein A: comparisons of the solution and crystal structures. *Biochemistry*, 31(40), pp. 9665-9672.
- Hagel, L. (2011) Gel Filtration: Size Exclusion Chromatography. in *Protein Purification*: John Wiley & Sons, Inc. pp. 51-91.
- Hahn, R., Schlegel, R. and Jungbauer, A. (2003) Comparison of protein A affinity sorbents. *J Chromatogr B Analyt Technol Biomed Life Sci*, 790(1-2), pp. 35-51.

- Haimer, E., Tscheliessnig, A., Hahn, R. and Jungbauer, A. (2007) Hydrophobic interaction chromatography of proteins IV. Kinetics of protein spreading. *Journal of Chromatography A*, 1139(1), pp. 84-94.
- Hari, S. B., Lau, H., Razinkov, V. I., Chen, S. and Latypov, R. F. (2010) Acid-induced aggregation of human monoclonal IgG1 and IgG2: molecular mechanism and the effect of solution composition. *Biochemistry*, 49(43), pp. 9328-38.
- Harris, R. P. and Kilby, P. M. (2014) Amino acid misincorporation in recombinant biopharmaceutical products. *Curr Opin Biotechnol*, 30C, pp. 45-50.
- Hermanson, G. T. (2013a) Chapter 3 - The Reactions of Bioconjugation. in Hermanson, G. T., (ed.) *Bioconjugate Techniques (Third edition)*, Boston: Academic Press. pp. 229-258.
- Hermanson, G. T. (2013b) Chapter 6 - Heterobifunctional Crosslinkers. in Hermanson, G. T., (ed.) *Bioconjugate Techniques (Third edition)*, Boston: Academic Press. pp. 299-339.
- Hermanson, G. T. (2013c) Chapter 15 - Immobilization of Ligands on Chromatography Supports. in Hermanson, G. T., (ed.) *Bioconjugate Techniques (Third edition)*, Boston: Academic Press. pp. 589-740.
- Hober, S., Nord, K. and Linhult, M. (2007) Protein A chromatography for antibody purification. *J Chromatogr B Analyt Technol Biomed Life Sci*, 848(1), pp. 40-7.
- Houde, D., Arndt, J., Domeier, W., Berkowitz, S. and Engen, J. R. (2009) Characterization of IgG1 Conformation and Conformational Dynamics by Hydrogen/Deuterium Exchange Mass Spectrometry. *Analytical Chemistry*, 81(7), pp. 2644-2651.
- Hubbard, R. E. and Kamran Haider, M. (2001) Hydrogen Bonds in Proteins: Role and Strength. in *eLS*: John Wiley & Sons, Ltd.
- Hughes, A. V. (2013) RasCAL. in, <http://sourceforge.net/projects/rscl/>: Sourceforge.
- Ishikawa, T., Ito, T., Endo, R., Nakagawa, K., Sawa, E. and Wakamatsu, K. (2010) Influence of pH on Heat-Induced Aggregation and Degradation of Therapeutic Monoclonal Antibodies. *Biological and Pharmaceutical Bulletin*, 33(8), pp. 1413-1417.

- Jansson, B., Uhlén, M. and Nygren, P.-Å. (1998) All individual domains of staphylococcal protein A show Fab binding. *FEMS Immunol Med Microbiol* 20(1), pp. 69-78.
- Jeyachandran, Y. L., Mielczarski, J. A., Mielczarski, E. and Rai, B. (2010) Efficiency of blocking of non-specific interaction of different proteins by BSA adsorbed on hydrophobic and hydrophilic surfaces. *Journal of Colloid and Interface Science*, 341(1), pp. 136-142.
- Jungbauer, A., Machold, C. and Hahn, R. (2005) Hydrophobic interaction chromatography of proteins. III. Unfolding of proteins upon adsorption. *Journal of Chromatography A*, 1079(1-2), pp. 221-8.
- Kienzle, P. A., O'Donovan, K. V., Ankner, J. F., Berk, N. F. and Majkrzak, C. F. (2000-2006) Reflpak reflectometry software. in <http://www.ncnr.nist.gov/reflpak>.
- Kong, J. and Yu, S. (2007) Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures. *Acta Biochimica et Biophysica Sinica*, 39(8), pp. 549-559.
- Krapp, S., Mimura, Y., Jefferis, R., Huber, R. and Sonderrmann, P. (2003) Structural Analysis of Human IgG-Fc Glycoforms Reveals a Correlation Between Glycosylation and Structural Integrity. *Journal of Molecular Biology*, 325(5), pp. 979-989.
- Krężel, A. and Bal, W. (2004) A formula for correlating pKa values determined in D2O and H2O. *Journal of Inorganic Biochemistry*, 98(1), pp. 161-166.
- Lain, B. (2013) Protein A. *BioProcess International*, 11(8), pp. 29-38.
- Le Brun, A. P., Holt, S. A., Shah, D. S. H., Majkrzak, C. F. and Lakey, J. H. (2011) The structural orientation of antibody layers bound to engineered biosensor surfaces. *Biomaterials*, 32(12), pp. 3303-3311.
- Liepinsh, E. and Otting, G. (1996) Proton exchange rates from amino acid side chains—implications for image contrast. *Magnetic Resonance in Medicine*, 35(1), pp. 30-42.
- Liu, B., Huang, P.-J. J., Zhang, X., Wang, F., Pautler, R., Ip, A. C. F. and Liu, J. (2013) Parts-per-Million of Polyethylene Glycol as a Non-Interfering Blocking Agent for

