UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES

School of Biological Sciences

Studies on the interaction of a single domain of protein G from Streptococcus with human Fc and Fab

By

Nicki Muir

Thesis for the degree of Doctor of Philosophy

February 2009

DECLARATION OF AUTHORSHIP

I, Nicola Mary Muir

declare that this thesis entitled

Studies on the interaction of a single domain of protein G from Streptococcus with human Fc and Fab

and the work presented in it are my own. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have consulted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I contributed myself;

Signed:.....

Date:.....

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES

SCHOOL OF BIOLOGICAL SCIENCES

Doctor of Philosophy

STUDIES ON THE INTERACTION OF A SINGLE DOMAIN OF PROTEIN G FROM STREPTOCOCCUS WITH HUMAN FC AND FAB

By Nicola Mary Muir

Protein G (SpG) is a Type III bacterial Fc receptor that binds selectively and non-antigenically to both the Fc and Fab regions of IgG. Here studies are described aimed at understanding the individual contributions to complex stability made by amino acid residues known to make contacts with either the Fab or Fc. Specific amino-acid residues that contribute to these binding interactions via a network of hydrogen bonds have been identified and studied by a programme of site directed mutagenesis. Unique tryptophan residues (W42 or W14) have been used to facilitate equilibrium and pre-equilibrium fluorescence studies to observe binding to Fc or Fab and to determine the Kd for the various equilibria. The effect of secondary substitutions of residues implicated in bond formation to either the Fc or Fab IgG fragment have been determined in order to generate SpG domains with a range of binding affinities for biotechnological uses. SpG species that only bind to Fc or Fab have also been developed. The mutant W42F has been shown by ITC studies to have lost the ability to bind to Fc, and the double mutant T15A-T16A shows no detectable binding to Fab. ITC and kinetic studies at various temperatures have been employed to characterise the thermodynamic parameters underlying binding of wild type or mutated SpG to Fc or Fab. Some restoration of Fc binding activity has been brought about by the placement of a tryptophan at residue 14 (E14W), which may be a useful protein for affinity chromatographic applications as a Kd in the micromolar range is ideal for affinity ligands where easy elution of the bound target protein is required. A dramatic reduction in the affinity of SpG for Fc has been obtained for the E26A mutation, and the N34 residue has been found to be significant in Fc binding, with a 50-fold decrease in affinity for the N34A mutation. The electrostatic nature of the interactions and their sensitivity to changes in pH has also been studied and results showed that Q31E has increased Kd by approximately 6-fold, demonstrating that a negative charge at this position is detrimental to binding. The lower affinity and the influence from the ionisable group would offer advantages for the elution of the Ig from the column.

In loving memory of Bertie

ACKNOWLEDGEMENTS

I would like to thank my supervisor Professor M.G. Gore for all of his help and guidance during the last eight years, as well as the University of Southampton for the funding of this research. I would not have been able to complete this thesis without his continued support and patience, for which I am very grateful. I would also like to thank all past members of the research group for all their help and support.

I am also very grateful to Dr A. Popplewell from UCB-Celltech for the kind gift of the pTTOD vector containing the Fab gene, and Professor M. Glennie at Southampton General Hospital for the kind donation of the human Fc fragments, both of which were essential for this work.

I would like to thank all my friends - you know who you are - for not giving up on me through the stressful times and being there when I needed you.

Finally, I would like to say a very big and heartfelt thank you to my family, who have not only supported me, but who believed in me and gave me the encouragement I needed to persevere, as well as having to endure my bad moods. I owe you all so much and I could not have completed this thesis without you - thank you.

ABBREVIATIONS

λ	Wavelength (nm)
λ_{EX}	Excitation wavelength
λ_{EM}	Emission wavelength
λ-chain	Lambda light chain of immunoglobulin
κ-chain	Kappa light chain of immunoglobulin
α	Alpha
Φ_F	Fluorescence quantum yield
ε_X	Molar extinction coefficient at wavelength X
ΔG	Gibbs free energy
ΔG_U	Difference in free energy between folded and unfolded states
ΔH	Enthalpy change
ΔS	Entropy change
ABTS	2, 2'-Azino-di-[3-ethylbenzthiazoline]
Ala	Alanine amino acid
APS	Ammonium persulphate
Arg	Arginine amino acid
Asn	Asparagine amino acid
Asp	Asparatate amino acid
ATP	Adenosine triphosphate
A_X	Absorbance at wavelength X nm
β	Beta
BCA	Bicinchroninic acid
BSA	Bovine serum albumin
BSE	Bovine spongiform encephalopathy
CD	Circular dichroism
CDR	Complementarity-determining region
C_H	Constant domain of immunoglobulin heavy chain
C_L	Constant domain of immunoglobulin light chain
CJD	Creutzfeldt-Jacob disease
D	Denatured state
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
DNase I	Deoxyribonuclease I
dNTP	Deoxy-nucleotide triphosphate
DTT	Dithiothreitol
Ea	Activation energy
EDTA	Ethylenediamineteraacetic acid
EGTA	Ethylene glycol-bis(β-aminoethyl ether)
ELISA	Enzyme-linked immunosorbant assay
E.coli	*Escherichia coli*
Fab	Antigen-binding fragment of immunoglobulin
Fc	Crystallising fragment of immunoglobulin
F_{corr}	Fluorescence corrected
F_F	Fraction of folded protein
F_{obs}	Fluorescence observed
FR	Framework region
F_U	Fraction of unfolded protein
Gln	Glutamine amino acid
Glu	Glutamate amino acid
Gdn-HCl	Guanidine hydrochloride
HCl	Hydrochloride
His	Histidine amino acid
HRP	Horse radish peroxidase
HSA	Human serum albumin
Ig	Immunoglobulin

IgG	Immunoglobulin G
IgX	Immunoglobulin of heavy chain subclass X
IPTG	Isopropyl-β-D-thiogalactosidase
ITC	Isothermal titration calorimetry
I_R	Right circularly polarized light
I_L	Left circularly polarized light
K_2HPO_4	Di-potassium hydrogen phosphate
KH_2PO4	Potassium di-hydrogen phosphate
KPO_4	Potassium phosphate
Ka	Affinity constant
k_{app}	Observed rate constant
k_{ass}	Association rate constant
Kd	Dissociation constant
k_{diss}	Dissociation rate constant
K_{eq}	Equilibrium constant
kJ	Kilo joules
LB	Luria-broth
Leu	Leucine amino acid
Lys	Lysine amino acid
MCS	Multiple cloning site
M	Molar
mM	Millimolar
mRNA	Messenger ribonucleic acid
ms	Milliseconds
NaCl	Sodium chloride
Na_2EDTA	Di-sodium Ethylenediamineteraacetic acid
NHS	N-hydroxysuccinimide
nm	Nano meters
NMR	Nuclear magnetic resonance
NTP	Nucleotide triphosphate
OD_{Xnm}	Optical density at wavelength X nm
ompA	Outer membrane of protein A
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
Pfu	*Pyrococcus furiosus*
pH	Potential of hydrogen
Phe	Phenylalanine amino acid
p*Ka*	Measure of proton binding affinity
p*K*	State of ionisation or dissociation
PMSF	Phenylmethane sulphonyl fluoride
PpL_{3316}	Single Ig-binding domain of protein L from *Peptostreptococcus magnus* strain 3316
PpL_{312}	Single Ig-binding domain of protein L from *Peptostreptococcus magnus* strain 312
RNA	Ribonucleic acid
RNase	Ribonuclease
ScFv	Single-chain fragment variable
sdH_2O	Sterile distilled water
SDS	Sodium dodecyl sulphate
SpG	Protein A from *Staphylococcus aureus*
S_0	Ground state orbital
S_1	First excited singlet state
SOE-ing	Splicing by overlap extension
SpG	Protein G from Groups C and G streptococci
T_1	Triplet excited state

TBE	Tris borate EDTA
TE	Tris EDTA
TEMED	N, N, N', N'-tetraethylmethylene diamine
Thr	Threonine amino acid
Tris	Tris (hydroxymethyl) aminomethane
Trp	Tryptophan amino acid
Tween™	Polyoxyethylenesorbitan monolaurate
U	Unfolded state
UV	Ultra-violet
V_H	Variable domain of immunoglobulin heavy chain
V_L	Variable domain of immunoglobulin light chain
v/v	Volume per volume
Wt	Wild-type
w/v	Weight per volume
y	Observed fluorescence at chosen denaturant concentration
y_F	Fluorescence characteristics of folded protein
y_U	Fluorescence characteristics of unfolded protein

CONTENTS PAGE

List of Figures

Chapter 1.0 - Introduction

Chapter 2.0 - Materials & methods

Chapter 3.0 - The design, cloning & expression of Fc & Fab constructs

Chapter 4.0 - Stability of Fc & Fab constructs

Chapter 5.0 - Equilibrium & pre-equilibrium binding studies of SpG-Fc binding interaction

Chapter 6.0 - Equilibrium & pre-equilibrium binding studies of SpG-Fab binding interaction

Chapter 7.0 - Equilibrium and pre-equilibrium binding studies of SpG-Fab binding interaction

List of Tables

Chapter 7 - Equilibrium and pre-equilibrium binding studies of SpG-Fab binding interaction

CHAPTER 1.0

INTRODUCTION

1.1 - Immunoglobulins

The immune system is a network of cells and organs that has evolved to defend the body against attacks by foreign invaders. Every cell in the body carries distinctive molecules that distinguish it as 'self' and under normal conditions the body's defences do not attack tissues that carry these self markers. Thus, the immune cells can coexist with other body cells in a state known as self-tolerance. Cells, destined to become immune cells, arise in the bone marrow from stem cells, some developing from lymphoid precursors, into small white blood cells called B- and T-lymphocytes. Foreign particles, such as bacteria or viruses, usually carry distinctive epitopes that protrude from their surfaces. These non-self antigens are recognised by the body's immune system triggering a humoral immune response, which responds by producing molecules known as immunoglobulins.

The term "immunoglobulin" encompasses all globulin proteins associated with the immune system, including all antibody molecules. Antibodies are soluble proteins that are expressed on the cell surface of all B-lymphocytes where they function as the receptors for the antigen, and as the secreted products of the plasma cell. Each B-lymphocyte is programmed to make one specific antibody, and when it encounters its triggering antigen, along with various accessory cells, it gives rise to a clone of many large plasma cells. Each plasma cell is essentially a factory for producing that one specific antibody.

The secreted antibody molecules circulate in the blood and lymph, and bind specifically to epitopes from the pathogen that elicited a humoral immune response, as illustrated in figure 1.1.

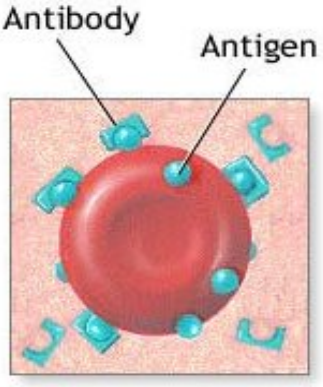


Figure 1.1 - An illustration of the interaction between antibody and antigen

Antibodies are produced by the immune system in response to the presence of an antigen. They may be free-floating in the serum and combine with antigen to form immune complexes, or maybe attached to cell-surface antigens thus acting as receptors.

This induces the B-lymphocytes to proliferate and differentiate into antibody-secreting cells, which can then recruit other cells and molecules to destroy the pathogen. These functions are structurally separated in the antibody molecule, one part, the Fab, recognises antigen, while the other part, the Fc, engages the effector mechanisms that will eliminate it [Goward *et al.*, 1993].

1.1.1 - The structure of an immunoglobulin molecule

Immunoglobulins (Igs) are composed of four polypeptide chains of varying length that are interconnected by covalent cysteine disulphide linkages (-S-S-) and non-covalent bonds to form a Y-shaped molecule. There are two types of polypeptide chains: two light chains which consist of approximately 220 amino acid residues and two heavy chains of about 440 amino acids [Porter, 1973]. They are synthesised separately, but connected by -S-S- bonds. Sequence studies of various immunoglobulins have shown that both the heavy and light polypeptide chains contain repeated homologous sequences that are approximately 110 residues in length [Schiffer *et al.*, 1973]. These homologous regions fold individually into similar structural domains, arranged as a bilayer of antiparallel β-sheets, and each domain contains an intra-chain disulphide bond. Figures 1.2 and 1.3 represent adapted illustrations of immunoglobulin molecules.

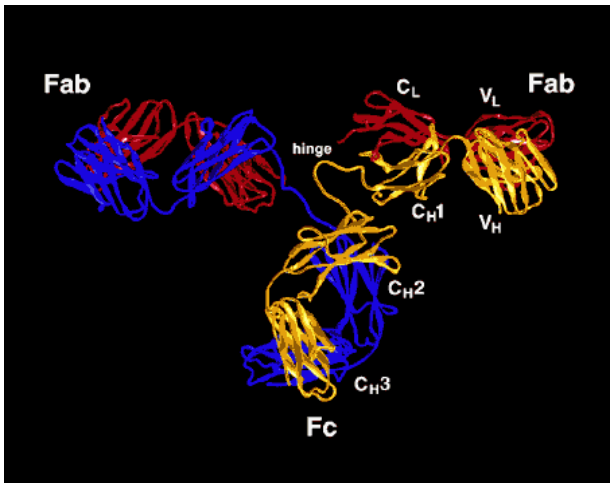


Figure 1.2 – <u>A ribbon representation of mouse IgG_2</u>

This three-dimensional ribbon representation of a mouse monoclonal antibody of subclass IgG_2 illustrates the secondary structure of the alpha-carbon backbone traces. The beta strands are indicated as ribbons. The structure shows that most of the domains have one face in intimate contact with a second domain, thus the heavy and light V-region domains are in contact, as are the light chain C-region domain and heavy chain C_H1 domain, and the two heavy chain C_H3 domains [Harris et al., 1992].

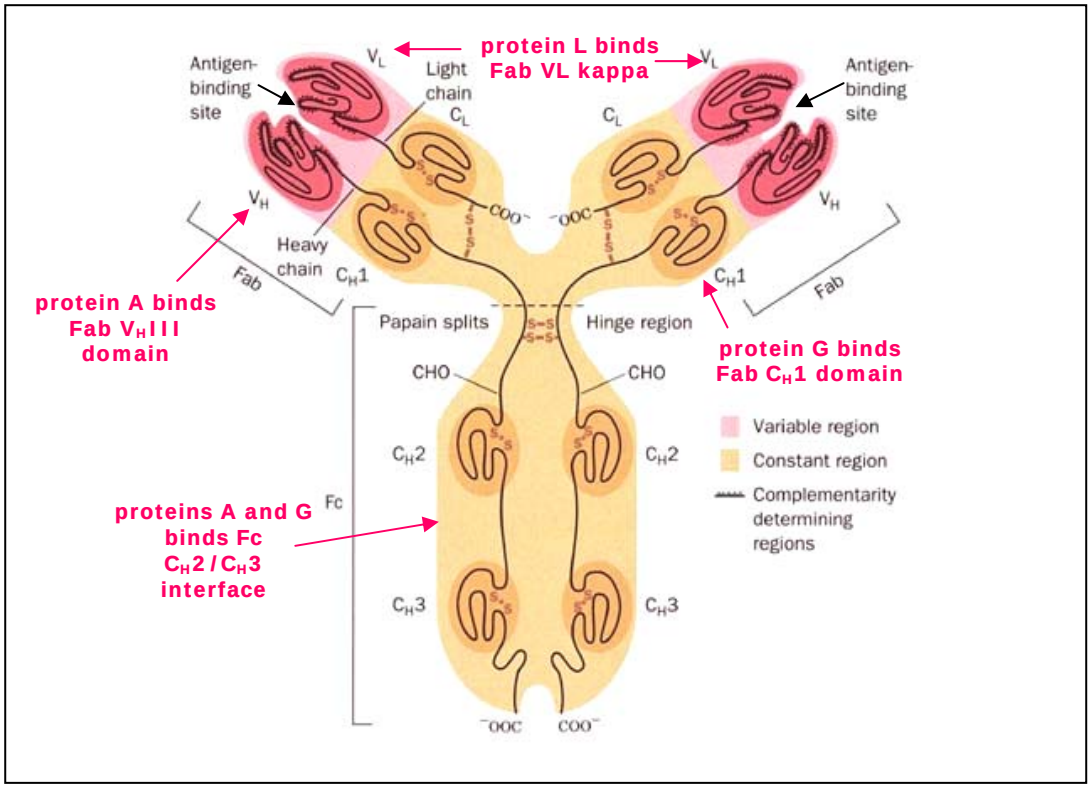


Figure 1.3 - <u>An adapted illustration of an immunoglobulin molecule</u>

An immunoglobulin molecule is presented together with the pattern of disulphide cross-links, which joins the heavy and light chains. The subscripts $_L$ and $_H$ refer to light and heavy chains respectively. C and V refer to regions of the sequence that are relatively constant or variable respectively. Digestion of the proline rich hinge region with proteolytic enzymes splits the antibody molecule into two Fab fragments. Each fragment contains a single κ or λ light chain linked by a disulphide bond to the NH_2-terminal half of the heavy chain; and one Fc fragment, composed of the two terminal halves of the heavy chain combined by a disulphide bond. It is only the Fab fragments that bind antigen, indicating that the antibody-binding site is entirely in the Fab fragment. The putative binding sites for several Ig-binding proteins are indicated in red [Tonegawa, 1985].

The Ig molecule in its entirety is formed of twelve structural domains, of which eight are formed by two heavy chains and the remaining four by the light chains [Tonegawa, 1985]. Immunoglobulin heavy and light chains each contain an amino-terminal variable (V) region, which differs from one Ig to the next depending upon their level of sequence variation. The carboxyl-terminal half of the molecule contains a small number of constant (C) regions, which mediate the biological effector functions of the immune system, and thus determines the mechanism used to destroy antigen, and does not vary in the same way [Lucas, 2001]. Antibodies are divided into five major classes, IgA, IgD, IgE, IgG and IgM, based on their constant region structure and immune function.

1.1.1.1 – Light and Heavy chains

Light chains contain one variable domain (V_L) and one constant domain (C_L), and variations in the sequence of the C_L domain have meant that the light chains can be of two classes, kappa (κ) and lambda (λ). A single immunoglobulin molecule can contain either κ light chains or λ light chains, but never both. In humans 60% of the light chains are kappa, and 40% are lambda, whereas comparatively, in mice 95% of the light chains are kappa, and only 5% are lambda [Porter, 1973].

Heavy chains contain one variable domain (V_H), and three (or four in the case of IgE and IgM) constant domains (C_H1, C_H2, C_H3 and C_H4). There are five classes of heavy chains, which give rise to the five classes of antibodies, IgA, IgD, IgE, IgG and IgM. Figure 1.4 represents the general structure of the five classes of secreted antibody.

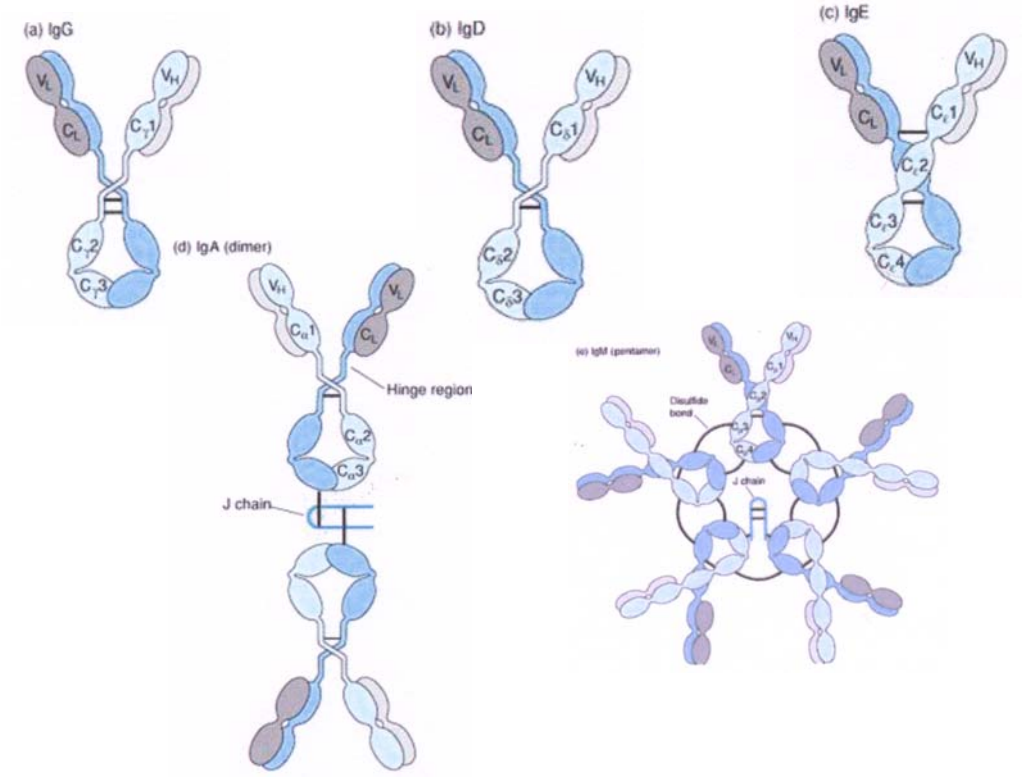


Figure 1.4 – <u>An illustration of the general structure of the five major immunoglobulin classes</u>

There are five major classes of secreted antibody, IgA, IgD, IgE, IgG and IgM. Light chains are shown in shades of grey and heavy chains in shades of blue; thick black lines indicate disulphide bonds [taken from Alzari et al., 1988].

IgA, IgD and IgG all contain only three C_H domains, whereas IgE and IgM contain all four. Minor differences in the amino acid sequences of IgA and IgG have led to further classification of the heavy chains into subclasses. In humans there are two subclasses of IgA, IgA_1 and IgA_2, and IgG can be further divided into four subclasses IgG_1, IgG_2, IgG_3 and IgG_4 [Porter, 1973]. In mice there are four subclasses of IgG heavy chains IgG_1, IgG_{2a}, IgG_{2b} and IgG_3.

1.1.1.2 – Hinge region and Fab fragments

The IgA, IgD and IgG heavy chains also contain an extended peptide sequence between the C_H1 and C_H2 domains that has no homology with the other domains. This region is known as the hinge region, and is rich in proline residues and flexible, giving IgA, IgD and IgG segmental flexibility. IgE and IgM chains lack a hinge region, but have an additional 110 amino acid domain (C_H2/C_H2) instead, that has hinge-like features [Tonegawa, 1985]. Early digest studies performed on rabbit IgG [Porter, 1959], demonstrated that papain hydrolyses at this hinge region, the midpoint of the heavy chain. Digesting the immunoglobulin molecule before the heavy inter-chain disulfide bond produces two identical Fab (*f*ragment *a*ntigen-*b*inding) fragments, which contain the light chain and the V_H and C_H1 domains of the heavy chain. The molecular size of an antibody is too large to study directly so the production of these fragments, which could be studied individually, meant that the clarification of the structure of immunoglobulins was revealed. The fragments produced by antibody digestion are summarised schematically in figure 1.3.

Each Fab fragment is composed of one constant and one variable domain of the light chain linked by a disulphide bond to the heavy chain, which continued into the Fc fragment before it was cleaved, as shown in figure 1.3. Together this composition shapes the antigen-binding site at the amino-terminal end of the monomer, and the two variable domains bind their specific antigens. Also, as a result of the hinge region flexibility, the two Fab arms can assume various angles relative to each other when an antigen is bound.

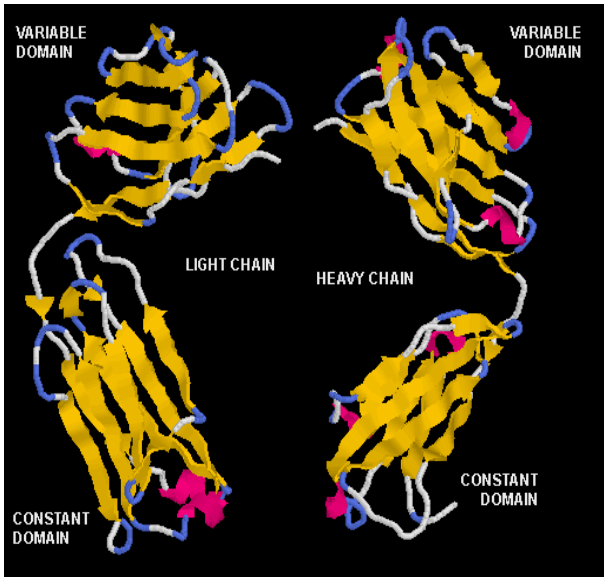


Figure 1.5 – A ribbon representation of two Fab fragments

Each Fab fragment contains two chains, a light and a heavy chain. The chains are not separated as depicted above. Each chain consists of two beta-sheet sandwiches, one the variable domain (V) and one the constant domain (C). Corresponding domains (V_H+V_L and C_H+C_L) are held together by hydrogen bonds and non-hydrophobic interactions.

1.1.1.3 – Antigen-binding sites

Various domains may associate to form discrete structural regions and the first 110 amino acids at the N-terminal end of each heavy and light chain provides specificity for binding an epitope (the portion of an antigen bound by an antibody). This amino acid sequence of both the light and heavy chains shows great variation from antibody to antibody and constitutes the variable domains (containing hypervariable regions or complementarity determining regions CDRs) of the antibody [Padlan, 1994]. Each B-lymphocyte during early development becomes genetically programmed through a series of gene-splicing reactions to produce unique three-dimensional shapes capable of fitting an epitope with a complementary shape. The various genes that the cell splices together determine the order of the amino acids and this sequence determines the final three-dimensional shape. Therefore, different antibody molecules produced by different B-lymphocytes will have different amino acid sequences to give them their unique shapes. These unique three-dimensional shapes are the antigen-binding sites, formed by folded V_H and V_L regions and can accommodate approximately four to seven amino acids [Mariuzza and Poljak, 1993].

The antigen-binding region is a unique feature of these molecules and distinguishes them from all other known proteins [Davies *et al.*, 1990]. Each

immunoglobulin molecule has at least two antigen-binding sites so that it can either bind two molecules of antigen, or if the antigen is a cell with multiple copies of the antigen on its surface, it can form multiple attachments to a single cell surface.

The variable region of an antibody molecule is further subdivided into hypervariable (HV) regions. Three segments of high variability can be identified in both the V_H and V_L domains and are denoted HV1, HV2 and HV3. The hypervariable regions, also known as CDRs, are where most of the variations occur and have a high ratio of different amino acids in a given position, relative to the most common amino acid in that position [Essen and Skerra, 1994]. The most variable part of the domain is the HV3 region. The regions between the hypervariable regions, which comprise the rest of the variable domain, show less variability and have more stable amino acid sequences. These are known as the framework regions and there are four such regions in each variable domain, designated FR1, FR2, FR3 and FR4. The framework regions form the β-sheets that provide the structural framework of the domain, whereas the hypervariable sequences correspond to three loops located at the outer edge of the β barrel, which are juxtaposed in the folded domain. This is shown in figure 1.6.

Folding brings the hypervariable loops from each domain together to create a single hypervariable antigen-binding site at the tip of each arm of the molecule. The three hypervariable loops, or CDRs, determine antigen specificity by forming a surface, which is complementary to the antigen [Ward *et al*., 1989]. The FR regions form a beta-sheet structure creating a pocket of three-dimensional internal form, into which a part of an antigen molecule can fit closely and bind tightly, with little cross-reaction from different antigens [Foote and Winter, 1992]. This also serves as a scaffold to hold the hypervariable regions in position to contact the antigen.

As already mentioned, the amino acid sequences of the CDRs are different in different antibodies, so therefore the shapes of the surfaces created by these CDRs are also different. Amino acid side chains in most or all of the hypervariable loops make contact with antigen and determine both the specificity and affinity of the interaction. Other parts of the variable region play little part in the direct contact with the antigen but provide a stable structural framework for the hypervariable loops and help determine their position and conformation.

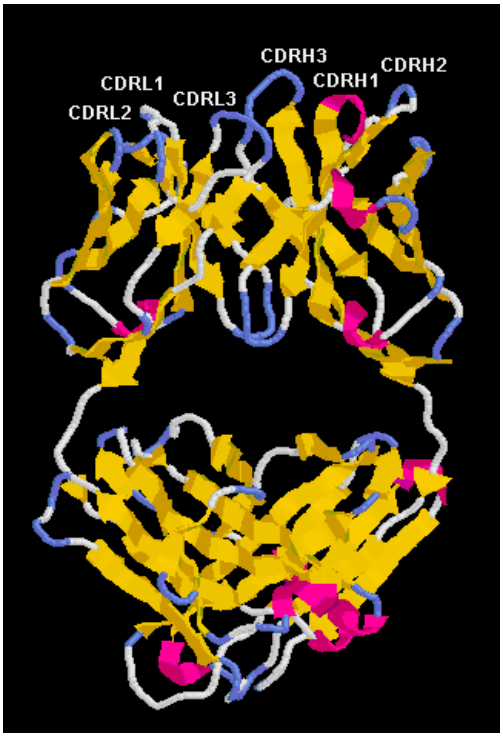


Figure 1.6 – A ribbon representation of the hypervariable regions

The hypervariable regions lie in discrete loops of the folded structure. When the hypervariable regions or CDRs (also known as hypervariable loops), are positioned on the structure of a variable domain it can be seen in the above diagram that they lie in loops that are brought together in the folded structure. In the antibody molecule, the pairing of a heavy and a light chain brings together the hypervariable loops from each chain to create a single hypervariable surface, which forms the antigen-binding site at the tip of each arm.

The strength of the interaction between a single antigen-binding site and its specific antigen epitope is known as the binding affinity. The antigen-binding specificity of an antibody resembles that of an enzyme binding to a substrate, and the higher the affinity, the tighter the association between the antigen and the antibody. Antibody affinity generally increases with repeated exposure to antigen, a process called affinity maturation, as B-lymphocytes with higher affinity antigen receptors are selected to produce larger clones of antibody-secreting plasma cells [Borrebaeck *et al*., 1992], particularly in conditions where the concentration of epitope is low.

Antigen-antibody interactions involve a variety of noncovalent forces that hold together the complex and contribute to binding. Electrostatic interactions occur between charged amino acid side chains, as in salt bridges, and interactions also occur between electric dipoles as in hydrogen bonds, or can involve short-range van der Waals forces and hydrophobic interactions. High salt concentrations, extremes of pH and detergents can disrupt this reversible interaction between antigen and antibody by weakening electrostatic interactions and/or hydrogen

bonds.

1.1.1.4 - Fc fragments

The other product of a papain digestion of an Ig is a fragment that contains the remainder of the two heavy chains, each containing a C_H2 and C_H3 domain. This fragment, shown in figure 1.7, was called Fc (*f*ragment *c*rystallising) because it was easily crystallized. It is composed of the carboxyl terminal portions of both heavy chains linked by a disulphide bond. The Fc portion is responsible for the biological activity of the antibody [Tonegawa, 1985].

Fc binds to various cell receptors and complement proteins, and evidence suggests that it is the C_H2 domains that are of primary importance in determining Fc effector functions, such as complement activation (C1q binding) and Fc Receptor (Fc gamma R) binding [Goulding *et al*., 1992]. In this way, Fc mediates different physiological effects of antibodies, such as: opsonisation; cell lysis; mast cell, basophil and eosinophil degranulation and other processes. However, other parts of the molecule may also be critical, in particular the hinge region and hinge link, or lower hinge region, where the genetically determined hinge joins on to a part of the hinge encoded genetically as the N-terminal end of the C_H2 domain.

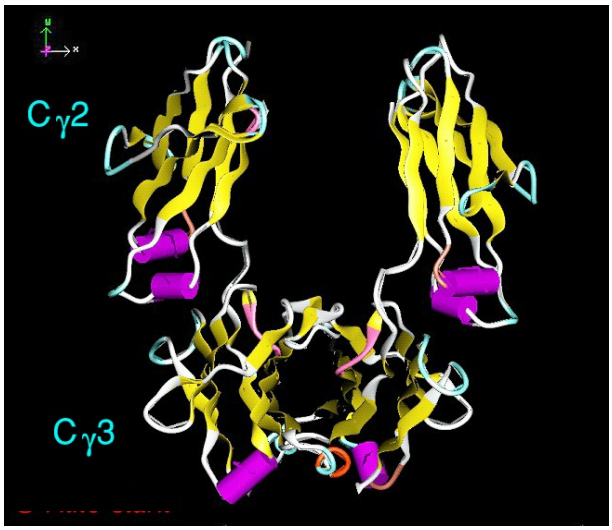


Figure 1.7 – A ribbon representation of the Fc fragment

The bottom part of the "Y", the C-terminal region of each glycoprotein chain is known as the Fc fragment. It is composed of two heavy chains that each contributes two constant domains, C_H2 and C_H3.

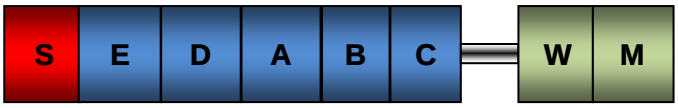
1.2 - Immunoglobulin-binding proteins

Several species of infectious bacteria express immunoglobulin-binding proteins on their cell surfaces. When a bacterium infects a mammal it encounters a hostile environment, where it is attacked by the host's immune system. The cell surfaces of some bacteria are covered with a number of proteins that may be involved in the mechanism of pathogenesis and they have to be robust to withstand the host's hostile environment. There are a number of these cell surface proteins of pathogenic bacteria that are able to bind to a variety of host plasma proteins, such as the immunoglobulins and clotting factors, which accumulate after trauma. The ability to bind plasma proteins is thought to help the bacterium to colonise a site, including wounds [Goward *et al*., 1993]. The function of these cell surface proteins *in vivo* is still unclear, but it is thought that they coat themselves with host proteins, so that the bacterium is no longer recognised by the host's defence mechanism as foreign and thus, enable it to avoid destruction. This, in turn, increases their virulence [Frick *et al*., 1992].

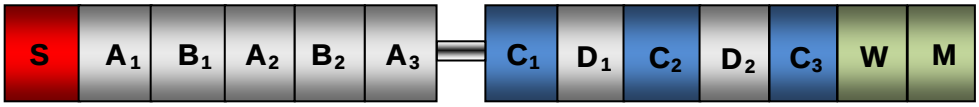
The immunoglobulin-binding capabilities of some cell-surface proteins are interesting in that they are related to the site of infection. For example, the IgG-binding proteins are found on bacteria present in the plasma, where IgG is the dominant antibody; the IgA-binding proteins are found on bacteria in mucosal secretion, where IgA is the major antibody; and IgD-binding proteins are found in the upper respiratory tract, which is a common site of infection for *Haemophilus influenza* and rich in IgD [Goward *et al*., 1993]. Although there is a lack of information to identify the role of these proteins in infectivity, their ability to interact with immunoglobulins *in vitro* without affecting their antigen-binding properties, makes them invaluable for a wide range of immunological techniques.

Several cell surface immunoglobulin-binding proteins from microbial cell walls have been identified. Protein A from *Staphylococcus aureus* (*S.aureus*) [Verway, 1940; Oeding *et al*., 1964; Boyle and Reis, 1987], protein G from Streptococci groups C and D [Kronvall, 1973], protein H from Streptococci group A [Akesson *et al*., 1990], protein D from *Haemophilus influenza* [Forgsren and Grubb, 1979] and protein L from *Peptostreptococcus magnus* [Myhre and Erntell, 1985]. The relative binding sites on IgG for proteins A, G and L are identified in figure 1.3, and their primary sequence organisation is shown in figure 1.8.

Protein A



Protein G



Protein L

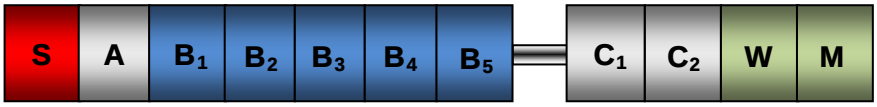


Figure 1.8 - Primary sequence organisation of Ig-binding proteins A, G, and L

All cell surface Ig-binding proteins share similar primary sequence organisation. The secretion signal peptide sequences (red area S) and regions that interact with the cell wall (green area W) and cell membrane (green area M) are shown. The blue areas indicate the regions that are involved in binding to immunoglobulins, the number of which is dependent on the strain of bacteria. The other pale grey areas denote various repetitive regions of the proteins which demonstrate further binding functions [Goward et al., 1993].

Protein	Source	Ligand	Comments
Protein A	*Staphylococcus aureus*	IgG Fab (V_HIII) + Fc (C_H2/C_H3 interface); IgA Fab; IgE Fab; IgM Fab; α_2-macroglobulin and Kininogen	Binds all Igs except IgG_3 and possesses a three α-helix structure
Protein G	*Streptococci* Group C and G	IgG Fab (C_H1) + Fc (C_H2/C_H3 interface); Human serum albumin; α_2-macroglobulin and Kininogen	Only binds IgG, with a much broader range and possesses a four β-strand and one α-helix structure
Protein L	*Peptostreptococcus magnus*	IgG Fab V_L kappa chains; HSA; α_2-macroglobulin and Kininogen	Binds to kappa light chains of Igs and possesses a four β-strand and one α-helix structure
Protein D	*Haemophilus influenza*	IgD	Only Gram-negative Ig-binding protein
Protein H	*Streptococcus pyogenes* Group A	IgG Fc (C_H2/C_H3); HSA and α_2-macroglobulin	Binds only to human and rabbit IgG

Table 1.1 - Comparison of bacterial cell surface Ig-binding proteins

1.2.1 - Protein A (SpA)

Staphylococcal protein A (SpA) is a Type 1 bacterial Fc receptor found on the surface of, or secreted by the cell wall of, the bacterium *S. aureus*. SpA is the most extensively characterised of the immunoglobulin-binding proteins and was first isolated in 1940 by Verway [Verway, 1940], and was later identified in 1964 by Oeding [Oeding *et al*., 1964].

1.2.1.1 - Structure of SpA

The gene for SpA has been cloned and over-expressed in *Escherichia coli* (*E.coli)* [Duggleby and Jones, 1983] and [Lofdahl *et al.*, 1983], and the DNA sequence has been determined [Uhlén *et al*., 1984a and 1984b]. Sequence studies have shown that the SpA molecule consists of six regions. The extracellular part of SpA consists of a series of five highly homologous regions [Sjödahl, 1977a and 1977b], each able to bind specifically to the Fc region of certain classes of immunoglobulin. In addition to the mutual homology between the regions, there is also significant internal homology within the regions [Sjödahl, 1977b]. The five functional monovalent Fc-binding domains are composed of approximately 60 amino acids each consecutively arranged in the order E, D, A, B and C, from the N-terminus of the protein [Moks *et al*., 1986].

The C-terminal part of SpA, the sixth region designated as domain X, is a cell wall binding domain comprising of 180 amino acid residues, which does not bind to the Fc portion. Its structure X shows a repetitive region, X_r, followed by a constant region, X_c [Uhlén *et al*., 1984]. The X_r region consists of an octapeptide sequence repeated twelve times, and it is thought to be involved in the binding to peptidoglycan of the cell wall. Region X_c represents a non-repeated carboxyterminal domain containing an extended stretch of hydrophobic amino acids, implying a cell membrane association [Uhlén *et al*., 1984]. A comparison of the Ig-binding domains of SpA is illustrated in figure 1.8.

1.2.1.2 - Structure of a single Ig-binding domain from SpA

The three-dimensional structure of Ig-binding domains of SpA (SpA_B) has been extensively studied using NMR and X-ray crystallography, and determined as free proteins or Ig-bound complexes. These structural analyses revealed that all of these SpA Ig-binding domains adopt a classical "anti-parallel" three helical bundle fold, as shown in figure 1.9. Each domain is composed of three nearly perfectly anti-parallel α-helices in orientation, with the first helix tilting slightly away from the other two helices. Helices II and III are anti-parallel to one other with helix I

tilted at a ~30° angle with respect to the other two helices [Gouda *et al*., 1992].

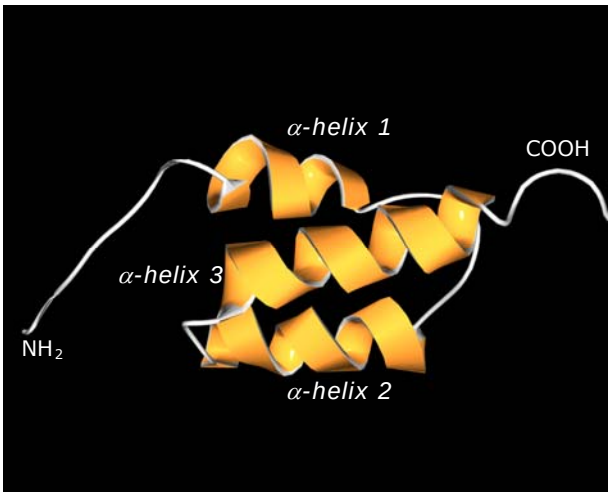


Figure 1.9 – A ribbon representation of the B domain of SpA

The three-dimensional structure of a single IgG-binding domain fragment B of SpA is composed of three anti-parallel α-helices. The figure illustrates the B domain viewed perpendicular to the α-helices.

1.2.1.3 - Interaction with Fc

Interest in SpA is based on its unique reactivity to immunoglobulins (Igs), defined by Forsgren and Sjöquist [Forsgren and Sjöquist, 1966]. They concluded that SpA's reactivity is directed to the Fc region of IgG. Further studies have shown that binding occurs through the Fc region to SpA on the surface of bacteria [Forsgren and Forsum, 1970], thus eliciting the ability to behave as an Fc receptor. The affinity and stoichiometry of the interaction between SpA and the Fc region has also been studied and results have shown that the binding site for SpA is localised at the C_H2-C_H3 domain interface [Lancet *et al*., 1978]. However, it has been demonstrated that one molecule of SpA is only able to bind two IgG molecules [Sjöquist *et al*., 1972].

SpA interacts with the Fc component of immunoglobulin from most mammalian species, including man. The class of immunoglobulin bound is predominantly IgG, but interactions are restricted to certain subgroups of IgG, namely IgG_1, IgG_2 and IgG_4. The inability of human IgG_3 to form a complex with SpA has been explained by the presence of the positive charge of arginine 435 [Deisenhofer, 1981]. The substitution of arginine 435 in IgG_3 from histidine 435 present in the other IgG subclasses in the Fc fragment, results in a steric clash between arginine 435 and other residues [Björck *et al*., 1984b]. SpA can also bind to human IgE and in some species with IgA and IgM as well [Lindmark *et al*., 1983]. Subclass specificity has also been suggested for human IgA, with SpA binding only to IgA_2

and not IgA_1 [Surolia *et al*., 1982].

The three-dimensional structure of a single IgG-binding domain fragment B (FB) bound to human Fc was initially determined by X-ray crystallographic analysis at 2.8Å resolution [Deisenhofer, 1981]. The crystal data indicated that two anti-parallel helical regions and a disordered C-terminal region exist in the Fc-bound FB. Furthermore, it has been demonstrated that the FB forms two contacts with the Fc fragment in the crystal [Deisenhofer, 1981]. Contact one is predominantly hydrophobic, and involves residues Gln10-Leu18 existing in the first helical region, and residues Glu26-Asp37 in the second helical region, together with those involved in the binding to IgG at the site located between C_H2 and C_H3. Contact two is smaller and involves residues Gln33, Ser34 and Asp37, which are located in the C-terminal segment of the second helical region, as well as Asp38 and Gln41 [Nilsson *et al*., 1987]. Figure 1.9 represents a three-dimensional ribbon diagram of the B domain of SpA and figure 1.10 represents a three-dimensional ribbon diagram of the B domain of SpA and the Fc fragment.