- Homogeneous Biosensor Development. *Analytical Chemistry*, 85(21), pp. 10045-10050.
- Lowe, C. R. (2001) Combinatorial approaches to affinity chromatography. *Current Opinion in Chemical Biology*, 5(3), pp. 248-256.
- Lu, J. R., Zhao, X. and Yaseen, M. (2007) Protein adsorption studied by neutron reflection. *Current Opinion in Colloid & Interface Science*, 12(1), pp. 9-16.
- Lumry, R. and Eyring, H. (1954) Conformation Changes of Proteins. *The Journal of Physical Chemistry*, 58(2), pp. 110-120.
- Lund, L. N., Christensen, T., Toone, E., Houen, G., Staby, A. and St. Hilaire, P. M. (2011) Exploring variation in binding of Protein A and Protein G to immunoglobulin type G by isothermal titration calorimetry. *Journal of Molecular Recognition*, 24(6), pp. 945-952.
- Lydyard, P. M., Whelan, A. and Fanger, M. W. (2011) *BIOS Instant Notes in Immunology*, Taylor & Francis Ltd.
- Mahler, H. C., Friess, W., Grauschopf, U. and Kiese, S. (2009) Protein aggregation: pathways, induction factors and analysis. *J Pharm Sci*, 98(9), pp. 2909-34.
- McGee, W. M., Mentinova, M. and McLuckey, S. A. (2012) Gas-Phase Conjugation to Arginine Residues in Polypeptide Ions via N-Hydroxysuccinimide Ester-Based Reagent Ions. *Journal of the American Chemical Society*, 134(28), pp. 11412-11414.
- Moks, T., AbrahmsÉN, L., Nilsson, B., Hellman, U., SjöQuist, J. and UhlÉN, M. (1986) Staphylococcal protein A consists of five IgG-binding domains. *European Journal of Biochemistry*, 156(3), pp. 637-643.
- Nakanishi, K., Sakiyama, T., Kumada, Y., Imamura, K. and Imanaka, H. (2008) Recent Advances in Controlled Immobilization of Proteins onto the Surface of the Solid Substrate and Its Possible Application to Proteomics. *Current Proteomics*, 5(3), pp. 161-175.
- Nilsson, B., Moks, T., Jansson, B., Abrahmsén, L., Elmblad, A., Holmgren, E., Henrichson, C., Jones, T. A. and Uhlén, M. (1987) A synthetic IgG-binding domain based on staphylococcal protein A. *Protein Engineering*, 1(2), pp. 107-113.