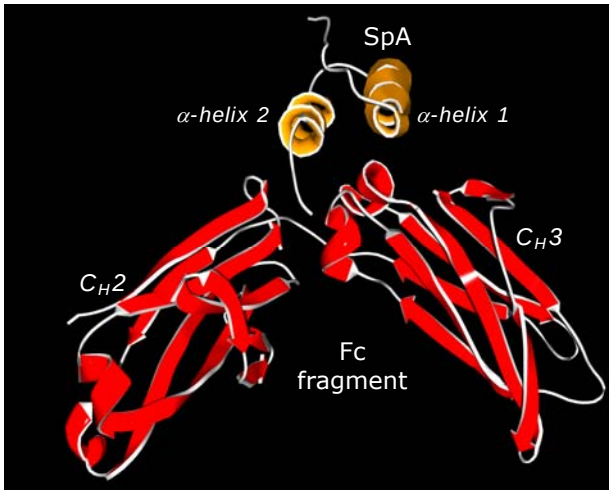


Figure 1.10 – <u>A ribbon representation of the B domain of SpA interacting with the C_H2/C_H3 interface of an IgG Fc fragment</u>

In this crystal structure complex the binding site for the B domain of SpA (gold) is localised at the C_H2/C_H3 domain interface of the Fc region of IgG (red). Only α-helices 1 and 2 were found by x-ray crystallographic analysis at 2.8Å resolution. Two contacts are formed with the Fc fragment in the crystal, with contact one involving residues from the first and second helical regions, together with residues involved in the binding to the IgG at the C_H2/C_H3 interface. Contact two is smaller and involves residues located in the C-terminal segment of the second helical region. PDB accession code: 1FC2.

However, NMR studies of the structure of SpA_B in solution identified a third α-helix corresponding to the unstructured region in the x-ray crystal structure. Thus, the ten residues that were thought to form the majority of the contact points with the Fc fragment, actually forms a third C-terminal α-helix [Torigoe, 1990a and 1990b]. This α-helix is packed against the other two forming an anti-

parallel three helical bundle, and NMR studies have shown the second and third α-helices to be in intimate contact regardless of whether the domain is in free solution or bound to IgG [Gouda *et al*., 1992]. More recent NMR analyses have concluded that the absence of helix III in the Fc complex is due to crystal contacts [Gouda *et al*., 1998]. A further NMR study into an engineered Ig-binding Z domain of SpA, supports all previous findings and also demonstrates that there is not a large conformational rearrangement of the three-helix Ig-binding domains upon Fc binding [Zheng *et al*., 2004].

1.2.1.4 - Interaction with Fab

In recent years an additional site for the interaction of SpA with the immunoglobulin molecule has been located in the Fab region, the Ig fragment responsible for antigen recognition, involving a site on the Ig heavy chain [Lindmark *et al*., 1983; Sasso *et al*., 1991]. This Fab-binding by SpA is mediated by the V domains and is associated with certain V_H genes. The V_H genes of family III encode all human immunoglobulins that are targets for Fab binding, but not all V_HIII encoded molecules are bound, thus binding is a marker of a group of human V_H genes, a subset of V_HIII [Ibrahim *et al*., 1993a]. Figure 1.11 represents a ribbon diagram of the interaction between the B domain of SpA and the Fab fragment of an IgG molecule.

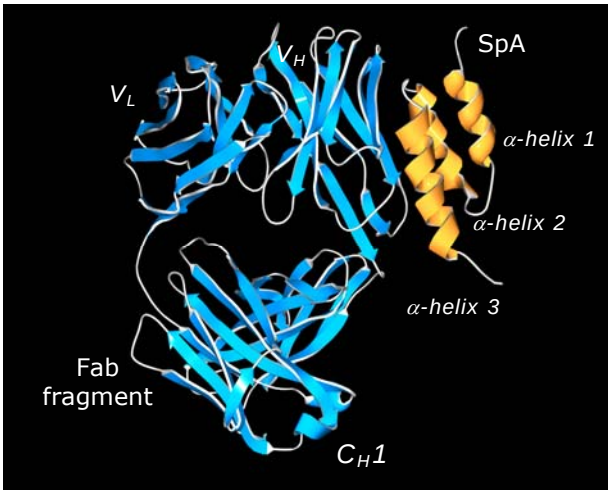


Figure 1.11 – A ribbon representation of the B domain of SpA interacting with the V_HIII domain of a Fab fragment of an IgG molecule

The interaction of SpA (gold) with the immunoglobulin molecule has been located in the variable region of the Ig heavy chain. In the complex, α-helices II and III of SpA Ig-binding domain interact with the variable region of the Fab (blue) heavy chain (V_H) through framework residues, without the involvement of the hypervariable regions implicated in antigen recognition. PDB accession code: 1DEE.

1.2.1.5 - Interaction with Fc and Fab

Previous studies have suggested that not all of the five homologous domains of SpA were involved in the Fab interaction [Ibrahim, 1993b], and it was suggested that Fab-binding was confined to the D domain [Roben *et al.*, 1995]. However, more recent studies have demonstrated that all five domains show affinity for both Fab and Fc [Jansson *et al.*, 1998]. Thus, SpA possesses two distinct Ig-binding activities, with each domain able to bind Fc and Fab, with a distinct specificity for the Fab region of some IgA, IgE, IgG and IgM. Figure 1.12 represents a ribbon diagram of the interaction between the B domain of SpA and both the Fc portion and V_HIII domain of Fab fragment of an IgG molecule.

The site responsible for Fab binding is structurally different from the domain surface that mediates Fc binding. The interaction with Fc involves residues in helix I of SpA, with lesser involvement of helix II. With the exception of residue Gln32, a minor contact in both complexes, none of the residues that mediate Fc binding are involved in Fab binding. The Fc and Fab structures form a sandwich around the opposite faces of helix II, without any evidence of steric hindrance of either interaction; with the variable region (V_HIII) lying toward the antigen-binding end of Fab and the Fc bound Ig projecting away from the complex.

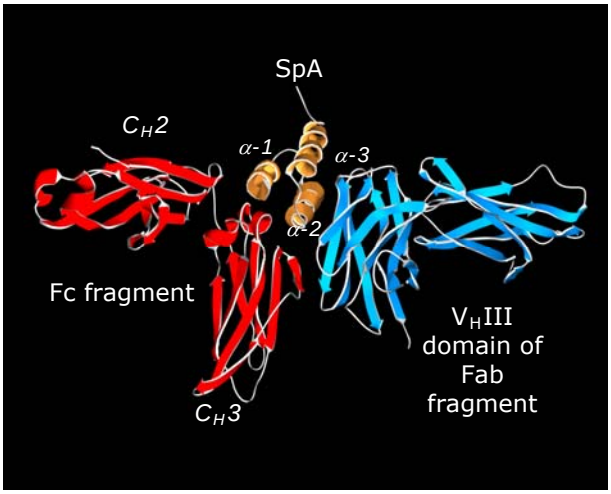


Figure 1.12 – <u>A ribbon representation of the interaction between the B domain of SpA interacting with both the Fc and Fab fragments of an IgG molecule</u>

SpA possesses two distinct Ig-binding activities, with each domain able to bind Fc and Fab simultaneously. Here SpA is complexed with Fc and Fab to illustrate how the site responsible for Fab binding is structurally separate from the domain surface that mediates Fc binding. The Fab and Fc are shown to form a sandwich around the opposite faces of helix II of SpA. The structure of SpA is shown in gold, the C_H2/C_H3 interface of the Fc fragment is shown in red and the V_H domain of Fab is shown in blue. PDB accession code: 1FC2 and 1DEE.

Recent X-ray crystallography studies have reported the structure of domain D of SpA complexed with the Fab fragment of human IgM antibody [Graille *et a*l., 2000], as shown in figure 1.13. In this complex, domain D assumes the triple α-helical bundle reported for domains B and E, and it is helix I and III which is reported to interact with the Fab, via a surface composed of four V_H region β-strands: B, C, D and E. The site of interaction on Fab is remote from the Ig light chain and the heavy chain constant region, and is mediated by thirteen heavy chain residues, six of which are in framework region β strands, with the other seven located in the V_H framework region interstrand loops on the side farthest from the antigen-binding pocket. The interaction also involves eleven residues from domain D, Gln26, Gly29, Phe30, Gln32, Ser33 and Asp36 of helix II; Asp37 and Gln40 in the loop between helix II and helix III; and Asn43, Glu47 and Leu51 of helix III [Graille *et a*l., 2000].

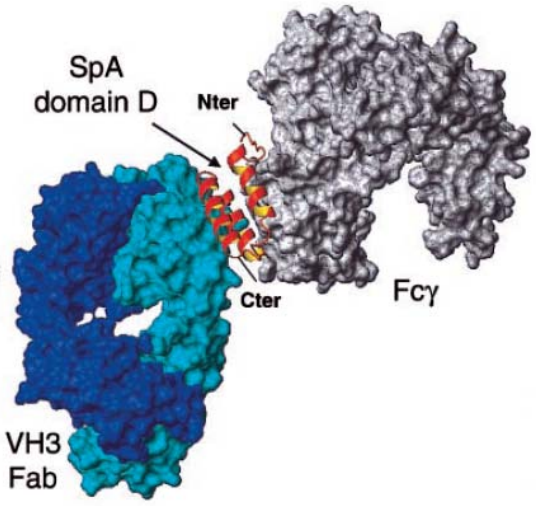


Figure 1.13 - <u>An illustration of the interaction between the D domain of SpA interacting with both the Fc and Fab fragments</u>

This model illustrates an individual SpA domain showing dual reactivity with both Fc and Fab fragments. Helix I and II of the SpA domain in the Fab-domain D complex and the previously determined Fc-domain B complex have been superimposed. The V_HIII Fab surface is shown as cyan, the Fc fragment surface is shown in grey and the single domain of SpA is represented by a red ribbon [Graille et al., 2000].

1.2.2 - Protein G (SpG)

Protein G (SpG) is a Type III bacterial Fc receptor found on the surface of, or secreted by the cell wall of, the β-haemolytic Gram-positive bacterium *Streptococcus* of the C and G strains [Kronvall, 1973]. Less is known about SpG than about SpA, primarily because SpG was not purified until 1984 [Björck and Kronvall, 1984]. However, interest in SpG has been maintained due to its ability to bind with IgG of different subclasses of most mammalian species.

β-haemolytic streptococci comprise a large and biological diverse group of microorganisms, and can be differentiated into groups in which the specific antigens are carbohydrates. Group C and G streptococci may inhabit various parts of the human body and cause severe and even fatal infections with indistinguishable clinical symptoms, the most common being pharyngitis [Efstratiou, 1989] and septicaemia [Auckenthaler *et al.*, 1983]. Studies have shown no difference in the structure of SpG when isolated from tonsillar strains and blood isolates, thus suggesting that SpG is not a virulence determinant, and may be involved in more basic biological events [Sjöbring *et al*., 1991].

1.2.2.1 - Structure of SpG

The gene for SpG has been cloned and can now be isolated from bacterial over-expression systems [Fahnestock *et al.*, 1986; Guss *et al.*, 1986]. Analysis of the amino acid sequence revealed a primary organisation similar to the genes of SpA and other Ig-binding proteins [Guss *et al.*, 1986]. For SpA, the part responsible for IgG binding consists of five highly homologous units with sizes ranging from 50 residues to 61 [Uhlén *et al.*, 1984a]. In contrast, the corresponding part of SpG consists of three domains namely C1, C2 and C3, with each of these units containing 55 amino acid residues with two spacers of 16 residues, D1 and D2. A comparison of the amino acid composition of regions C1, C2 and C3 is presented in figure 1.14, and illustrates how their sequences are highly homologous, with only six residues varying between them. It is also striking that regions D1 and D2 have been exactly conserved [Guss *et al.*, 1986].

It is not known why the SpG genes from strains G148 and GX7809 consist of two and three IgG-binding domains, respectively, while the SpA gene consists of five IgG-binding domains. However, it has been suggested that the number of IgG receptors may affect the strength and the kinetics of binding to different IgG molecules form different species [Olsson *et al.*, 1987]. Different strains may have adopted this receptor response to various external environments.

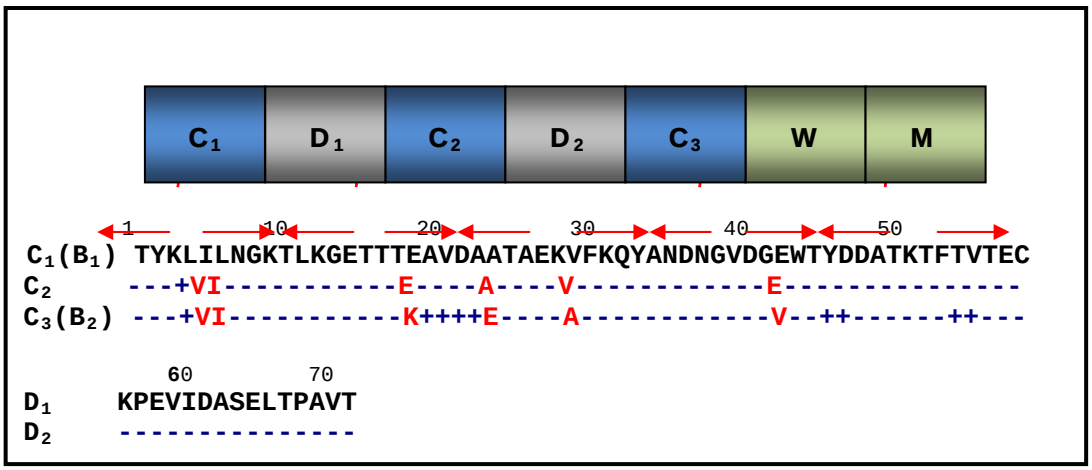


Figure 1.14 - A comparison of the amino acid composition of the Ig-binding domains of SpG

A comparison of the three C Ig-binding domain and the two D linker regions is presented as sequence alignments to achieve maximal homology. Identical amino acid residues are indicated with a blue bar, silent mutations, which have not altered the amino acid, are marked with a blue cross, and amino acids that have changed are marked with the appropriate amino acid letter, in red. The comparison indicates highly homologous sequences between the C1, C2 and C3 Ig-binding domains, with only six amino acids varying between them. The linker regions D1 and D2 have been exactly conserved. Fahnstock isolated a two IgG-binding domain SpG strain GX7809 and labelled the domains B1 and B2 [Fahnestock et al., 1986], whereas Guss isolated a three IgG-binding domain SpG strain G148 and labelled the domains C1, C2 and C3 [Guss et al., 1986]. As illustrated above, the sequences of C1 and B1, and C3 and B2 are identical, containing the same six amino acid variations. C2, a hybrid of C1 and C3, contains four amino acids variations from C1 and two from C3 [Olsson et al., 1987].

Following the IgG-binding domains there are cell wall-binding (W) and membrane-binding (M) domains near the C-terminus, which are involved in cell wall interactions. The sequence of region W is highly regular, with the first part being rich in proline residues followed by a pentapeptide consensus sequence, which is repeated six times. The whole region is very hydrophilic and secondary structure prediction by Chou and Fasman procedures [Chou and Fasman, 1974] illustrate low probabilities for the presence of alpha and beta structures. These predictions are comparable with other cell wall-binding regions of proteins such as SpA [Uhlén *et al.*, 1984a]. It has been suggested that this structure might serve to orient the protein on the outer surface of the cells, and thus sterically facilitate the function of the molecule [Uhlén *et al.*, 1984b].

Apart from the IgG-binding domains and the cell wall-spanning region the SpG gene also contains a third independent repetitive structure A and B. N-terminal regions A1, B1, A2, B2 and A3 are highly conserved when compared with each other, which strongly suggests an important biological function. Studies have

shown that these regions are involved in the binding to human serum albumin (HSA) [Åkerström *et al.*, 1987]. This would indicate that SpG is a bi-functional receptor possessing two independent regions with affinity to albumin and IgG.

1.2.2.2 - Structure of a single Ig-binding domain from SpG

High-resolution three-dimensional structures of the single IgG-binding domains from SpG have been determined by NMR spectroscopy [Gronenborn *et al.*, 1991; Lian *et al.*, 1992] and X-ray crystallography [Achari *et al.*, 1992; Derrick and Wigley, 1992; Gallagher *et al.*, 1994; Derrick *et al.*, 1994; Sauer-Eriksson *et al.*, 1995]. All domains have the same arrangement of secondary structure, consisting of a central α-helix and a four-stranded β-sheet, as presented in figure 1.15. NMR has shown that the two inner strands of the β-sheet, β1 and β4, are parallel to one another, and the two outer strands, β2 and β3 are anti-parallel to the two inner strands, β1 and β4 respectively [Lian *et al.*, 1991]. Strands β1 and β2 are connected by a type I turn, whereas strands β3 and β4 are connected by an unusual six-residue turn. The long axis of the helix lies at approximately 140° to the axes of β2 and β3, and lies diagonally across the β-sheet with an angle of Ω ~ -40° [Gronenborn *et al.*, 1991].

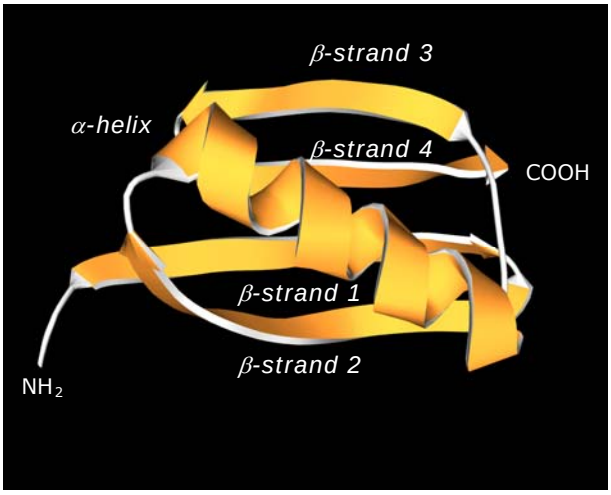


Figure 1.15 – A ribbon representation of an Ig-binding domain of SpG

The three-dimensional structure of a single IgG-binding domain (B1) of SpG consists of a central α-helix and a four-stranded β-sheet. NMR has shown that the two inner strands of the β-sheet, β1 and β4, are parallel to one another, and the two outer strands, β2 and β3 are anti-parallel to the two inner strands, β1 and β4 respectively. The long axis of the helix lies at approximately 140° to the axes of β2 and β3, and lies diagonally across the β-sheet with an angle of Ω ~ -40° [Gronenborn et al., 1991].

The single B1 and B2 IgG-binding domains of SpG are very stable, and differential scanning calorimetry studies have demonstrated extreme thermal stability with melting temperatures of 87.5°C and 79.4°C, respectively [Alexander *et al.*, 1992].

This feature is unusual considering the small size of each of the domains and the absence of disulphide bridges or tight ligand binding. It has been suggested that this unusual stability for a small protein is due to a number of factors. Firstly, the involvement of nearly all the residues in the regular secondary structure ensures that a large number of hydrogen bonds stabilises the protein. The secondary structure content of SpG is 95% compared to an average 75% in other proteins, with a total of 45 atoms participating in hydrogen bonds [Gronenborn *et al.*, 1991].

Secondly, the interior of SpG is highly hydrophobic, while the exterior is very hydrophilic. The increase in heat capacity of the unfolded state is directly related to the amount of hydrophobic surface that is buried in the folded state relative to the unfolded state [Achari *et al.*, 1992]. Thus, the B domain appears to be stable because it efficiently buries a significant hydrophobic surface in spite of its small size. Nine hydrophobic residues are completely buried, with only five hydrophobic residues exposed to solvent [Gronenborn *et al.*, 1991; Alexander *et al.*, 1992].

However, the features mentioned above are probably not sufficient in their own right to fully account for the stability of SpG. The overall polypeptide fold of the protein may itself play a key role in stabilising the protein. In particular, the hydrogen bonds between the N-terminal and C-terminal strands, β1 and β4, which constitute the two internal strands of the four stranded β-sheet, and their network of hydrophobic interactions with the overlaying helix, require disruption before unfolding can be accomplished. Thus, unfolding is highly cooperative and involves all parts of the protein structure simultaneously.

1.2.2.3 - Interaction with Fc

The structures of isolated Fc-binding repeats of SpG have been well documented and successfully determined by NMR and X-ray crystallography [Gronenborn *et al.*, 1993; Sauer-Eriksson *et al.*, 1995]. NMR data have provided the structure of the Fc-binding domain B1 of SpG [Gronenborn *et al.*, 1991], and the contact regions of the B1 IgG-binding domain, when complexed with human Fc, identified [Gronenborn *et al.*, 1993]. Figure 1.16 represents a ribbon diagram of the interaction between the B1 domain of SpG and the Fc fragment. Results show that the interaction site with Fc lies within the α-helix, β-strand three and the loop that connects them.

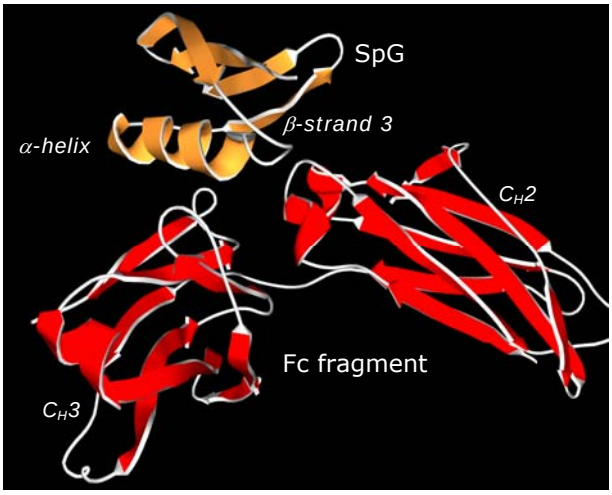


Figure 1.16 – A ribbon representation of an Ig-binding domain of SpG interacting with the C_H2/C_H3 interface of the Fc fragment from an IgG molecule

In this complex, the B1 Ig-binding domain of SpG (gold) binds Fc (red) at the hinge region that connects the C_H2 and C_H3 domains. The interactions with Fc lies within the α-helix and β-strand three of SpG, and involves residues 35 to 45, which are situated in the C-terminal part of the α-helix, the N-terminal part of β-strand 3, and the extended loop region connecting these two structures. The residues that interact with SpG are situated within loop regions of Fc [Sauer-Eriksson et al., 1995]. PDB accession code: 1FCC.

Corroboratory evidence has come from studies that demonstrated that a small 11 residue peptide fragment involving residues 35 to 45, situated in the C-terminal part of the α-helix, the N-terminal part of β-strand 3, and the extended loop region connecting these two structures, blocks the interaction between IgG•Fc and SpG [Frick *et al.*, 1992].

The crystal structure of the C2 domain in complex with Fc was solved by Sauer-Eriksson *et al.*, (1995). The B1 domain is equivalent to the C1 domain, and the amino acid sequences of the C1 and C2 domains are identical except for residues 5 and 6. Thus, it might be assumed that the structures of C1 and C2 are identical [Olsson *et al.*, 1987]. Studies confirmed that SpG binds Fc at the interface region that connects the C_H2 and C_H3 domains. There are three residues of the C_H2 domain of Fc, which are involved in the interfacial interactions, namely Ile253, Ser254 and Gln311. In the C_H3 domain, there are two areas which contribute to the interface, namely Glu380 and Glu382 and the residues His433 to Gln438 [Sauer-Eriksson *et al.*, 1995]. These residues that interact with SpG are situated within loop regions of Fc. The residues in SpG that interact with Fc have been identified as Glu27, Lys28, Lys31, Gln32, Asn35, Asp40, Glu42 and Trp43. The interfacial interactions between Fc and SpG consist of a network of hydrogen bonds and salt links, involving mainly charged and polar residues.

1.2.2.4 - Interaction with Fab

SpG has been shown to interact with the Fab regions of IgG, but with a 10-fold lower affinity than that determined for the Fc region (Bjorck and Kronvall, 1984). NMR and X-ray crystallography has been used to study complexes of SpG domains with Fab fragments [Derrick and Wigley, 1992; Lian *et al*, 1994; Derrick and Wigley, 1994; Sauer-Eriksson *et al*, 1995]. In this complex, the outer second β-strand of SpG, at the edge of the β-sheet, forms an anti-parallel extension with the last β-strand of the C_H1 domain of IgG, thus extending the β-sheet into SpG [Derrick and Wigley, 1994]. Figure 1.17 illustrates the interaction between the B1 domain of SpG and the Fab fragment of an IgG molecule.

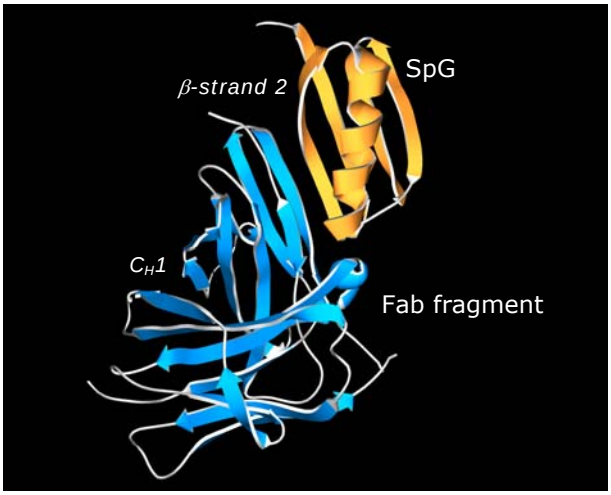


Figure 1.17 – A ribbon representation of an Ig-binding domain of SpG interacting with the C_H1 domain of a Fab fragment

The overall structure of the SpG:Fab complex is presented. In this complex, the outer second β-strand of the B1 Ig-binding domain of SpG, at the edge of the β-sheet, forms an anti-parallel interaction with the last β-strand of the C_H1 domain of IgG, thus extending the β-sheet into SpG. [Derrick and Wigley, 1992]. PDB accession code: 1IGC.

The atoms involved in these interactions are predominantly from the backbone of the C_H1 domain of the IgG, but include both backbone and side-chain atoms from the binding domain of SpG. The alpha-beta structure of SpG in solution is conserved on binding to Fab, and this complex is stabilised through a network of hydrogen bonds from the last β-strand of the C_H1 domain to the second β-strand in SpG. A second, minor region of contact occurs between the C-terminal end of the α-helix in SpG and the first β-strand of the C_H1 domain [Derrick and Wigley, 1994]. This complex is stabilised further by the association of exposed non-polar residues from Fab and SpG, providing a continuous hydrophobic core, which is shielded from solvent [Derrick and Wigley, 1992]. The SpG and Fab complex provides an explanation for the wide specificity of SpG for immunoglobulins, as

the interactions involve backbone atoms of a highly conserved sequence in C_H1 [Tashiro and Montelione, 1995].

1.2.2.5 - Interaction with Fc and Fab

SpG binds selectively to both the Fc and Fab regions of most mammalian IgGs, and SpG has also been reported to bind F(ab')2 fragments. Previous studies have indicated that the binding of Fc and Fab to SpG does not interfere with one another [Erntell *et al.*, 1983], suggesting that there are independent and separate binding regions for both Fab and Fc fragments of IgG on the SpG molecule [Erntell *et al.*, 1988]. Figure 1.18 illustrates the specific amino acid residues of SpG that contribute to both Fc and Fab binding sites, indicating that the binding interactions for these fragments occur in different regions of the protein.

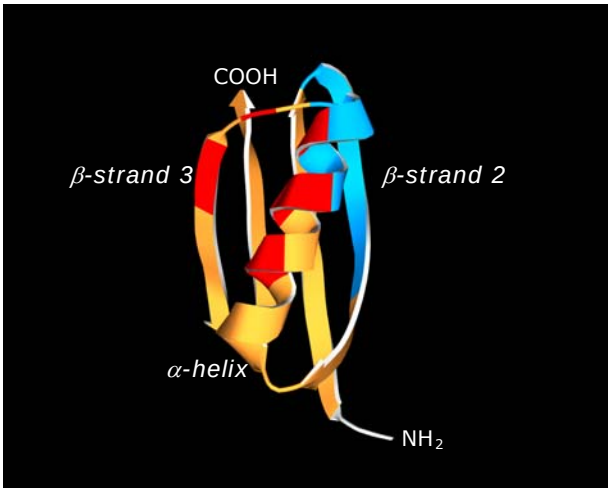


Figure 1.18 – **A ribbon representation of SpG indicating both Fc and Fab binding regions**

Specific amino-acid residues that contribute to these binding interactions via a network of hydrogen bonds have been identified. The Fc binding region, α-helix and β-strand 3, is shown in red; and the Fab binding region, α-helix and β-strand 2, is shown in blue.

It has been suggested that the elongated structure of the SpG molecule [Åkerström and Björck, 1986] permits the simultaneous binding of both Fab and Fc fragments [Erntell *et al.*, 1988]. Figure 1.19 demonstrates SpG in complex with both Fc and Fab fragments, and clearly demonstrates how the site responsible for Fab binding is structurally separate from the domain surface that mediates Fc binding. SpG uses its β-strand 2 to form an anti-parallel interaction with the last β-strand of the C_H1 domain of IgG-Fab and the α-helix and β-strand 3 to bind predominantly to the interface of C_H2-C_H3 heavy chains of IgG-Fc.

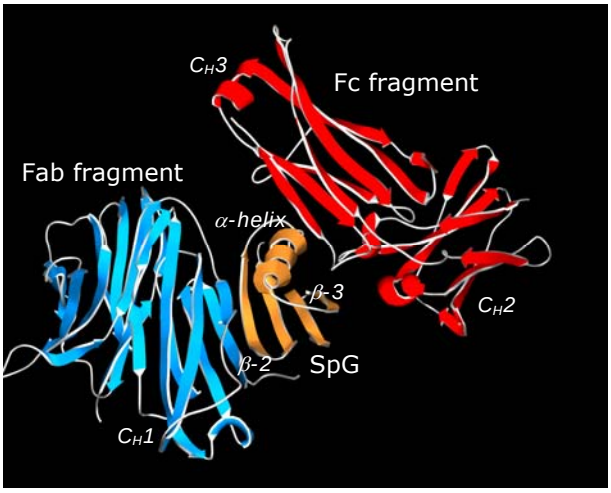


Figure 1.19 – A ribbon representation of the interaction between SpG and both Fc and Fab fragments of an IgG molecule

SpG is complexed with Fc and Fab fragments to illustrate how the site responsible for Fab binding is structurally separate from the domain surface that mediates Fc binding. The structure of SpG is shown in gold, the C_H2/C_H3 interface of the Fc fragment is shown in red and the C_H1 domain of Fab is shown in blue. PDB accession code: 1FCC and 1IGC.

1.2.2.6 - Reactivity with immunoglobulins

Studies have shown that SpG expresses separate, repetitively arranged domains for both IgG and human serum albumin (HSA), and has also been found to interact with plasmin together with α_2-macroglobulin (α_2M) and kininogen, two proteinase inhibitors of human plasma [Sjöbring *et al*, 1989]. SpG has also been shown to bind all varieties of rat and mouse monoclonal antibodies and a large number of polyclonal IgGs from cow, goat, rabbit and mouse.

SpA has been an extremely useful chemical reagent, used in a variety of immunochemical and immunological techniques. However, its usefulness has been affected by the fact that some species and subclasses of IgG do not react with this protein. In contrast, SpG has been demonstrated to be more generally IgG binding, with all mammalian IgG molecules showing affinity [Åkerström and Björck, 1986]. This makes SpG a potentially useful tool where binding and detection of IgG antibodies are required. Studies have also demonstrated higher affinity for the binding of SpG to IgG than SpA [Åkerström *et al*, 1985], especially for rat and goat polyclonal IgG and human IgG3, thus suggesting the binding site on SpG is subtly different from that on SpA due to the presence of a single helix.

SpG is responsible for the binding of both human serum albumin (HSA) and IgG, and these reactions occur at different sites on the SpG molecule. Albumin-binding has been located to be the amino-terminal half and the IgG-binding to the

carboxyl-terminal half of the bacterial cell wall molecule [Åkerström *et al.*, 1987]. This means that SpG has evolved separate binding sites for IgG Fc, IgG Fab and albumin. The albumin-binding activity of SpG resides in the amino-terminal 56 amino acid peptide, which contains sequence unit A_3 and part of B_2, as shown in figure 1.8. This sequence is repeated three times in the SpG structure reported by Guss *et al.* 1986, and twice in the sequence of Fahnestock *et al.* 1986, and corresponds to amino acid residues 104 to 140, 179 to 215 and 254 to 290. It has also been shown that some strains of C and G streptococci bind HSA poorly [Wideback and Kronvall, 1982], whereas IgG-binding is consistently high. This suggests that the IgG-binding domains (C repeats) are associated with essential biological functions, whereas the albumin-binding A and B repeats seem less indispensable.

Apart from IgG and albumin, SpG also shows affinity for plasmin, kininogen and α_2-macroglobulin [Sjöbring *et al.*, 1989]. Studies have also demonstrated that Ig-binding proteins, namely SpA and PpL, interact with α_2-macroglobulin and kininogen with the same pH optima for the binding of these proteins to IgG. The interaction between SpG and α_2-macroglobulin has been characterised in most detail as this protein is a major plasma protein, which has the capacity to block a wide range of proteinases [Barrett and Starkey, 1973]. It has been shown that α_2-macroglobulin and IgG compete for the same binding sites in the C domains of SpG [Sjöbring *et al.*, 1989], despite the lack of structural similarities between the two, and with IgG being bound with double the affinity of α_2-macroglobulin. It has also been shown that this interaction takes place irrespective of whether α_2-macroglobulin is complexed with proteinase or not, and that SpG does not influence the binding of proteinases to either α_2-macroglobulin or kininogen [Sjöbring *et al.*, 1989].

Proteolytic activity is important for the multiplication and tissue penetration of microorganisms, and the binding of these proteinase inhibitors from the infected host with the streptococcal cell surface through interaction with SpG could interfere with a variety of proteolytic events. Alternatively, proteolytic inhibition can be beneficial to the parasite by helping the bacteria to avoid proteolytic attack by host enzymes. α_2-macroglobulin also has immunoregulatory functions through interactions with lymphokines and immune complexes [James, 1980], these functions also provide advantages to the host, as α_2-macroglobulin: proteinase complexes at the bacterial surfaces facilitate phagocytosis of the bacteria [Sjöbring *et al.*, 1989]. Thus, it is unclear why SpG has retained its

binding affinity for α_2-macroglobulin.

1.2.3 - Protein L (PpL)

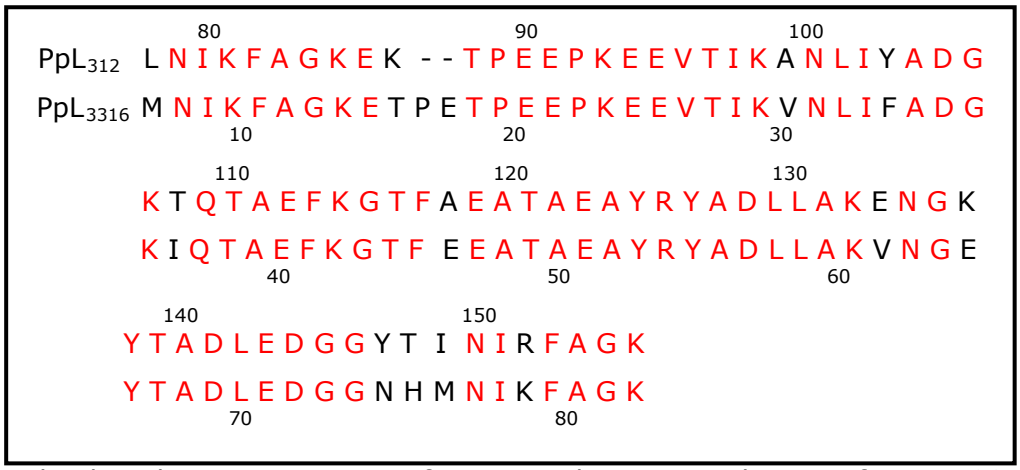
Protein L is a novel immunoglobulin-binding protein that has been isolated from the surface of bacteria belonging to the anaerobic species *Peptostreptococcus magnus (P. magnus)* [Myhre and Erntell, 1985]. It has been shown that protein binds an immunoglobulin through a variable domain light-chain interaction, giving rise to the name PpL [Björck, 1988]. It is a multi-domain protein, which is expressed at the surface of some strains of the anaerobic bacterium [Björck, 1988], and appears to be correlated to bacterial virulence [Kastern *et al*., 1990].

PpL is preferentially expressed by strains of *P. magnus* isolated from patients with bacterial vaginosis [Ricci *et al*., 2001], and only strains that express PpL possess the PpL-encoding gene, which can bind different classes of immunoglobulin to their cell surface with high affinity. This is thought to dramatically change their surface properties and the binding of different classes of Ig could influence important cellular and molecular events, which are significant to a balanced coexistence between the host and parasite [Kastern *et al*., 1990]. *P. magnus* are normally found on most body surfaces, where they are members of the indigenous flora of the skin, the oral cavity and gastrointestinal and genitourinary tracts, and are also causative agents in a variety of infections.

Studies have shown that PpL_{312} (PpL from *P. magnus* strain 312) induces *in vitro* mediator release from human basophils and mast cells isolated from lung parenchyma and skin tissues, which could be one possible virulence mechanism [Patella *et al*., 1990]. The releasing activity of PpL appears to be mediated by an interaction with the kappa-light chains of IgE present on human basophils and mast cells to induce the release of pro-inflammatory mediators. These results were the first to indicate that a bacterial SpA associated with bacterial virulence [Kastern *et al*., 1990] can induce the release of preformed histamine and *de novo* synthesised chemical mediators from human inflammatory cells. SpA also acts through an IgE-dependent mechanism, but induces the release of mediators from basophils by interacting through the $F(ab')_2$ region of polyclonal IgE on the basophil membrane. It does not activate lung and skin mast cells [Genovese *et al*., 2000]. SpG, however, failed to induce the release of mediators from either basophils or mast cells.

1.2.3.1 - Structure of PpL

PpL is an elongated fibrous protein, and the gene encoding PpL has been sequenced from two strains of *Peptostreptococcus*, strain 312 (PpL_{312}) [Kastern *et al.*, 1992] and strain 3316 (PpL_{3316}) [Murphy *et al.*, 1994]. The primary structure of PpL from both strains shows 59% homology, comprising of a hydrophobic signal sequence, an NH_2-terminal domain, and a repeated region with Ig-binding activity, with each domain containing approximately 72 to 76 amino acid residues. PpL_{312} possesses five homologous Ig-binding domains, namely B_1-B_5 [Kastern *et al.*, 1992], whereas PpL_{3316} possesses four homologous Ig-binding domains, C_1-C_4, and four albumin-binding domains, D_1-D_4 [Murphy *et al.*, 1994]. The sequence of these Ig-binding domains within each strain shows 80% homology, with only 60% homology between these strains. A comparison of



the Ig-binding domain sequences of PpL_{312} and PpL_{3316} is shown in figure 1.20.

Figure 1.20 – A comparison of the Ig light chain-binding domain sequences of PpL_{312} and PpL_{3316}

The amino acid residues comprising the Ig light chain-binding domains from both strains of Peptostreptococcus magnus PpL, PpL_{312} and PpL_{3316}, are presented as sequence alignments and compared. Regions of homology between the two strains are shown in red. Residue numbering is that given by Wikström et al., 1993 and Murphy et al., 1994 for strains 312 and 3316 respectively.

Following a spacer region of 8 amino acids, there are additional non-related repeated regions of unknown function, and further downstream there is a highly hydrophilic cell wall-spanning domain and the transmembrane COOH-terminal region. The COOH-terminal regions, seen in other Gram-positive bacterial cell surface proteins, are thought to mediate the binding of the PpL to the bacterial cell wall [Kastern *et al.*, 1992].

1.2.3.2 - Structure of a single Ig-binding domain from PpL

The three-dimensional solution structure of a single Ig-binding domain of PpL has been investigated by using ^{1}H NMR [Wikström *et al.*, 1993; Wikström *et al.*, 1994; Wikström *et al.*, 1995]. The first of the five homologous Ig light chain-binding domains of PpL_{312}, the B1 domain, is 76 amino acid residues in length (Lys80 to Gly155), and has been shown to adopt a well-defined structure in solution. The domain consists of a 15 residue (Glu78 to Glu94), highly flexible disordered N-terminus followed by a 61 residue (Val95 to Gly155), folded unit comprising a four-stranded β-sheet with a central α-helix crossing over on top. The three-dimensional structure of PpL is presented in figure 1.21.

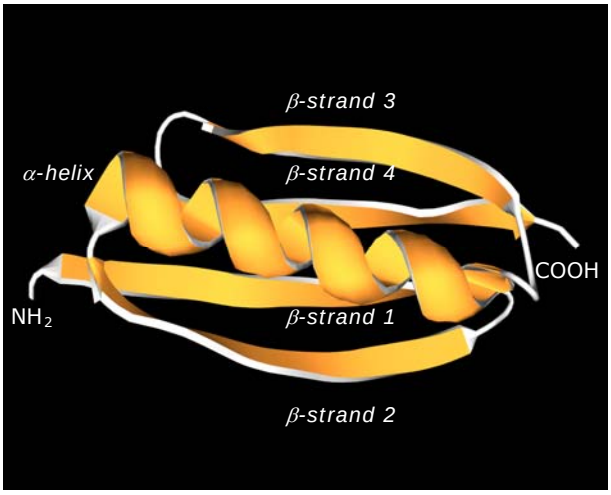


Figure 1.21 – A ribbon representation of an Ig-binding domain of PpL

The three-dimensional structure of PpL is composed of a disordered N-terminal polypeptide segment, followed by a folded portion composed of a central α-helix packed against a four stranded β-sheet, with a -1, +3x, -1 topology. The orientation of the α-helix runs almost parallel with the β-sheet, at an angle of -10° [Kastern et al., 1992].

The N- and C- terminal strands of the B_1 domain, β1 (Val95 to Phe103) and β4 (Tyr147 to Ala154) respectively, form a central parallel pair of strands which places the β2 (Ser107 to Gly115) and β3 (Glu137 to Arg144) strands on opposite sides of the domain [Wikström *et al.*, 1993]. The α-helix involves residues Phe117 to Leu131, and crosses the β-sheet at an angle of -10°. The segment following the helix (Lys132) is a compact loop connecting the helix to the third β-strand (Gly137). Both the helix and β-sheet, which shows a characteristic right-handed twist, combine their hydrophobic surfaces into an inner hydrophobic core.

The disordered polypeptide segments are similarly charged, and comprise the N-terminus. These regions are present in all five domains, and have been proposed

to function as linking segments between Ig-binding domains, and were thought to facilitate the interaction with Ig light chains by increasing the individual flexibility of the domains [Wikström *et al.*, 1993; Wikström *et al.*, 1994]. Recent NMR studies have confirmed that the flexible N-terminal is not involved in binding, as there are no major conformation changes upon binding [Wikström *et al.*, 1995].