- Pabst, T. M., Palmgren, R., Forss, A., Vasic, J., Fonseca, M., Thompson, C., Wang, W. K., Wang, X. and Hunter, A. K. (2014) Engineering of novel Staphylococcal Protein A ligands to enable milder elution pH and high dynamic binding capacity. *Journal of Chromatography A*, 1362, pp. 180-5.
- Paul, A. J., Schwab, K. and Hesse, F. (2014) Direct analysis of mAb aggregates in mammalian cell culture supernatant. *BMC Biotechnol*, 14(1), pp. 99.
- Perevozchikova, T., Nanda, H., Nesta, D. P. and Roberts, C. J. (2015) Protein Adsorption, Desorption, and Aggregation Mediated by Solid–Liquid Interfaces. *J Pharm Sci*, pp. n/a-n/a.
- Peysner, J. R. (2010) Nucleic acids encoding recombinant protein A. in: Google Patents.
- Philo, J. (2006) Is any measurement method optimal for all aggregate sizes and types? *AAPS J*, 8(3), pp. E564-E571.
- Pradyot, P. (2004) Optical Activity and Rotatory Dispersion. in *Dean's Analytical Chemistry Handbook, Second Edition*: McGraw Hill Professional, Access Engineering.
- Pynn, R. (2009) Neutron Scattering - A Non-destructive Microscope for Seeing Inside Matter. in Liang, L., Rinaldi, R. and Schober, H., (eds.) *Neutron applications in earth, energy and environmental sciences*, US: Springer.
- Rabe, M., Verdes, D. and Seeger, S. (2011) Understanding protein adsorption phenomena at solid surfaces. *Advances in Colloid and Interface Science*, 162(1–2), pp. 87-106.
- Raschke, T. M. and Marqusee, S. (1998) Hydrogen exchange studies of protein structure. *Curr Opin Biotechnol*, 9(1), pp. 80-86.
- Rayner, L. E., Hui, G. K., Gor, J., Heenan, R. K., Dalby, P. A. and Perkins, S. J. (2014) The Fab Conformations in the Solution Structure of Human Immunoglobulin G4 (IgG4) Restrict Access to Its Fc Region: Implications For Functional Activity. *Journal of Biological Chemistry*, 289(30), pp. 20740-20756.
- Roberts, C. J. (2007) Non-native protein aggregation kinetics. *Biotechnol Bioeng*, 98(5), pp. 927-38.

- Roberts, C. J. (2014) Therapeutic protein aggregation: mechanisms, design, and control. *Trends Biotechnol*, 32(7), pp. 372-380.
- Roque, A. C. A., Silva, C. S. O. and Taipa, M. Â. (2007) Affinity-based methodologies and ligands for antibody purification: Advances and perspectives. *Journal of Chromatography A*, 1160(1-2), pp. 44-55.
- Salvalaglio, M., Zamolo, L., Busini, V., Moscatelli, D. and Cavallotti, C. (2009) Molecular modeling of protein A affinity chromatography. *Journal of Chromatography A*, 1216(50), pp. 8678-86.
- Samuel, Z. and Harris, J. M. (1997) Introduction to Chemistry and Biological Applications of Poly(ethylene glycol). in *Poly(ethylene glycol)*: American Chemical Society. pp. 1-13.
- Sayle, R. and Milner-White, E. J. (1995) RasMol: Biomolecular graphics for all. *Trends in Biochemical Sciences*, 20(9), pp. 374.
- Seshadri, S., Khurana, R. and Fink, A. L. (1999) [36] Fourier transform infrared spectroscopy in analysis of protein deposits. in *Methods in Enzymology*. Academic Press. pp. 559-576.
- Sharma, V. K., Chih, H.-W., Mrsny, R. J. and Daugherty, A. L. (2009) The Formulation and Delivery of Monoclonal Antibodies. in *Therapeutic Monoclonal Antibodies*: John Wiley & Sons, Inc. pp. 673-709.
- Shukla, A. A., Gupta, P. and Han, X. (2007a) Protein aggregation kinetics during Protein A chromatography Case study for an Fc fusion protein. *Journal of Chromatography A*, 1171(1-2), pp. 22-28.
- Shukla, A. A. and Hinckley, P. (2008) Host cell protein clearance during protein A chromatography: development of an improved column wash step. *Biotechnol Prog*, 24(5), pp. 1115-21.
- Shukla, A. A., Hubbard, B., Tressel, T., Guhan, S. and Low, D. (2007b) Downstream processing of monoclonal antibodies--application of platform approaches. *J Chromatogr B Analyt Technol Biomed Life Sci*, 848(1), pp. 28-39.

- Sjodahl, J. (1977) Structural studies on the four repetitive Fc-binding regions in protein A from *Staphylococcus aureus*. *Eur J Biochem*, 78(2), pp. 471-90.
- Su, T. J., Lu, Thomas, R. K., Cui, Z. F. and Penfold, J. (1998) The Conformational Structure of Bovine Serum Albumin Layers Adsorbed at the Silica–Water Interface. *The Journal of Physical Chemistry B*, 102(41), pp. 8100-8108.
- Su, T. J., Lu, J. R., Thomas, R. K. and Cui, Z. F. (1999) Effect of pH on the Adsorption of Bovine Serum Albumin at the Silica/Water Interface Studied by Neutron Reflection. *The Journal of Physical Chemistry B*, 103(18), pp. 3727-3736.
- Swinnen, K., Krul, A., Van Goidsenhoven, I., Van Tichelt, N., Roosen, A. and Van Houdt, K. (2007) Performance comparison of protein A affinity resins for the purification of monoclonal antibodies. *Journal of Chromatography B*, 848(1), pp. 97-107.
- Taylor, J., Arnold, O., Bilheaux, J., Buts, A., Campbell, S., Doucet, M. and Draper, N. (2012) Mantid, A high performance framework for reduction and analysis of neutron scattering data. *Bulletin of the American Physical Society*, 57.
- Tetin, S. Y., Prendergast, F. G. and Venyaminov, S. Y. (2003) Accuracy of protein secondary structure determination from circular dichroism spectra based on immunoglobulin examples. *Analytical Biochemistry*, 321(2), pp. 183-187.
- Townsend, M. W. and Deluca, P. P. (1990) Stability of ribonuclease A in solution and the freeze-dried state. *J Pharm Sci*, 79(12), pp. 1083-1086.
- Ueberbacher, R., Rodler, A., Hahn, R. and Jungbauer, A. (2010) Hydrophobic interaction chromatography of proteins: thermodynamic analysis of conformational changes. *Journal of Chromatography A*, 1217(2), pp. 184-90.
- Vermeer, A. W. and Norde, W. (2000) The thermal stability of immunoglobulin: unfolding and aggregation of a multi-domain protein. *Biophysical Journal*, 78(1), pp. 394-404.
- Wales, T. E. and Engen, J. R. (2006) Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrometry Reviews*, 25(1), pp. 158-170.