1.2.3.3 - Interaction with Ig light chains

PpL possesses specific binding sites for human immunoglobulin, and early studies demonstrated that PpL bound Fab and $F(ab')_2$, but not Fc fragments, implying a reactive site in the Fab portion of the molecule [Myhre and Erntell, 1985]. Reactivity was shown with human Ig molecules, regardless of heavy chain class, through interaction with Ig kappa light chain binding sites, exclusively located in the V_L domain [Nilson *et al.*, 1992]. No binding interactions occur with the C_L domain. The κ and λ light chains are representative of the major human variable region (V_L) subgroups, comprising four $V_{\kappa I}$, $V_{\kappa II}$, $V_{\kappa III}$ and $V_{\kappa IV}$ subgroups, and five $V_{\lambda I}$, $V_{\lambda II/V}$, $V_{\lambda III}$, $V_{\lambda IV}$ and $V_{\lambda VI}$ subgroups. The interaction between PpL and the variable light chains is limited to the κ-chains and, more specifically, to proteins of the $V_{\kappa I}$, $V_{\kappa III}$ and $V_{\kappa IV}$ subgroups. No binding occurs with proteins of the $V_{\kappa II}$ subgroup or with any λ light chain subgroups [Nilson *et al.*, 1992].

^{1}H-^{15}N NMR spectroscopy studies were utilised to define the interaction between an Ig-binding domain of PpL (B1) and the variable domain of human Ig κ-chain strand [Wikström *et al.*, 1995]. Results show that the Ig-binding region of the PpL domain involves residues in the second β-strand (residues 109-114), the C-terminal residues of the α-helix (residues 127, 129-131) and the loop connecting the α-helix with the third β-strand (residues 132-133 and 136) [Wikström *et al.*, 1995]. Figure 1.22 represents a three-dimensional ribbon representation of PpL interacting with the V_L domain of Fab. V_L domain residues important for the interaction with PpL have been localised and comprises parts of three β-strands that are all in close proximity (8-11, 19-23 and 72-74). Thus, PpL binds to the outer surface of the framework region of the V_L domain, primarily involving the second β-strand [Enokizono *et al.*, 1997].

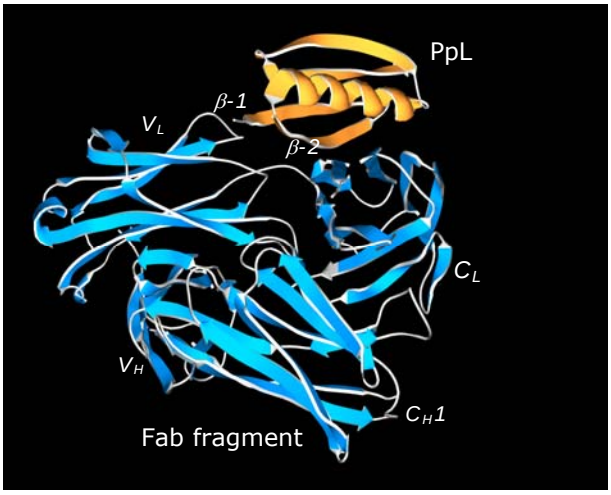


Figure 1.22 – A ribbon representation of an Ig-binding domain of PpL interacting with the V_L region of Fab

A three-dimensional ribbon representation of a single Ig-binding domain of PpL and the V_L region of a Fab fragment is presented. Strand β2 of PpL interacts with an Fab V_L β strand, forming a continuous β sheet with the Fab through a β-zipper interaction [Wikström et al., 1995]. PDB accession code: 1MHH.

1.2.3.4 - Interaction with F(ab')$_2$

Recent crystallography studies have shown that an Ig-binding domain of PpL has two distinct regions that can interact with kappa light chains and is capable of binding two Fab fragments (F(ab')$_2$) simultaneously [Graille *et al.*, 2001]. The crystal structure of human antibody F(ab')$_2$ complexed through its V_L region to a PpL Ig-binding domain has been resolved to 2.7 Å, and revealed that the single PpL domain is positioned between the V_L regions of two Fab fragments. A three-dimensional ribbon representation of an Ig-binding domain of PpL•F(ab')$_2$ complex is presented in figure 1.23.

The interaction between PpL and the first V_L domain involves 13 residues from the Fab molecule and 12 residues from the PpL Ig-binding domain [Graille *et al.*, 2001]. Ten of the residues from the Fab fragment involved in the interaction are located in framework region one without contacting the hypervariable loops, thus not affecting the antigen-binding site. The remaining three residues being Lys107 from the segment connecting the V_L to the C_L region, Glu143 from the C_L region and Arg24 from the second β-strand. All of the twelve residues in the PpL domain are located mainly on the second β-strand and the α-helix. This first interface is characterised by the formation of six hydrogen bonds positioned between the second β-strand of PpL and β-strand A from the first Fab fragment. These hydrogen bonds join the β-sheets of the Fab and PpL into a unique sheet through a β-zipper interaction [Graille *et al.*, 2001].

The interaction between PpL and the second V_L domain involves 15 residues from the Fab fragment and 14 residues from the PpL domain. The amino acid side chains contributed by the Fab fragment to the interaction are located on β-strands A and B, with some participation from β-strands D and E. It has also been shown that ten of these residues are also involved in the first interface. The residues donated from PpL are located on the third β-strand and the α-helix, but none of these residues are involved in the first interaction. Again six hydrogen bonds are formed, with two additional salt bridges mediating the interaction.

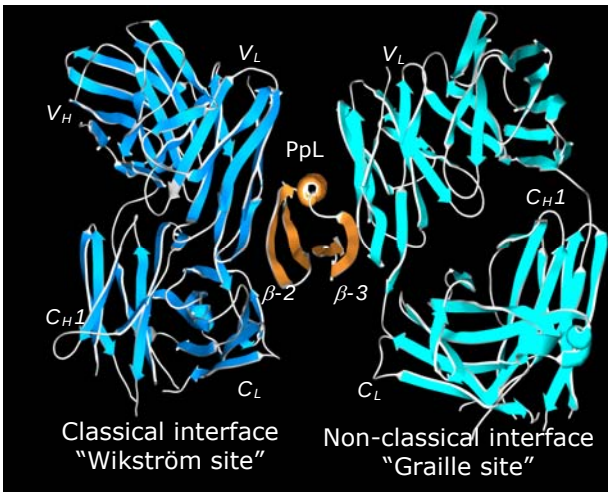


Figure 1.23 - A ribbon representation of an Ig-binding domain of PpL interacting with the V_L region of two Fab fragments

A three-dimensional ribbon representation of two human antibody Fab fragments and a single PpL domain complex is presented. The structure contains two Fab molecules sandwiching a single PpL domain, which contacts similar V_L framework regions of two light chains via independent interfaces. The first interface involves residues located on β2 and the α-helix, six hydrogen bonds join the β sheets of the Fab and the PpL into a unique sheet through a β zipper interaction. The second interface involves residues located mainly on β3 and the α-helix, and is mediated by six hydrogen bonds and two salt bridges [Graille et al., 2001]. PDB accession code: 1HEZ.

1.2.3.5 - Reactivity with immunoglobulins

PpL is of great interest because the interaction is with the light chains of the immunoglobulin carrying the kappa isotype. Specifically, PpL targets the framework regions of the kappa chain V_L domain [Enokizono *et al.*, 1997], without interfering with the antigen-binding capacity of the antibody, which involves the hypervariable regions [Åkerström and Björck, 1989]. Compared to the other Ig-binding proteins, PpL has the advantage of interacting with Ig regardless of the class specific heavy chains, and can be used to detect Ig of all classes that contain kappa chains. However, the reactivity with human kappa chains is limited to the κI, κIII and κIV subgroups, with approximately 65% of human immunoglobulins carrying kappa light chains [Kastern *et al.*, 1990], while

light chains of the κII subgroup show no affinity for PpL [Nilsson *et al*., 1992]. The dissociation constant (Kd) between PpL and kappa chains is 1.5 x 10^9 M^{-1}, compared with 1 x 10^{10} M^{-1} for IgG, IgM and IgA [Kastern *et al*., 1990]. Studies have also shown that the binding of PpL is highly dependent upon the tertiary structural integrity of the kappa chain V_L domain [Nilsson *et al*., 1992].

PpL also binds to the light chains of immunoglobulins from species other than man, such as human, baboon, chimpanzee, pig, mouse, rat, rabbit and guinea pig, but not to cow or goat Ig [De Chateau *et al*., 1993].

Cell surface proteins from Gram-positive bacteria are known to bind a number of serum proteins other than immunoglobulins. These proteins include albumin, fibrininogen, α_2-macroglobulin and kininogen [Sjöbring *et al*., 1989]. Studies have shown that in addition to binding Ig, peptostreptococcal strain PpL_{3316} binds human serum albumin (HSA), and also contains extra sets of different repeat units to the kappa light chain-binding units that both peptostreptococcal strains, PpL_{312} and PpL_{3316}, display [Murphy *et al*., 1994]. Western blot analysis has demonstrated that the function of the PpL_{3316} D1-D4 units is to bind albumin, and this separate location of albumin-binding regions to the Ig-binding regions is similar to that found in SpG [Björck *et al*., 1987]. A comparison of the D repeats with the N-terminal regions, A1, B1, A2, B2 and A3, of SpG shows considerable sequence homology with 67% shared amino acid sequence identity [Murphy *et al*., 1994].

Due to its site of interaction on V_L domains PpL also binds single chain Fv (scFv) antibodies. ScFv antibodies are single chain fragments of Ig consisting of heavy and light chain variable domains covalently linked together by a designed peptide linker sequence, which links the carboxyl terminus of the V_L sequence to the amino terminus of the V_H. This is the smallest possible fragment of an antibody which still retains an intact antigen binding site [Bird *et al*., 1988]. Kd values were shown to be between 1.4 and 4.5 x 10^9 M^{-1} when tested with scFv antibodies [Åkerström *et al*., 1994], compared to those obtained for the binding of PpL to isolated V_L domains at 0.9 x 10^9 M^{-1} [Nilsson *et al*., 1992], demonstrating that PpL was not influenced by the introduction of the peptide linker. Results also showed that all V_L domains belonging to the κI subgroup bound to PpL, regardless of which V_H domain it was connected to, suggesting that the binding is independent of the V_H domain. Thus, all available data are indicative that the binding of PpL to the V_L domain is independent of the

interactions with antigen, V_H domains and C_L domains [Åkerström *et al*., 1994].

1.2.4 - Protein H

Protein H is a surface protein of Group A *Streptococcus pyogenes (S. pyogenes),* belonging to the M protein family, which interacts with the constant Fc region of IgG [Gomi *et al*., 1990]. An IgG-binding fragment is released from the streptococcal surface by an extracellular cysteine proteinase produced by the bacteria. This fragment covers the entire surface-exposed part of the molecule, and exhibits high affinity for IgG, one of the most abundant soluble extracellular human proteins. *S. pyogenes* is an important human pathogen causing primary disease manifestations such as pharyngitis, impetigo and tonsillitis, which may lead to serious sequelae including rheumatic fever and glomerulonephritis [Kihlberg *et al*., 1999].

Studies suggest that protein H is always complexed with IgG, and when the protein H•IgG complexes are released from the surface of the bacteria, they are capable of activating the classical complement pathway, but prevent the activation of complement when associated with the bacterial surface [Berge *et al*., 1997]. Complement plays an important role in defence against pathogenic micro-organisms by facilitating antigen clearance and generating an inflammatory response. The complement reaction amplifies the initial antibody-antigen reaction and converts the reaction into a more effective defence mechanism. As the antigen becomes coated with complement reaction products, it is more readily phagocytosed by phagocytic cells that bear receptors for these complement products. Protein H belongs to the M protein family, which specifically binds to the complement proteins factor H and C4b-binding protein that regulates complement activation. *S. pyogenes* also expresses a surface-associated peptidase, which degrades C5a, a chemotactic fragment of C5 [Berge *et al*., 1997]. These effects provide a selective advantage to *S. pyogenes* and may help to explain how Ig-binding surface proteins contribute to the virulence of this human pathogen. When the protein H•IgG complexes are released complement is activated, which leads to its breakdown in the vicinity of the bacteria, thus preventing assembly and activation of complement at the bacterial surface. Conversely, complement activation will generate C5a and therefore attract phagocytic cells. *S. pyogenes* then efficiently and simultaneously releases the biologically active fragment of a C5a peptidase which destroys C5a [Kihlberg *et al*., 1999]. To initiate complement activation away from the bacterium and to inactivate the C5a that is generated represents a sophisticated microbial defence mechanism.

Protein H has also been shown to bind HSA, α_2-macroglobulin and IgG, and the Ig-binding domains have been identified as the C1-C3 domains [Åkesson *et al*., 1990]. Studies have also demonstrated that although the affinity for human IgG is ten times lower than that of proteins A and G, protein H showed the same binding patterns as these proteins. Thus, protein H binds at the C_H2/C_H3 interface on the Fc of IgG, and results show that this binding is competitive with that of SpA and SpG, despite the complete lack of sequence homology between the Ig-binding domains of these proteins [Åkesson *et al*., 1994]. Protein H also interacts with α_2-macroglobulin *in vivo*, which could have implications for virulence, as α_2-macroglobulin is the major proteinase inhibitor in human plasma and exhibits affinity for the three main Ig-binding proteins A, G and L [Sjöbring *et al*., 1989]. The reactivity with Ig classes from various species other than human has been identified and is limited to the IgG of baboons, rabbits and pigs [Åkesson *et al*., 1994].

1.2.5 - Protein D

Protein D is a surface protein of the bacterial species *Haemophilus influenza (H. influenza)*, and is the only Ig-binding receptor to date that has been identified in a Gram-negative bacterium [Ruan *et al*., 1990]. *H. influenzae* is known to cause invasive and non-invasive diseases in children and adults, and can cause pneumonia in the elderly and immune-compromised patients [Sasaki and Munson, 1993]. This bacterium was first identified to be capable of binding IgD [Forsgren and Grubb, 1979], with affinity for all four IgD myeloma proteins. Protein D interacts with the highest affinity to the Fc region on an IgD molecule and can also interact with the Fab region. The constant parts of the heavy chains both in the Fab and Fc fragments have been identified as the contact regions for the interaction with the protein [Ruan *et al*., 1990]. Unfortunately the structure of protein D has not yet been determined, however, since *Haemophilus influenzae* colonises the upper respiratory tract, protein D can be used to study bacteria-IgD interactions on lymphocytes and in secretion, which may play an important role in pathogenesis and host defence in upper respiratory tract infections.

1.2.6 - Comparison of Ig-binding proteins A, G and L

1.2.6.1 - Proteins A and G

Immunoglobulin-binding proteins are important reagents for a variety of biochemical and immunological applications. Staphylococcal SpA and streptococcal SpG are examples of cell-surface proteins that bind to the constant

Fc region of mammalian IgG with the advantage that they do not interact with the antigen-binding part of the antibody. SpA and SpG share no sequence or structural homology even though they compete for the same binding site on Fc, at the hinge region that links the second and third constant domains of the heavy chain (C_H2/C_H3 interface). The interaction with Fc is comparable for both proteins, and it has been suggested that the α-helix of SpG might occupy the same position as the first helix of SpA [Gronenborn *et al*., 1993]. This is supported by similar contact regions when complexed with Fc, involving residues located mainly on the N-terminal helix of SpA, and residues within the α-helix and β-strand three of SpG. Conversely, although these interactions seem similar, it has been shown that the helix of SpG does not coincide with any of the helices of SpA when complexed with Fc. The interfacial interactions in the two complexes are also very different. The interface between Fc and SpG consists of a network of hydrogen bonds and salt links, involving mainly charged and polar residues, whereas the SpA•Fc complex is stabilised by non-specific hydrophobic interactions and fewer polar contacts.

Crystallography data has shown that by superimposing the two complexes, about one-quarter of SpG overlaps with about one-third of SpA. The helix of SpG makes an angle of approximately 45° to the plane formed by the two longest helices of SpA. Two helices of SpA are shown to be located mostly on the C_H2 side of the Fc molecule, whereas the helix of SpG lies wedged in the C_H2-C_H3 cleft [Sauer-Eriksson *et al.*, 1995], thus the two helices cannot bind to the same region of Fc simultaneously. Figure 1.24 presents a ribbon representation of the SpA•Fc complex superimposed with the SpG•Fc complex.

Comparatively, SpG does offer some other similarities with SpA. Both proteins are flanked by, and interact with, three loops from Fc. Four residues from Fc, Ile253, Ser254, Gln311 and Asn434, are particularly important because their side chains interact with side chains of both SpA and SpG. Several of the Fc residues interact with both proteins, which are possible because the third β-strand in SpG is situated in approximately the same region on Fc as the first α-helix of SpA [Tashiro and Montelione, 1995]. These and other interactions involving main-chain atoms in these three loops suggest how the two proteins compete for binding to Fc. However, SpA and SpG each have a set of unique interactions with Fc, which accounts for the differences strategies in Fc binding.

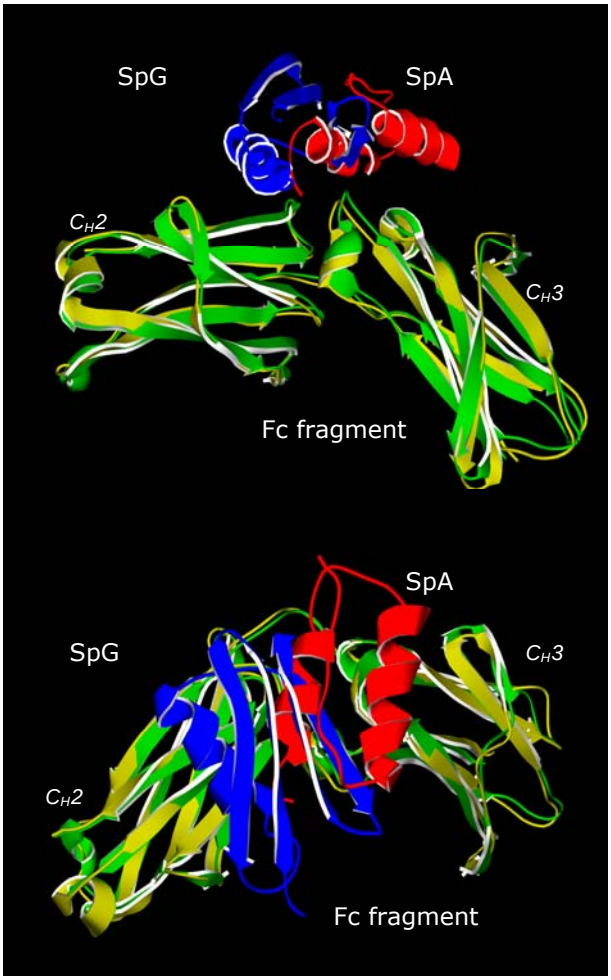


Figure 1.24 – A superimposed ribbon representation of SpA and SpG interacting with an Fc fragment from an IgG molecule

Crystallography data have shown that by superimposing the two complexes, about one-quarter of SpG overlaps with about one-third of SpA. Two helices of SpA are shown to be located mostly on the C_H2 side of the Fc molecule, whereas the helix of SpG lies wedged in the C_H2-C_H3 cleft, thus the two helices cannot bind to the same region of Fc. The two superimposed C_H2/C_H3 interfaces of Fc are shown in green and yellow, SpA is shown in red and SpG is shown in blue. The bottom picture represents a top view of the superimposed complexes [Sauer-Eriksson et al., 1995].

1.2.6.2 - Proteins G and L

SpG and PpL share many common features; such is the extent of these that it has been proposed that a transfer of gene fragments may have occurred between these two infective bacteria [Goward *et al*., 1993]. In spite of different binding specificities, the Ig-binding domains of PpL and SpG demonstrate a similar structural fold, both having a central alpha helix on top of a mixed four-stranded beta-sheet in a (-1, +3x, -1) topology [Wikström *et al.*, 1994]. The first two β-strands are also of the same length in both proteins, containing nine amino acids residues apiece. Figure 1.25 presents an adapted three-dimensional ribbon comparison of PpL-V_L and SpG-C_H1 interactions.

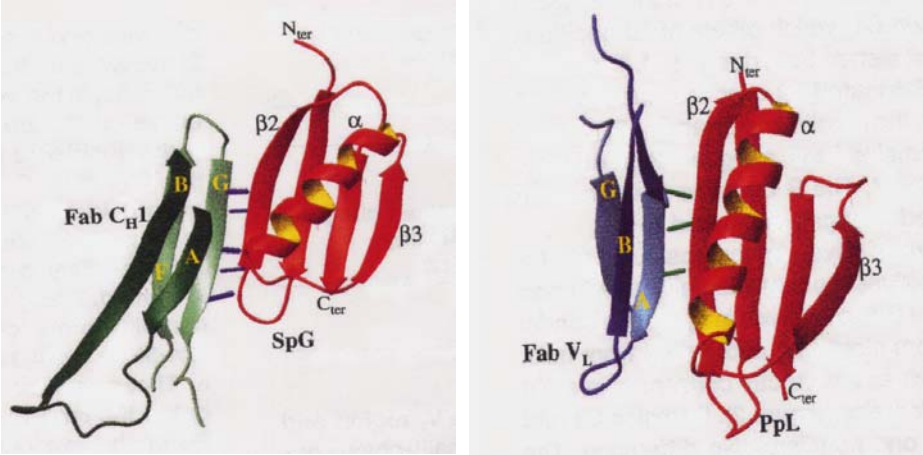


Figure 1.25 – A ribbon comparison of PpL-V_L and SpG-C_H1 interactions

The interaction between proteins G and L and a Fab fragment is accomplished by an alignment of β-strands. Strand β2 of both domains interacts with, respectively, a Fab V_L (blue) or a Fab C_H1 (green) β strand, and the similarities of their interactions are clearly demonstrated. In both cases β2 extends a β-sheet of the Fab through a β-zipper interaction. This interaction involves five main-chain hydrogen bonds in the SpG-C_H1 interface, but only three hydrogen bonds in the PpL-V_L first interface, and is represented by coloured lines [Graille et al., 2001].

In spite of the overall resemblance of the Ig-binding domains of SpG and PpL, several differences are found when their three-dimensional structures are compared. One main structural difference is the orientation of the α-helices, where the α-helix in PpL runs almost parallel to the β-sheet at an angle of Ω ~-10°, whereas in SpG it runs diagonally across the β-sheet with an angle of Ω ~-40°. This arises from a difference in the length of the loop between the α-helix and the third β-strand, which is one residue shorter in PpL. This is the region of SpG that is responsible for Fc binding, so therefore, the shorter loop in PpL may help to explain the inability of PpL to bind Fc [Graille *et al*., 2001].

Other differences pertain not only to the helix orientation, but also to differences in the turns between the β-strands, as well as the lengths of strands β3 and β4. The turns connecting the first two β-strands differ in that in SpG, strands β1 and β2 are connected by a type I β-turn, while in PpL the linking segment is a five residue turn. The turns between strands β3 and β4 in SpG is a six residue turn as opposed to a distorted type I' β-turn in PpL [Wikström *et al.*, 1994]. Also in PpL the length of strands β3 and β4 is eight amino acid residues, whereas in SpG the corresponding strands are only five residues long.

Both proteins L and G use their second β-strands when interacting with the framework region on the V_L domain of kappa light chain and the C_H1 region in

IgG Fab respectively [Wikström *et al*., 1996], thus burying equivalent surface areas upon binding, and this also shares similar features. In both interactions the second β-strand, extends a β-sheet of the Fab through a β-zipper interaction, involving five main-chain hydrogen bonds in the SpG-C_H1 interface [Derrick and Wigley, 1994], but only three hydrogen bonds are involved in the PpL-V_L first interface [Graille *et al*., 2001]. However, although this second β-strand is important for Fab binding with both proteins G and L, different parts of the strand participate in the interaction. In the contact between PpL and the Ig light chain variable domain, residues from almost the whole strand are affected, whereas the first turn and the two N-terminal thirds of the second β-strand are involved in the binding of SpG to C_H1.

When the primary and tertiary structures of the Ig-binding domains of L and G are compared, the third β-strands are the most homologous. The involvement of this strand in the interaction between L and kappa chain is highly unlikely since PpL cannot inhibit the binding of SpG to Fc. The least homologous and surface-exposed regions in the comparison are the second β-strand, parts of the α-helix and the loop following the helix [Wikström *et al*., 1994].

Streptococci and staphylococci bacteria strains are virulent and cause a range of clinical infections in humans, but their respective Ig-binding proteins cannot be unambiguously assigned a role in the virulence of the bacteria. Peptostreptococci is a member of the normal flora on most body surfaces, and are even found in the same habitats as streptococci, such as the human intestinal and genital tracts, and compared to staphylococci and streptococci are much less virulent. However, strains of *P. magnus* that have the PpL gene are more frequently connected with clinical infections suggesting that PpL is a virulence determinant. This is in contrast to SpG, even though it is expressed in group C and G streptococci strains isolated from patients with severe clinical infections.

The reason why PpL-expressing bacteria appear more virulent is not fully understood, although it is thought that this could be that Ig bound to PpL or SpG at the bacterial surface is exposed differently. An antibody bound to PpL through light chain interactions can also bind to Fc receptors on host cells, whereas the Fc region of IgG antibodies bound to SpG is blocked. Furthermore, PpL-expressing Peptostreptococci strains can bind all classes of Ig, and are also capable of inducing histamine release from human basophils and mast cells by binding to and cross-linking surface IgE [Patella *et al*., 1990]. Thus, PpL and SpG seem to

have been assigned functionally different roles, regardless of their similar structural organisation.

1.3 - Applications of cell surface Ig-binding proteins

Of the cell surface Ig-binding proteins SpA was the method of choice for many preparative and analytical purposes in immunology. It has been employed in the purification and analysis of human antibodies on Sepharose affinity columns, and for the fractionation of subclasses, which are difficult to separate by other means. SpA is also utilised in the detection of immunoglobulin in an ELISA, used as a solid phase adsorbant for isolating antigen-antibody complexes, membrane antigens and receptors. Finally, SpA has proven useful for the study of antigens and receptors on the surface of intact cells, and for the detection of antibody-secreting cells [Goding, 1978].

The ability of SpA's to bind immune complexes has promoted its use of SpA as a therapeutic agent in cancer and transplants. 20-30% of patients with organ transplants reject the new tissue due to the production of anti-histocompatability (anti-HLA) antibodies [Palmer *et al*., 1989]. Extracorporeal immunoadsorption with immobilised SpA can be used to remove these anti-HLA antibodies, since the majority of these are IgG, prior to the operation. This technique has been used successfully in kidney transplants without side effects, and is more effective than traditional plasma exchange methods [Palmer *et al*., 1989, Barocci *et al*., 1993]. Binding SpA to immune complexes has also been used in extracorporeal shunt therapy for cancer therapy [Bansal *et al*., 1978.

SpG has a much broader specificity for IgG, and its resistance to denaturation either by temperature or urea, give this protein immense potential as an immunological tool. In antibody purification, SpG offers a means of separating IgG from other immunoglobulins and could be used for a wider selection of mammalian antibodies than SpA. Affinity for a wider range of species and subclasses would put less constraint on the IgG that could be used in ELISA where immunoglobulin-binding proteins can be used in place of a secondary antibody. Its greater stability should also make this protein less susceptible to degradation by host proteases, making it a better choice than SpA, as a fusion protein in expression vectors [Nilsson *et al*., 1992].

PpL represents a new tool for the binding and detection of immunoglobulins. The advantage of using PpL is that it possesses the ability to interact with all classes

of immunoglobulin, which possess Ig kappa light chains. Thus, the specific interaction between PpL and the immune system could be of significance to the understanding of the host-parasite relationship. PpL has potential use for the isolation of antigen-binding V_L domains and scFv fragments prepared from monoclonal antibodies or produced in recombinant DNA systems [Nilson *et al*., 1992]. The use of recombinant antibodies in the form of scFv fragments is becoming increasingly common, and it has been shown that both PpL and SpA may be valuable tools for the purification of such recombinant antibodies from crude cell culture media [Åkerström *et al*., 1994]. In addition they can be used as detectors of antibodies in immunoblotting and immunocytochemistry as they do not impair antigen-binding to the antibody fragments.

Protein H also has considerable potential as an immunological tool. Its specificity for human IgG enables purification of human monoclonal IgG antibodies from a culture medium [Åkesson *et al*., 1990]. Its affinity for IgG is lower than that of SpA or SpG, but is sufficiently high for immunoglobulin purification. The benefit of this lower affinity is, however, advantageous as the elution of human IgG from immobilised Protein H requires less denaturing conditions.

Thus, it can be noted that, at present, no particular Ig-binding protein is without limitations.

1.4 - Applications of recombinant antibody fragments

The unique ability of antibodies to specifically recognise and bind with high affinity to virtually any type of antigen, such as microorganisms, larger parasites, viruses and bacterial toxins, rendering them harmless has made them interesting molecules for medical and scientific research. Since monoclonal antibody (mAb) technology was developed in 1975 [Köhler and Milstein, 1975], the isolation and production of mAbs against protein, carbohydrate and nucleic acids antigens has resulted in a rapid development of the use of antibodies and their fragments in diagnostics, human therapeutics and as fundamental research tools [Porter *et al*., 1988]. Applications outside research and medicine include consumer uses such as the use of antibodies in shampoos to prevent dandruff [Dolk *et al*., 2005] or in toothpaste to prevent against tooth decay [Joosten *et al*., 2003].

However, using production systems based on mammalian expression is expensive, producing antibodies in bulk amounts is difficult and the low stability and solubility of some antibodies under specific (harsh) conditions presents major

problems. Therefore, the use of recombinant human monoclonal antibodies is becoming increasingly popular due to their advantages compared to conventional monoclonal or polyclonal antibodies. The production of recombinant antibodies is much faster compared to conventional antibody production and they can be generated against an enormous number of antigens. In contrast, in the conventional method, many antigens prove to be non-immunogenic or extremely toxic, and therefore cannot be used to generate antibodies in animals. Moreover, increasing the affinity and specifity of recombinant antibodies is very simple and relatively fast, and large numbers of different antibodies against a specific antigen can be generated in one selection procedure [Hudson, 1999].

Recombinant antibodies and their fragments now represent over 30% [Holliger and Hudson, 2005] of all biological proteins undergoing clinical trials for diagnosis and therapy. Over recent years, important advances have been made in the design, selection and production of new types of engineered antibodies. Innovative selection methods have enabled the isolation of high-affinity cancer-targeting and antiviral antibodies, the latter capable of redirecting viruses for gene therapy applications. In other strategies for cancer diagnosis and immunotherapy, recombinant antibody fragments have been fused to radioisotopes, drugs, toxins, enzymes and biosensor surfaces, effectively enhancing the human immune response in anti-cancer vaccines and T cell recruitment strategies [Hudson, 1999]. The growth of therapeutic antibodies is therefore dependent on the development of new approaches for faster antibody development, improved methods of selection and optimisation, alternatives for production, as well as the evaluation for novel applications [Glasgow *et al*., 2006].

In the future, such recombinant antibodies may form the basis for the treatment of autoimmune diseases, but there remains a long way to go.

1.5 - Aims and objectives

The aim of this research is to study a single Ig-binding domain construct of SpG, with respect to its stability and binding interactions with both Fc and Fab fragments of IgG molecules. The 208 bp Ig-binding domain gene was based on the sequence for SpG-C_2, which is a hybrid of the C-terminus of the C_1 domain and the N-terminus of the C_3 domain. This domain was chosen due to its combined characteristics of its counterparts. Despite being almost identical in sequence, C1 has a lower affinity for IgG, but is more stable and has a higher

melting temperature than C3.

The aim of this research is to study the role of the individual interfacial interactions between Fc and SpG, which consist of a network of hydrogen bonds and salt links, involving mainly charged and polar residues. An additional aim of this research is to study the role of various amino acids that form the major contact regions in the binding of SpG with Fab, as very little is known of the binding properties and characteristics of this interaction. To facilitate these studies a program of site-directed mutagenesis will be carried out.

Thus, it may prove possible to generate a library of constructs that offer a range of binding affinities to enable different biotechnological uses. For example, to remove Ig or Ig complexes from sera, a high affinity SpG ligand would be beneficial. Alternatively, an affinity ligand with a lower affinity would offer advantages for the elution of the Ig from the column.

Finally, the stability of the constructs is to be studied using guanidine hydrochloride denaturation experiments, using both fluorimetry and circular dichroism (CD). Since the product must offer, not only the appropriate affinity, but also be thermodynamically stable. The binding properties of these mutant constructs will then be characterised thoroughly to obtain affinity constants by a combination of fluorescence titration and isothermal titration calorimetry (ITC), and kinetic measurements of real-time binding using stopped-flow fluorescence spectroscopy.

CHAPTER 2.0

MATERIALS AND METHODS

2.1 - General chemicals and specific materials

To enable the interaction of SpG with Fc and Fab fragments to be studied, a selection of microbiological, DNA and protein analysis techniques were employed using a wide range of general chemicals and specific materials. General chemicals were obtained from Fisher Scientific, Loughborough, UK and Sigma Chemical Company Ltd, Poole, Dorset. Specialised chemicals, kits and services were purchased from the following companies:

- Agarose: Flowgen, Rockland, USA.
- Agar No.1: Lab M, Bury, Lancaster, UK.
- Dithiothreitol; DNase 1: Roche Molecular Biochemicals, East Sussex, UK.
- Mis-match oligonucleotide primers for site-directed mutagenesis; DNA Sequencing: MWG-Biotech, Milton Keynes, UK.
- Protogel (30%acrylamide/0.8%bis-acrylamide): National Diagnostics, Hessle, Hull, UK.
- *Pfu* DNA polymerase; Restriction and modification enzymes with respective buffers: Promega, Chilworth, Southampton, UK.
- QIAprep Spin miniprep Kit; QIAquick Gel Extraction Kit; QIAquick Nucleotide Removal Kit; QIAquick PCR Purification Kit: Qiagen Ltd., Crawley, West Susssex, UK.
- Tryptone; Yeast Extract: Melford Laboratories, Ipswich, Suffolk, UK.
- 100bp DNA ladder: New England Biolabs, Hitchin, Herts, UK.
- Human Fc fragments were kindly provided by Prof. M. Glennie, Tenovus Labs, Southampton General Hospital, UK.
- Plasmid vector containing Fab gene and purified Fab protein were kindly provided by Dr. A.G. Popplewell, UCB-Celltech, Slough, UK.

2.2 - Preparation of buffers and common reagents

2.2.1 - Media

2.2.1.1 - Luria-Broth

Routine growth of bacterial cultures was achieved using Luria-Broth (LB) (tryptone 10 g/litre, yeast extract 5 g/litre, NaCl 5 g/litre) and supplemented with antibiotics, ampicillin (50 μg/ml) or tetracycline (7.5 μg/ml). LB was solidified for plates by the addition of 1.5% w/v agar no.1.

2.2.2 - 6X Gel Loading Buffer

Gel loading buffer (0.25% w/v bromophenol blue, 30% v/v glycerol) was used to increase the density of the DNA sample to ensure that the DNA dropped evenly into the wells of agarose gels (section 2.4.3).

2.2.3 - 1X and 10X TBE Buffer

1X TBE buffer was diluted from a 10X TBE stock (108 g Tris, 55 g Boric acid, 9.3 g Na_2EDTA/litre distilled water) for the preparation of agarose and as a running buffer during agarose gel electrophoresis (section 2.4.3).

2.2.4 - TE Buffer

Sterile TE buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA) was used during the preparation of phenol/chloroform extraction of DNA (section 2.4.2.1).

2.2.5 - Sonication Buffer

Sonication buffer (20 mM Tris-HCl pH 8.0, 0.1 mM EGTA, 0.1 mM EGTA, 0.5% v/v Triton X100) was used to resuspend pelleted cells, which had been harvested from overnight cultures, grown to express recombinant SpG (section 2.5.1). The cell paste was used immediately or stored frozen at -20 °C.

2.2.6 – SpG Purification Buffers

Buffer A (20 mM Tris-HCl pH 8.0) was used to pre-equilibrate a Q-Sepharose anion exchange column, and as part of a linear gradient with elution buffer B (20 mM Tris-HCl pH 8.0, 1 M NaCl) (section 2.5.1).

2.2.7- Buffer for Extracting Fab from Periplasm

Periplasmic extraction buffer (20 mM Tris-HCl pH 8.0, 10 mM EDTA, 20% sucrose, 1 ml broad range protease inhibitors) was used to resuspend pelleted cells that had been harvested from a 32 hour fermentation culture, grown to

express recombinant Fab, and to extract the Fab protein from the resulting cell suspension (section 2.5.2).

2.2.8 - Fab Purification Buffers

Buffer A (100 mM Tris pH 8.0, 0.5 M NaCl) was used to pre-equilibrate a 50 ml PpL column bound (5 mg/ml) to the NHS-Sepharose column medium, and as part of a linear gradient with elution buffer B (50 mM citrate, pH 2.2) (section 2.5.2).

2.2.9 - SDS High-Resolution Polyacrylamide Gel

A 6 ml high-resolution 17.5% acrylamide separating gel (3.5 ml 30% w/v protogel, 2.4 ml 3.6 M Tris-HCl pH 9.3, 0.06 ml 10% w/v SDS, 0.03 ml 25% w/v APS, 0.010 ml TEMED) was used for SDS-PAGE electrophoresis (section 2.5.3). The gel was completed with a 4 ml stacking gel (0.600 ml protogel, 1.6 ml 3.6M Tris-HCl pH 8.85, 1.7 ml distilled water, 0.04 ml 10% SDS, 0.02 ml 25% APS, 0.005 ml TEMED).

2.2.10 - 5X SDS High-Resolution Running Buffer

1X SDS running buffer was diluted from a 5X SDS running buffer stock (30.25 g Tris, 142.63 g glycine, 10 g SDS/litre distilled water) and used during SDS gel electrophoresis (section 2.5.3).

2.2.11 - 2X Reducing Sample Buffer

Sample buffer (2 ml glycerol, 4 ml 10% SDS, 5 ml distilled water, 4 ml 2-Mercaptoethanol, 15 mg bromophenol blue) was used to increase the density of the SpG samples to ensure that the protein dropped evenly into the wells of SDS gels during electrophoresis (section 2.5.3).

2.2.12 - Non-Reducing Sample Buffer

Non-reducing sample buffer (2 ml glycerol, 4 ml 10% SDS, 5 ml distilled water, 15 mg bromophenol blue) was used to increase the density of the protein sample to ensure that the Fab protein samples dropped evenly into the wells of SDS gels during electrophoresis (section 2.5.3).

2.2.13 - Coomassie Brilliant Blue Stain

Coomassie brilliant blue stain (454 ml methanol, 96 ml acetic acid, 454 ml distilled water, 2.5 g Coomassie blue/litre distilled water) was used to stain SDS polyacrylamide gels (section 2.5.3).

2.2.14 – Gel Destain Solution

Destain (180 ml methanol, 100 ml acetic acid, 1720 ml distilled water) was used to destain SDS polyacrylamide gels (section 2.5.3).

2.2.15 – 20 mM Potassium Phosphate Buffer

20 mM potassium phosphate (75 ml 1M K_2HPO_4, 5 ml 1 M KH_2PO_4/920 ml distilled water), with varying pH, was used for dialysis (4 litres 20 mM KPO_4 buffer) and the preparation of protein for Fc and Fab binding characterisation studies (section 2.6 and 2.7).

2.2.16 – 20 mM Triple Buffer System

A triple buffer system (75 ml K_2HPO_4, 5 ml KH_2PO_4, 80 ml 1M Tris-HCl, 80 ml 1 M sodium acetate/3760 ml distilled water), ranging from pH 4.0 to pH 9.0, was used for dialysis and the preparation of protein used for pre-equilibrium studies (section 2.6.3), to determine the effect of pH on Fc and Fab binding.

2.3 - Microbiological Methods

2.3.1 - Sterilisation

All media, stock solutions, pipette tips and microfuge tubes were sterilised by downstream displacement autoclaving at 121 °C, 15 lb/sq.inch, for 20 minutes. Glassware was sterilised by heating to 160 °C for 2-3 hours. Heat labile stock solutions such as antibiotics, IPTG and salt supplements were sterilised through 0.45 μm Millipore disposable filters.

2.3.2 – Bacterial Strains

2.3.2.1 - *E.coli* JM103

JM103 cells were used for the expression and storage of SpG and Fab [Messing *et al*., 1981].

Genotype:-Δ(lacproAB),supE,thi,rsplL.endA,SbcB15,[F',traD36,proA,lac$I^qZ\Delta M15$]

2.3.2.2 - Antibiotic resistance

The genes for the SpG and Fab proteins used in this study were in vectors that confer different bacterial antibiotic resistance to host bacteria. The SpG expression vector (AP10-2 modified *pQE*-30) confers ampicillin resistance, and thus 50 μg/ml ampicillin was present in all SpG bacterial cultures. The Fab vector (pTTOD) confers tetracycline resistance, and 7.5 μg/ml tetracycline was present in all Fab bacterial cultures.

2.3.2.3 - Storage of bacterial strains

Bacterial strains containing plasmid DNA and antibiotic resistant gene were stored as streaks on LB-plates (section 2.2.1.1) supplemented with the appropriate antibiotic. Non-transformed cells were stored on LB-plates with no antibiotic present. These plates are viable stored at 4 °C for up to 4 weeks and used to produce overnight cultures.

For long-term storage of strains, 5 ml cultures were grown up in LB (section 2.2.1.1) to mid-log phase, at which point the cells were harvested and resuspended in 800 μl of LB and 200 μl sterile glycerol. These can be stored at -70 °C for approximately two years.

2.3.2.4 - Growth of bacterial strains

Individual colonies of transformed ligations or streaks of transformed plasmid DNA (section 2.3.3) were used to inoculate 10 ml LB (section 2.2.1.1) containing the appropriate antibiotic. These cultures were incubated at 37 °C overnight with shaking at 200 rpm, and used for the expression of protein (section 2.5.1) or DNA preparation (section 2.4.1).

2.3.3 - Transformation of E.coli

2.3.3.1 - Preparation of competent cells

Bacterial cells were made competent for the uptake of DNA by the method of Cohen *et al.*, [1972]. This technique induces a transient state of competence in the recipient bacteria, during which they are able to take up DNA derived from a variety of sources, and then used to confer bacterial antibiotic resistance.

A 10 ml culture grown overnight of the host strain *E.coli*, was diluted 1:1000 in LB and grown to mid-log (OD_{600nm} ~0.5) at 37 °C with shaking. The cells were harvested by centrifugation for 5 minutes in a bench top centrifuge, and resuspended in 5 ml ice-cold 50 mM calcium chloride. The cells were incubated on ice for a further 30 minutes, and then centrifuged as before, and resuspended in a final volume of 1 ml ice-cold 50 mM calcium chloride. The competent cells were kept on ice and used within 24 hours.

2.3.3.2 - Transformation of plasmid DNA

Aliquots of competent cells, 200 μl, were mixed with 20-200 ng plasmid DNA, or whole ligation reactions, and incubated for 45 minutes on ice. The cells were heat-shocked at 42 °C for 90 seconds to facilitate the uptake of DNA, and incubated on ice for a further 5 minutes to allow recovery. 0.8 ml of LB was

added to the cells and incubated at 37 °C for 1 hour to enable some initial expression of the antibiotic gene product conferring antibiotic resistance to the cells. Successful transformants were selected by plating the cells onto the relevant selective media, and incubated overnight at 37 °C.

2.4 - Molecular Biology Techniques

All molecular biology techniques were performed on the synthetic gene encoding a single Ig-binding domain of SpG, SpG_{C2}-1, as described in Maniatis *et al.*, [1982].