- Wang, X. L., Das, T. K., Singh, S. K. and Kumar, S. (2009) Potential aggregation prone regions in biotherapeutics A survey of commercial monoclonal antibodies. *MABs*, 1(3), pp. 254-267.
- Wong, C., Strachan-Mills, C. and Burman, S. (2012) Facile method of quantification for oxidized tryptophan degradants of monoclonal antibody by mixed mode ultra performance liquid chromatography. *Journal of Chromatography A*, 1270, pp. 153-61.
- Xu, H., Lu, J. R. and Williams, D. E. (2006a) Effect of Surface Packing Density of Interfacially Adsorbed Monoclonal Antibody on the Binding of Hormonal Antigen Human Chorionic Gonadotrophin. *The Journal of Physical Chemistry B*, 110(4), pp. 1907-1914.
- Xu, H., Zhao, X., Grant, C., Lu, J. R., Williams, D. E. and Penfold, J. (2006b) Orientation of a Monoclonal Antibody Adsorbed at the Solid/Solution Interface: A Combined Study Using Atomic Force Microscopy and Neutron Reflectivity. *Langmuir*, 22(14), pp. 6313-6320.
- Yeats, C. A. and Orengo, C. A. (2001) Evolution of Protein Domains. in *eLS*: John Wiley & Sons, Ltd.
- Zamolo, L., Busini, V., Moiani, D., Moscatelli, D. and Cavallotti, C. (2008) Molecular Dynamic Investigation of the Interaction of Supported Affinity Ligands with Monoclonal Antibodies. *Biotechnol Prog*, 24(3), pp. 527-539.
- Zhao, X., Pan, F., Cowsill, B., Lu, J. R., Garcia-Gancedo, L., Flewitt, A. J., Ashley, G. M. and Luo, J. (2011) Interfacial Immobilization of Monoclonal Antibody and Detection of Human Prostate-Specific Antigen. *Langmuir*, 27(12), pp. 7654-7662.
- Zhao, X., Pan, F., Garcia-Gancedo, L., Flewitt, A. J., Ashley, G. M., Luo, J. and Lu, J. R. (2012) Interfacial recognition of human prostate-specific antigen by immobilized monoclonal antibody: effects of solution conditions and surface chemistry. *Journal of The Royal Society Interface*, 9(75), pp. 2457-2467.
- Zhao, X., Pan, F. and Lu, J. R. (2009) Interfacial assembly of proteins and peptides: recent examples studied by neutron reflection. *J R Soc Interface*, 6 Suppl 5, pp. S659-70.

11 Appendix

11.1 Publication

Mazzer, A.R.; Perraud, X.; Halley, J.; O'Hara, J. and Bracewell, D.G. Protein A chromatography increases monoclonal antibody aggregation rate during subsequent low pH virus inactivation hold. *Journal of Chromatography A* (2015). doi:10.1016/j.chroma.2015.08.068. <http://creativecommons.org/licenses/by/4.0/>.

11.2 Conferences

American Chemical Society, March 2015, CO, USA, BIOT Division poster presentation: Neutron reflectivity for characterisation of chromatographic interfaces in bioprocessing.

Neutron Characterisation in Fundamental and Applied Biotechnology, October 2014, Abingdon, Oxfordshire, UK, poster presentation: Neutron reflectivity for characterisation of chromatographic interfaces.

BioProcessUK, December 2013, London, UK, poster presentation: Understanding on-column protein aggregation and its impact on bioprocessing.