2.4.1 - DNA preparation

This method is based upon the alkaline lysis of the bacterial cells followed by adsorption of the DNA onto silica in the presence of high salt [Birnboim and Doly, 1979]. The bacteria is harvested and lysed by sodium hydroxide and SDS under alkaline conditions in the presence of RNase A. The SDS solubilises the phospholipid and protein components of the cell membrane, leading to lysis and release of the cell contents, whereas the alkaline conditions denature the chromosomal and plasmid DNA. Under optimal lysis time this procedure allows the maximum release of plasmid DNA without the release of chromosomal DNA. The lysate is subsequently neutralised and adjusted to high salt binding conditions, ready for purification on a silica-gel membrane. The high salt concentration causes the denatured proteins, chromosomal DNA, cellular debris and SDS to precipitate, while the smaller plasmid DNA renatures and stays in solution. The separation of the plasmid is based on the precipitation of the cell wall-bound chromosomal DNA with insoluble complexes that contain salt, detergent and protein. Washing and eluting the plasmid DNA ensures that endonucleases are efficiently removed, which is essential with *endA*+ strains, such as JM103, since they can reduce the quality of DNA. These cells also produce large amounts of carbohydrates that are released during lysis and can inhibit enzyme activities if they are not completely removed [Vogelstein and Gillespie, 1979].

Double-stranded plasmid DNA was prepared from 10 ml cultures grown in Luria broth for 16 hours at 37 °C, shaking at 200 rpm. Each 10ml culture had been inoculated with a single colony and contained the appropriate antibiotic selection. The DNA was extracted and purified using a QIAprep spin miniprep kit, by following the manufacturer's recommendations.

2.4.2 - DNA purification

One of the most popular methods to purify DNA is that of Birnboim and Doly, [1979], as described and utilised by the QIAprep spin miniprep kit (section 2.4.1). However, subsequent purification after various enzymatic reactions is necessary before further DNA applications can be performed.

2.4.2.1 - Phenol/chloroform extraction of DNA

The standard procedure used in molecular cloning to purify nucleic acids is phenol/chloroform extraction. The removal of contaminating proteins is achieved by extracting aqueous solutions of nucleic acids with phenol: chloroform, and chloroform [Sambrook *et al*, 1989]. Such extractions are utilised whenever it is necessary to inactivate and remove enzymes that are used in one step of a cloning operation before proceeding to the next. The phenol: chloroform solution denatures any proteins present in the DNA sample, and this de-proteinisation is more efficient as two organic solvents are used instead of one. The nucleic acids partition into the aqueous phase and the denatured proteins form the organic phase. Back extraction of the organic phase and the interface is also performed to achieve optimum recovery by the addition of TE buffer (section 2.2.4). This is repeated until there are no proteins visible at the interface of the organic and aqueous phases, which is seen as a white interface. A subsequent extraction with chloroform removes any lingering traces of phenol from the preparation, and the nucleic acids are recovered by precipitation with ethanol.

Contaminating proteins were removed from DNA solutions by phenol: chloroform extraction. 100 μl phenol: chloroform solution was added to 100 μl DNA sample in a sterile microfuge tube. The mixture was vortexed for 30 seconds, allowed to stand for 1 minute before another 30 second vortex. The tube was then centrifuged for 4 minutes, and the upper aqueous layer was separated from the lower organic phase and transferred to a fresh tube. The procedure was then repeated by back extracting using TE buffer in place of the phenol: chloroform solution. The two upper layers were pooled and an equal volume of chloroform was added to remove any remaining traces of phenol. Nucleic acids were recovered by precipitation with 2½ times the initial volume of absolute ethanol and 1/10th volume of 3 M sodium acetate pH 5.5 at -70 °C, for 20 minutes. This was followed by centrifugation for 10 minutes. A 70% ethanol wash was then performed to remove any residual salts, followed by another centrifugation for 10 minutes. The resulting pellet was air dried and resuspended with sterile analar water.

2.4.2.2 - Nucleotide removal

A QIAquick nucleotide removal kit was used to purify oligonucleotides and DNA fragments up to 10 kilobases (kb) from enzymatic reactions such as dephosphorylation, ligations and restrictions digests [Birnboim and Doly, 1979]. The procedure is based upon the selective binding properties of the silica-gel membrane utilised in the spin columns. The special buffers provided are optimised for the efficient recovery of DNA and the removal of contaminants. The adsorption of oligonucleotides greater than 17 bases and DNA fragments up to 10 kb to the silica-gel membrane is promoted by the binding buffer and occurred in the presence of a high concentration of chatrophic salts, which modify the structure of water [Volgelstein and Gillespie, 1979]. During the DNA adsorption step, unwanted primers and impurities, such as salts and enzymes, do not bind and flow through the membrane. An ethanol-containing buffer quantitatively washes salts away, and any residual buffer that may interfere with subsequent enzymatic reactions, is removed by an additional centrifugation.

DNA fragments and oligonucleotides were purified by removing unwanted primers <10 bases, enzymes and contaminating salts and incorporated nucleotides using the QIAquick nucleotide removal kit, by following the manufacturer's recommendations.

2.4.2.3 - Agarose gel extraction

A QIAquick gel extraction kit was used to purify extracted DNA fragments between 70 base pairs (bp) and 10 kb from standard or low-melting agarose gels in 1X TBE buffer (section 2.2.3). This procedure is based on the same principles as illustrated by the QIAquick nucleotide removal kit (section 2.4.4.2) [Vogelstein and Gillespie, 1979].

DNA fragments ranging from 70 bp to 10 kb were extracted from agarose gels in TBE buffer and purified by removing primers, enzymes, such as restriction and phosphatases, and salts using the QIAquick gel extraction kit by following the manufacturer's recommendations.

2.4.3 - Agarose gel electrophoresis

Agarose gel electrophoresis is a technique used to separate, identify and purify DNA fragments [Sharp *et al.*, 1973]. It is simple and rapid to perform, and capable of resolving fragments that cannot be separated adequately by other procedures, such as density gradient centrifugation. Agarose, which is extracted

from seaweed, is a linear polymer of galactopyranose derivatives. Agarose gels can be cast by melting the agarose powder in the presence of the desired buffer, until a clear, transparent solution is achieved, which is then poured into a mould and allowed to harden. Upon hardening, the agarose forms a matrix, the density of which varies depending on the concentration of agarose. When an electric field is applied across the gel, DNA, which is negatively charged at a neutral pH, migrates towards the anode. The linear DNA fragment of any given size will migrate at different rates through the matrix of the agarose gel. Gel loading buffer (section 2.2.2) is mixed with the DNA to increase the density of the sample to ensure that the DNA drops evenly into the well. The added colour from the bromophenol blue makes the loading process easier, and contains dyes that move towards the anode at predictable rates.

The location of both single- and double-stranded DNA within the gel can be determined directly by staining with low concentrations of a fluorescent intercalating ethidium bromide dye, and detected by the direct examination of the gel under ultraviolet light. Ethidium bromide contains a planar group that intercalates between the stacked bases of DNA. The fixed position of this group close to the bases causes the dye bound to the DNA to display an increased fluorescent yield compared to that of the dye in free solution [Sharp *et al.*, 1973].

Identification, sizing and purification of DNA fragments were performed using submerged horizontal agarose gel electrophoresis. Agarose 0.5%-2.0% w/v in 1X TBE buffer (section 2.2.3) was heated until all agarose had dissolved. Upon cooling to approximately 45 °C, 0.1 μg/ml ethidium bromide was added. The gel solution was poured onto the gel plate with a comb, and allowed to set. The comb was removed and the gel placed into a tank containing 1X TBE buffer. The samples were prepared using 6x gel loading buffer (section 2.2.2) and approximately 2-10 μl was loaded into the wells. The gel was run at 100 volts and the resulting separated DNA fragments were visualised under UV light. Images of the gels were obtained by using the Syngene Gene Genius gel documentation system.

2.4.4 - Mutagenesis

Mutagenesis is the alteration of a DNA sequence by specific chemical or enzymatic methods. Cloned genes allow mutagenesis to be much more precise by replacing specific bases in a gene, or specific amino acids at selected sites in the protein, a procedure known as site-directed mutagenesis.

2.4.4.1 - Oligonucleotide mutagenic primers

Mis-match mutagenic oligonucleotide and flanking *pQE*-30 vector primers were produced and purified by MWG-Biotech, and dissolved in sterile distilled water to obtain a final 100 pmol/µl concentration.

2.4.4.2 - PCR-based site-directed mutagenesis

Polymerase chain reaction (PCR) amplification is used to amplify a segment of DNA that lies between two regions of a known sequence. The two oligonucleotides are used as primers for a series of synthetic reactions that are catalysed by a DNA polymerase [Saiki *et al*., 1988]. These primers have different complementary sequences that lie on the opposite strands of the template DNA, and flank the segment of DNA to be amplified. The template DNA is first denatured by heating, at approximately 95 °C, in the presence of a large molar excess of each of the two primers and four dNTPs. The reaction is cooled to a temperature that allows the primers to anneal to their target sequences, between 35 °C and 55 °C, after which the annealed primers are extended with DNA polymerase at approximately 72 °C. The cycle of denaturation, annealing and DNA synthesis is repeated many times. The products of one cycle of amplification serves as a template for the next, allowing each successive cycle to double the amount of desired DNA product [Erlich *et al*., 1988]. The product of this reaction is a segment of double-stranded DNA, which is defined by the 5' termini of the oligonucleotides primers, and the distance between the primers defines its length.

PCR-based site-directed mutagenesis was employed for introducing point mutations, using the principles of splicing by overlap extension (SOE-ing) [Horton *et al*., 1989]. A modified two-sided SOE-ing method was used for the production of all SpG mutations, and requires three separate PCR reactions to produce the final mutant gene product. PCR reactions A and B were set up according to the reaction mix shown in table 2.1, with one reaction mix prepared per sample. The two-sided SOE-ing mutagenesis reactions were prepared on ice, and each of the components was added in order whilst mixing gently.

The PCR amplification products A and B were analysed by 2% agarose gel electrophoresis (section 2.4.3), and fragments corresponding to the correct size were excised and purified using a QIAquick gel extraction kit (section 2.4.2.3). PCR products A and B were then used to produce the final mutagenic product PCR C according to the reaction mix shown in table 2.2.

Reagent	Volume (µl)
25 mM $MgCl_2$	1
10x *Pfu* DNA polymerase buffer	2.5
Template plasmid DNA	1
Mutagenic primer (100 pmol/µl)	0.5
Flanking pQE-30 primer (100 pmol/µl)	0.5
2 mM dNTPs	2.5
DMSO	2
Sterile distilled water	14
Pfu DNA polymerase enzyme (3u/µl)	1
TOTAL volume	**25**

Table 2.1 - Reaction mix for PCR A and B

Reagent	Volume (µl)
25 mM $MgCl_2$	1
10x *Pfu* DNA polymerase buffer	2.5
Product from PCR A	1
Product from PCR B	1
Flanking pQE-30 primer 2 (100 pmol/µl)	0.5
Flanking pQE-30 primer 1 (100 pmol/µl)	0.5
2 mM dNTPs	2.5
DMSO	2
Sterile distilled water	7
Pfu DNA polymerase enzyme (3u/µl)	1
TOTAL volume	**25**

Table 2.2 - Reaction mix for PCR C

For both sets of PCR reactions, *Pfu* DNA polymerase was added to the reaction mix last to ensure that the 3'-5' exonuclease proofreading ability of the enzyme did not degrade the single-stranded primers. All reactions were also overlaid with a drop of mineral oil to prevent the mixture evaporating at the high temperatures, before being amplified in a TechGene thermal cycler. Each PCR was performed using optimised thermal cycling parameters shown in table 2.3.

Temperature (°C)	Time (minutes)	
95	1	
95	0.5	repeated
50	0.5	20
72	1	cycles
72	5	
4	∞	

Table 2.3 - Thermal cycling parameters for two-sided SOE-ing reactions

2.4.5 - Restriction digestion of DNA

Restriction endonucleases are used to manipulate DNA by binding specifically to double-stranded DNA, and cleaving it at specific sites within or adjacent to a particular sequence known as the recognition sequence. The cuts in the two strands are frequently staggered, which create single-stranded overhangs or sticky ends, and allow the two strands to link up by base pairing to form recombinant DNA *in vitro* [Nathans and Smith, 1975].

Bam HI and *Hin*d III restriction digests were performed on plasmid DNA, pQE-30 vector, and the PCR gene products excised and purified from agarose gel. The mixture for a typical restriction digest is outlined in table 2.4.

Reagent	Volume (µl)
DNA	25
10x Restriction enzyme buffer E	4
Restriction enzyme *Bam* HI (10 u/µl)	2
Restriction enzyme *Hin*d III (10 u/µl)	2
Sterile distilled water	7
TOTAL VOLUME	**40**

Table 2.4 - Reaction mix for restriction digest

The mixture was incubated at 37 °C for eight hours, and the resulting gene fragments were excised and purified using a QIAquick gel extraction kit (section 2.4.2.3).

2.4.6 - Dephosphorylation of vector

Following the restriction digestion of vector DNA, the 5' phosphate group was removed from linear DNA molecules to reduce self-ligation in a process known as dephosphorylation. Phosphatase buffer was added along with 2 units of calf intestinal alkaline phosphatase (CIAP), a dimeric glycoprotein, and incubated at 37 °C for 40 minutes. CIAP activity was then destroyed by incubation at 65 °C for 20 minutes. The vector DNA was then purified using the methods of either phenol/chloroform extraction followed by ethanol precipitation (section 2.4.2.1) or by using the QIAquick nucleotide removal kit (section 2.4.2.2).

2.4.7 - Ligation of gene fragment into vector

Ligation of a segment of foreign DNA to a linearised plasmid vector involves the formation of new bonds between phosphate residues located at the 5'-termini of double-stranded DNA and adjacent 3'-hydroxyl residues [Ullrich *et al.*, 1977]. The formation of phosphodiester bonds between adjacent 5'-phosphate and 3'-hydroxyl residues is catalysed *in vitro* by bacteriophage T4 DNA ligase [Zimmerman and Pheiffer, 1983].

"Sticky-end" ligations were performed using 1 unit/µl T4 DNA ligase and 20-100 ng vector DNA with varying amounts of insert DNA, as shown in table 2.5. The mixture transferred to a 2 litre water bath at 25 °C, and then incubated at 4 °C

for 16 hours. The entire ligation product was then transformed into *E.coli* JM103 cells (section 2.3.3) and plated onto LB-agar plates supplemented with the appropriate antibiotic.

Reagent	Ligation control	1:1 ratio (v:v)	1:3 ratio (v:v)	1:7 ratio (v:v)
Vector	1	4	1	1
Gene insert	0	4	3	7
10x T4 DNA ligase buffer	1	1	1	1
T4 DNA ligase (1 u/µl)	0.5	0.5	0.5	0.5
Sterile distilled water	7.5	0.5	4.5	0.5
TOTAL VOLUME	**10**	**10**	**10**	**10**

Table 2.5 - Reaction mix for ligation of gene fragment into vector

2.5 - Protein Analysis Techniques

2.5.1 - Production of SpG

2.5.1.1 - Expression of recombinant Ig-binding proteins

One colony was grown overnight in 10 ml LB (section 2.2.1.1) containing 50 µg/ml ampicillin at 37 °C, and 2 ml was used to inoculate a 2 litre conical flask containing 700 ml LB and 50 µg/ml ampicillin. Six flasks were set up and incubated, with shaking at 37 °C for 4-5 hours. The *E.coli* carrying the desired recombinant SpGene in the pQE-30 expression vector were induced by the addition of 0.6 mM isopropyl-β-D-thiogalactoside (IPTG), a non-hydrolysable analogue whose structure is shown in figure 2.1, into the culture medium OD_{600nm} 0.600–0.900.

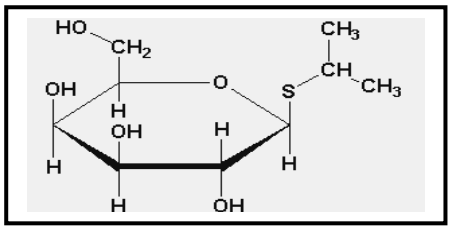


Figure 2.1 – A diagrammatic representation of the structure of IPTG

Isopropyl-β-D-thiogalactoside (IPTG) is a synthetic artificial inducer of the lac operon. The lac operon in E.coli is an example of a dual control system enabling gene transcription to be switched on and off by gene regulation proteins. IPTG is used in contrast to the lac operon's natural inducer allolactose (a minor product of β-galactosidase action), because IPTG cannot be hydrolysed and broken down in the cell by β-galactosidase [Bell and Lewis, 2000].

This analogue exhibits control over expression and allows cell growth to reach a maximum [Galloway *et al*., 2003] before cell resources are converted to target gene expression, thereby ensuring that on induction, the maximum amount of foreign protein is produced [Patnaik, 2001].

The cultures were grown overnight with shaking at 37 °C. The cells were harvested by centrifugation using a Beckman, Avanti XPI centrifuge at 6,000 rpm at 4 °C for 20 minutes. The supernatant was discarded and the cells resuspended in 30 ml of sonication buffer (section 2.2.5). The cell paste was used immediately or stored frozen at -20 °C.

2.5.1.2 - Extraction of recombinant Ig-binding proteins from *E.coli*

The cells, resuspended in buffer, were transferred to a 50 ml Oakridge tube, where 5 mg lysozyme and 200 µl 2 mg/ml DNase I were added to hydrolyse any released DNA. As extracts from cell disruption may contain proteases, they need to be inactivated to prevent the protein from being degraded. This was achieved through the inhibition of proteases and the subsequent addition of 1 ml of a broad range protease inhibitor cocktail and 50 µl of 1 mM phenylmethanesulfonyl fluoride (PMSF), which targets serine proteases such as trypsin, and thiol proteases such as papain. After 45 minutes incubation on ice, the cells were placed in a 100 ml sonication vessel and salt ice water bath, and disrupted by sonication using an XL-2020 sonicator fitted with a ½ inch probe, for 10 consecutive minutes of 30 seconds with 30 second intermissions at setting number eight. Insoluble cellular debris was removed by centrifugation at 18,000 rpm for 30 minutes at 4 °C using a Beckman J2-21 centrifuge.

2.5.1.3 - Purification of recombinant Ig-binding proteins

The soluble cell extract was removed from the pellet and partially purified by heat treatment at 80 °C for 45 minutes, with occasional stirring. The insoluble denatured proteins were removed by centrifugation at 15,000 rpm for 20 minutes at 4 °C.

2.5.1.4 - Anion-exchange chromatography

The refined supernatant from the heat treatment was loaded at a rate of 1.5 ml/min onto a 60 ml Q-Sepharose fast flow anion exchange column, pre-equilibrated in buffer A, pH 8.0 (section 2.2.6). The Ig-binding domains of SpG have an isoelectric point of 4.2 [Achari *et al*., 1992], so the proteins will have a negative charge above this pH. Anion-exchange, a form of ion-exchange

chromatography, is therefore the most appropriate technique to use for the purification of SpG [Lian *et al*., 1992].

The column was washed for 180 minutes with buffer A to remove any unbound proteins A, using a Bio-Rad Biologic LP system. Anion-exchange chromatography was then employed utilising a linear salt gradient, from 0 to 0.25 M NaCl in buffer B, pH 8.0 (section 2.2.6), run over 600 minutes to elute bound protein. Fractions of 7.5 ml were collected over 80 tubes from the elution gradient, which started from 100 minutes into the wash to ensure that all the desired protein was collected. The purification gradient was also followed by the conductivity of the eluting buffer B and UV (OD_{280nm}). This provided a spectral analysis of the column fractions, which was used to determine the presence of SpG.

10 µl of the fractions containing SpG were separated on a high-resolution 17.5% SDS-PAGE gel (section 2.5.3), to examine the purity of protein obtained. Fractions were pooled according to purity, and concentrated by an ammonium sulphate precipitation at the 80% saturation level (section 2.5.4).

2.5.2 - Production of Fab protein

2.5.2.1 - Expression of recombinant Fab fragments by fermentation

To allow the interaction of SpG with Fab fragment to be studied extensively, it was essential that enough Fab protein was available. This was achieved by employing a modified pTTOD vector [Humphreys *et al*., 2002]. The pTTOD vector, kindly supplied by UCB-Celltech, Slough, had been optimised for the expression of Fab in *E.coli*. It was used to clone the wild-type Fab heavy and light chains genes, of known sequence, which are encoded separately in different reading frames.

An extraction and purification system was devised by Dr SL Harrison and Dr AJ Cossins to produce large quantities of the recombinant Fab fragments using fermentation [Cossins, 2006]. A 2.2 L New Brunswick Scientific BioFlo 110 Modular Benchtop Fermenter was used, with 1.5 L of SM6E media, pH 7.0, for the 48 hour fermentation of JM103 *E.coli* cells containing the pTTOD vector with the genes for wild-type Fab. 1 mM IPTG was used to induce the cells. Dr Harrison and Dr Cossins provided all cell yields for this study [Cossins, 2006].

2.5.2.2 - Periplasmic extraction of recombinant Fab protein

Each Fab light and heavy chain has an N-terminal *omp*A signal sequence, which targets the translated protein to the periplasm where the chains can fold and associate. To selectively release the periplasmic contents of the cell without disrupting the cytoplasmic membrane, and to minimise the amount of contaminating proteins in the extract, an extraction protocol based on osmotic shock was employed.

Cells containing Fab protein were harvested from the fermenter and resuspended in periplasmic extraction buffer (section 2.2.7). Osmotic shock was achieved by using this solution, which also contained protease inhibitors to prevent degradation of extracted proteins and DNase I to hydrolyse any released DNA. The resulting cell suspension was incubated overnight at 30 °C with 200 rpm shaking. Following extraction, the cellular debris was removed from the cell suspension by centrifugation at 17,000 rpm for 20 minutes at 4 °C, to obtain the soluble periplasmic extract.

2.5.2.3 - Purification of recombinant Fab protein

The soluble periplasmic extract was removed from the pellet and partially purified by heat treatment at 60 °C for 45 minutes, with occasional stirring. The insoluble denatured proteins were removed by centrifugation at 17,000 rpm for 20 minutes at 4 °C.

2.5.2.4 - Affinity chromatography using immobilised PpL

The refined supernatant from the heat treatment was loaded at a rate of 1.5 ml/min onto a 50 ml column of PpL bound (5 mg/ml) to NHS-Sepharose, pre-equilibrated in buffer A, pH 8.0 (section 2.2.8). The column was washed for approximately five column volumes with buffer A to remove any unbound proteins, using a Bio-Rad Biologic LP system. Affinity bound proteins were then eluted with a gradient of 0-50 mM buffer B, pH 2.2 (section 2.2.8), over three column volumes. Fractions of 7.5 ml were collected from the elution gradient into tubes containing 1 ml 1 M K_2HPO4, pH 9.5 to neutralise the elution buffer. The purification gradient was also followed by the conductivity of the eluting buffer B and UV (OD_{280nm}).

In addition, to obtain highly pure samples of Fab following PpL affinity purification, a 1 ml concentrated sample was applied to a Superdex 200 Hi-Load

16/60 gel filtration column, and the purified Fab collected. This was performed by Dr A Cossins [Cossins, 2006].

10 µl of the fractions containing antibody fragments were separated on a high-resolution non-reducing 17.5% SDS-PAGE gel (section 2.5.3), to examine the purity of protein obtained. Fractions were pooled according to purity, and concentrated by an ammonium sulphate precipitation at the 80% saturation level (section 2.5.4).

2.5.3 - SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were routinely separated by gel electrophoresis using a discontinuous Tris-glycine system [Laemmli, 1970].

A high-resolution separating gel containing 17.5% acrylamide was used (section 2.2.9). SpG samples were run under reducing conditions, whereas Fab was run under non-reducing conditions. All gels were run in a Biometra tank at 20-30 amps, until the bromophenol blue dye had reached the bottom. The gel was stained using Coomassie brilliant blue stain (section 2.2.13) for two hours, agitating at room temperature, and destained by diffusion into destain solution (section 2.2.14) until blue bands were visible against a clear background.

2.5.4 - Ammonium sulphate precipitation

An 80% ammonium sulphate precipitation was achieved by the addition of 56.1 g of solid ammonium sulphate per 100 ml of protein solution. The solid was added slowly to the protein solution and stirred constantly to distribute the salt and to prevent the inclusion of any unwanted proteins in the precipitate. Once solubilisation had been achieved, the solution was centrifuged at 15,000 rpm for 20 minutes at 4 °C. The pellet was retained and resuspended in a minimal volume of 80% ammonium sulphate solution and stored at 4 °C. Prior to use, protein samples were dialysed overnight at 4 °C against 4 litres of the required buffer to remove the ammonium sulphate.

2.5.5 - Estimation of protein concentration

2.5.5.1 - Bicinchoninic assay

Protein concentrations were estimated by spectral analysis at 280nm or a bicinchoninic acid (BCA) assay based on chemical principles.

Impure protein samples were quantified using the BCA technique [Smith *et al*., 1985]. The protein to be analysed was reacted with Cu^{2+} and bicinchoninic acid

resulting in the production Cu^{+}. The Cu^{+} is chelated by the BCA, which converts the pale-green colour of the free BCA to a purple colour of the copper-BCA complex. The absorbance was read at 570 nm using a Dynex MX5 microtitre plate reader. The concentration of unknown proteins was determined using a standard calibration curve of varying known concentrations of BSA.

2.5.5.2 - Spectrophotometric determination

Pure protein samples were quantified using spectral analysis. The absorption spectra of a 50 fold diluted 1 ml protein sample was measured in a 1 ml quartz cuvette with a 1 cm light path in the near UV range between 240-340 nm, using a Hitachi U-2000 spectrophotometer. This procedure was repeated using dialysis buffer (section 2.2.15), and this value was subtracted from the protein sample to form a baseline. According to the Beer Lambert Law, the absorption of the solution is proportional to the molar concentration of the protein, thus by applying this concept, the concentration of the protein can be determined. Mathematically, the Beer Lambert Law can be expressed in equation 2.1.

$$\mathbf{A = \varepsilon Cl} \qquad \textit{(equation 2.1)}$$

Where, A is absorbance, ε is molar absorptivity or extinction coefficient in litres/mole/cm, C is concentration in moles/litre, and l is path length of the sample cell in cm.

The molar extinction co-efficient value for each protein was calculated from the amino acid sequence using values of ε_{280} = 1280, 5690 and 0 litre.mole^{-1}.cm^{-1} for tyrosine, tryptophan and phenylalanine respectively [Gore, 2000]. The spectrum of the long UV absorbing amino acids, tryptophan, tyrosine and phenylalanine is presented in figure 2.2.

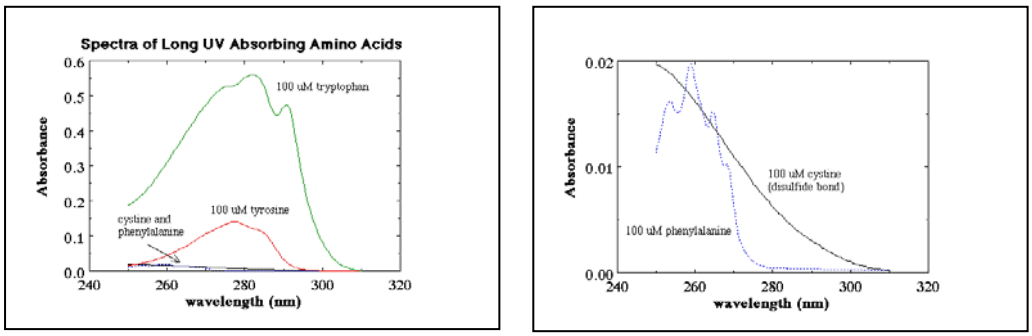


Figure 2.2 – <u>The absorption spectra of the aromatic amino acids</u>

Typical absorption spectra for the aromatic amino acids are presented. Proteins contain three aromatic amino acids; tryptophan, tyrosine and phenylalanine, of which tyrosine and tryptophan are the major absorbing species. These amino acids and disulphide bonds will principally absorb light in the near UV, in the range of 250-300nm [Weber and Teale, 1957].

In addition to the absorption spectrum being used to quantify protein concentrations, it also provides information about certain properties of the protein solution, which is important in evaluating the quality of the protein. Such information includes the identity of dominant aromatic groups and the presence of aggregated protein.

2.6 - Determination of protein binding

2.6.1 - Fluorescence measurements

All fluorescence measurements were made using a Hitachi F-2000 spectrofluorimeter with a 150 watt xenon power supply. Slit widths were routinely set to 5nm, and the photomultiplier voltage to either 400 or 700 volts depending on the sensitivity required, and a thermally jacketed cuvette holder maintained the temperature at 15 °C. Protein samples were at 10 μM in 20 mM potassium phosphate buffer, pH 8.0 (section 2.2.15) and filtered to minimise light scattering. Scans were run at 60 nm/minute and excitation wavelength set at both 280 nm and 295 nm.

All measurements were corrected, where necessary, for the inner filter effect using the formula in equation 2.2.

$$F_{corr} = F_{obs} \times 10^{(0.5\,A_{ex} + 0.5\,A_{em})} \qquad \textit{(equation 2.2)}$$

Where the A_{ex} is the absorbance of the sample measured at the excitation wavelength, and A_{em} is the absorbance at the emission wavelength.

2.6.2 - Fluorescence titrations

All titrations were performed using 1μM human Fc and aliquots of 100 μM SpG diluted in 20 mM potassium phosphate buffer (section 2.2.15), unless otherwise stated. Excitation wavelengths were kept at 280 nm or 295 nm and emission was measured at 340 nm. Readings were averaged and corrected for the inner filter effect and dilution, before they were used in subsequent calculations.

2.6.3 - Stopped-Flow fluorescence spectroscopy

All stopped-flow measurements were performed with an Applied Photophysics SX.17MV stopped-flow apparatus to enable fluorescence emission from the mixed solutions of SpG and the appropriate ligand to be observed. 2 ml drive syringes were used for each sample, equilibrated at either 15 °C or 25 °C unless otherwise stated, maintained by a Grant LTD6G circulator water bath. Protein solutions were prepared in 20 mM potassium phosphate buffer (sections 2.2.15) at either pH 6.0

or pH 8.0, which had been previously filtered and de-gassed. Studies to determine the effect of pH on the interaction of SpG, with both Fc and Fab, were performed using a triple buffer system (section 2.2.16) to buffer a wide range of pH.

To determine the rate of association, protein solutions were mixed at a 1:1 v/v mixing ratio between the ligand, Fc or Fab, which was kept at a constant initial concentration of 1 μM and 3 μM respectively, and with initial concentrations of SpG proteins varying between 60 μM and 100 μM. All measurements were scanned at an excitation wavelength of 280 nm, and tryptophan fluorescence emission was measured directly with a 335 nm glass cut-off filter. The rate of dissociation was determined by either the competitive displacement of Fc from a preformed 2:1 6 μM SpG:3 μM Fc complex with buffer, or the competitive displacement of Fab from a preformed 1:1 3 μM SpG:3 μM Fab complex with 30 μM non-fluorescent SpG mutant construct.

Stopped flow runs were completed using the maximum rate of data sampling, either set to 1000 data points at lower protein concentrations, and 400 data points at higher protein concentrations to reduce the amount of noise. Each run was performed in triplicate, with data being averaged and analysed using a single exponential curve fitting algorithm, using equation 2.3.

$$F_t = F_0 \exp(-kt) + C \qquad \textit{(equation 2.3)}$$

Where F_0 and F_t are the fluorescence intensities at time 0 and time t respectively, k is the rate constant for the reaction and C is a constant.

Data are presented as reaction progress curves, which were averaged as three sets of triplicate traces.

2.6.4 - Isothermal Titration Calorimetry (ITC)

ITC measurements were performed using a MicroCal VP-ITC MicroCalorimeter. SpG and ligands, Fc or Fab, were dialysed together overnight in 20 mM potassium phosphate buffer, pH 8.0 (section 2.2.15), otherwise an energy change due to buffer dilution will be observed. All protein solutions and buffers were degassed for 20 minutes prior to use. The cell was loaded with 10 μM Fc or 20 μM Fab, and the syringe was filled with 400-600 μM SpG, unless otherwise stated. The MicroCal VP-ITC system was under computer control, and a standard protocol was followed using an initial injection of 2 μl, preceded by varying number of

injections of 5-7 µl depending on the nature of the protein•ligand interaction, with titrations being performed at 15 °C. All data obtained were analysed using the manufacturer's Origin™ software.

2.7 – Conformational Stability Determination

2.7.1– Denaturation studies

Denaturation studies were carried out using guanidine hydrochloride at a stock solution of 6 M solution in 20 mM potassium phosphate buffer (section 2.2.15), which was filtered prior to use. For accurate determination of guanidine hydrochloride concentrations after filtration, the refractive index of the solutions were measured and calculated using equation 2.4 [Pace *et al.*, 1986].

$$[Gdn\text{-}HCl] = 57.147(\Delta N) + 38.68(\Delta N)^2 - 91.60\ (\Delta N)^3 \qquad \textit{(equation 2.4)}$$

Where, Δ*N* is the difference between the refractive index of the denaturant and the refractive index of the buffer used.

Controlled unfolding measurements were made in Gdn-HCl at increasing concentrations from 0-6 M. The final Gdn-HCl concentration with 10 µM SpG was determined by refractometry as before. The change in protein fluorescence was followed at an excitation wavelength of 280nm, and the fluorescence emission was measured at 350 nm. Each sample was pre-incubated at 15 °C for two hours to reach thermal equilibrium before being measured.

2.7.2 - Circular Dichroism (CD)

CD measurements were carried out using a Jasco J-720 CD spectropolarimeter. Spectra were measured in both the near and far UV range. Spectra in the far UV range, 190-250 nm, were measured in triplicate at room temperature using a 0.01 cm light path with a scan speed of 50 nm/min. Spectra in the near UV range, 250-330 nm, were measured in triplicate, again at room temperature using a 0.5 cm light path with a scan speed of 20 nm/min. Spectra were obtained using protein at a concentration of 100 µM, in 20 mM potassium phosphate buffer, pH 8.0 (section 2.2.15), with a response time of 4 seconds, band-width of 1.0 nm, resolution of 0.2 nm, and sensitivity between 20-50 mdeg. All data obtained were analysed using the manufacturer's Standard Analysis software.

Data sets shown are accumulations of nine spectral scans, adjusted for buffer signal, and expressed as Molar ellipticity (θ=deg.cm^2.dmol^{-1}).

CHAPTER 3.0

THE DESIGN, CLONING AND EXPRESSION OF FC & FAB CONSTRUCTS

3.1 - Introduction

The aim of this research is to study a single Ig-binding domain of SpG, which was based on the sequence for SpG_{C2}-1 designed by Walker [Walker, 1994], a hybrid of the C-terminus of the C1 domain and the N-terminus of the C3 domain. Specific amino acid residues that contribute to the Fc or Fab binding interactions via a network of hydrogen bonds were identified and studied by a programme of site-directed mutagenesis. This allowed information on the importance of these individual amino acids in the encoded protein to be obtained, the key residues that are critical for these interactions to be identified and the contribution of a specific amino acid residue to the biological function of SpG to be evaluated. To achieve these objectives, both currently available and new mutant constructs of D-SpG_{C2}-1 were used to alter the Fc and Fab binding affinities from that of the wild-type protein.

3.2 - Mutagenesis

Mutagenesis is the alteration of a DNA sequence by specific chemical or enzymatic methods, and is a fundamentally important DNA technology. DNA mutagenesis is used in many applications, such as DNA structural analysis, protein expression studies and structure-function relationships in cloned gene products. Cloned genes allow mutagenesis to be much more precise by replacing specific bases in a gene, or specific amino acids at selected sites in the protein.

3.2.1 - PCR-based site-directed mutagenesis

3.2.1.1 - The polymerase chain reaction (PCR)

Polymerase chain reaction (PCR) is a relatively new and versatile *in vitro* technique for amplifying defined target DNA sequences present within a source of DNA. Since the first reports describing this technology in the mid 1980s, there have been numerous applications in both basic and clinical research, including fundamental technologies such as DNA sequencing and mutagenesis, both of which can be accomplished using PCR-based methods.

A schematic representation of PCR is shown in figure 3.1. The template DNA is first denatured by heating, at approximately 94 °C, in the presence of a large molar excess of each of the two primers and DNA precursors, the four deoxynucleoside triphosphates (dNTPs) dATP, dCTP, dGTP and dTTP. The primers move around by Brownian motion, which results in hydrogen bonds being constantly formed and broken between them and the single-stranded template. They bind specifically to complementary DNA sequences at the target site and initiate the synthesis of new DNA strands, which are complementary to the individual DNA strands of the target DNA segment, and which overlap each other.

The sequence of the primers that complement the template exactly are more stable, and on this small section of double-stranded DNA (made up of the template DNA and primer), the DNA polymerase can attach and start copying the template. Once there are a few bases copied, the stability of the primer-template complex is so strong between the template and the primer that it no longer breaks.

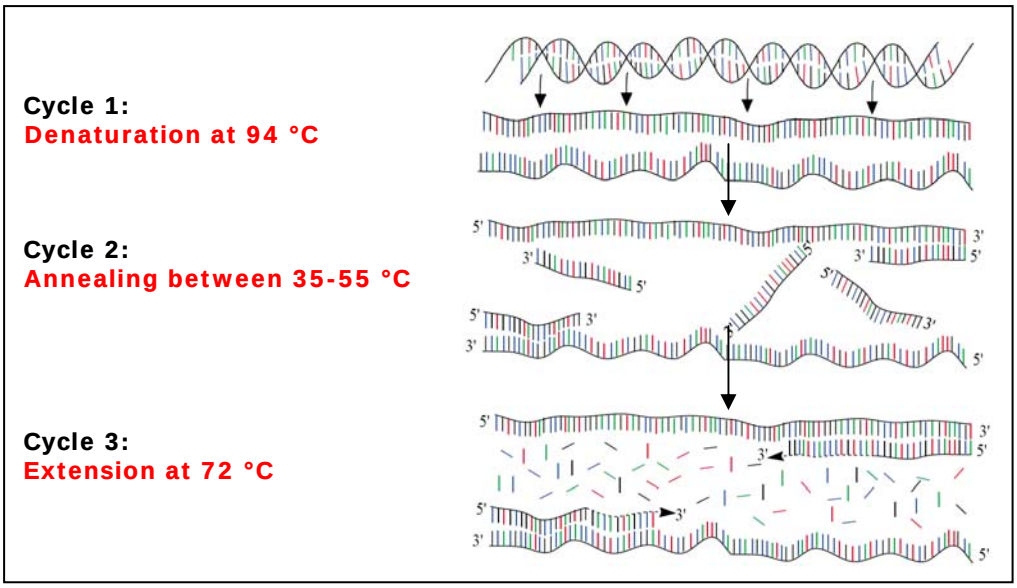


Figure 3.1 - An adapted schematic representation of PCR

There are three major steps in a PCR, which are repeated for approximately 30-40 cycles. During denaturation, at 94 °C, the double-stranded DNA melts open to two strands of single-stranded DNA. During annealing, at approximately 54 °C, the primers are moving around and bind specifically to complementary DNA sequences at the target site. The synthesis of new DNA strands is initiated and the DNA polymerase attaches and starts copying the template. The polymerase adds dNTPs from 5' to 3', reading the template from 3' to 5' side, and bases that are complementary to the template DNA are coupled to the primer on the 3' side [Ban, 2006].

The reaction is then cooled to a temperature (between 35 °C and 55 °C) that allows the primers to anneal to their target sequences, after which the annealed primers are extended with DNA polymerase at approximately 72 °C, the ideal working temperature for the polymerase. The bases that are complementary to the template DNA are coupled to the primer on the 3' side, as the polymerase adds dNTPs from 5' to 3', reading the template from 3' to 5' side.

The expected product of this reaction is a segment of double-stranded DNA and is produced at the end of the second cycle. It is identified by the 5’ termini of the oligonucleotides primers, and its length is defined by the distance between the primers. The product is then amplified in every subsequent cycle of denaturation, annealing and DNA synthesis, with the products of one cycle of amplification serving as a template for the next. This allows each successive cycle to double the amount of desired DNA product [Erlich *et al*., 1988], leading to an exponential increase in product.

3.2.1.2 - PCR-based mutagenesis

The polymerase chain reaction has proven to be particularly useful for a wide range of mutation procedures and applications, and PCR mutagenesis procedures make it possible to modify and engineer any target DNA sequence.

To permit the selective amplification of a specific target DNA sequence within a heterogeneous collection of DNA sequences, some prior DNA sequence information from the target sequence is required. This information is used to design two mutagenic oligonucleotide primers, which are often about 15–25 nucleotides long and specific for the target sequence. The sequence of these primers is perfectly complementary to the gene sequence in the region to be mutated, but with a single codon difference at the intended mutation site. This difference is complementary to the desired mutant nucleotide sequence rather than the original, which leads to the replication of a certain region, and allows for insertions, deletions or point mutations in the gene of interest.

Two further oligonucleotides are used as primers and have different complementary sequences that anneal on the opposite strands of the template DNA, and flank the segment of DNA to be amplified. New DNA synthesis is achieved in a series of synthetic reactions that are catalysed by a thermostable DNA polymerase [Saiki *et al*., 1988], to create a complementary full-length sequence containing the desired mutation. The newly formed heteroduplex is then used to transform cells, and the desired mutant genes can be identified by screening for the mutation.

3.2.2 - PCR-based site-directed mutagenesis using splicing by overlap extension

PCR-based site-directed mutagenesis using the principles of Splicing by Overlap Extension (SOE-ing) [Ho *et al*., 1989] was employed for this research for the deletion and insertion of sequences, as well as introducing point mutations into the SpG gene. Two-sided SOE-ing is a sequence-independent method for site-directed mutagenesis. It is based on the idea that a PCR product can be engineered by adding or changing sequences at its ends, so that the product itself can be used to prime DNA synthesis in a subsequent overlap extension reaction to create mutant or recombinant molecules [Horton, 1995].

3.2.2.1 - Mis-match primers

PCR based site-directed mutagenesis using splicing by overlap extension (see section 2.4.4.2) was employed to mutate the proposed Fc and Fab residues. SOE-ing requires

the use of two successive polymerase chain reactions and four primers; two flanking primers and two mutagenic primers. The two flanking primers, A and B, were designed to anneal to a region outside the gene to include any restriction digest site, whereas the mutagenic primers were designed to anneal to the gene to introduce the desired mutation. This results in an amplified product with a specific pre-determined mutation of choice located in a central segment. This is illustrated below in figure 3.2.

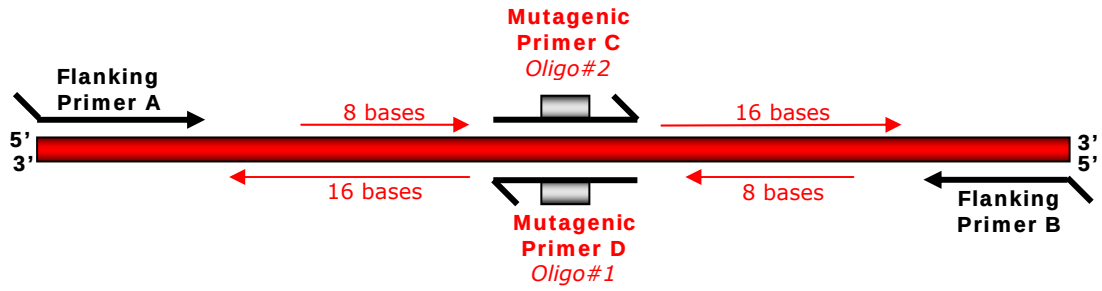


Figure 3.2 – An adapted schematic diagram of mis-match primer design

Oligonucleotide primers were 27 bases in length with the mis-match codon for each designated mutant central to the primer, illustrated by the grey rectangle.

All oligonucleotide primers were approximately 27 nucleotide bases in length, unless otherwise stated, and the mis-match mutation within the primer was placed at a unique site. The upper strand mutagenic primer, primer C, was ordered as oligo#2 and the central mis-match codon was flanked on its 5' side by 8 bases and by 16 bases on its 3' side, complimentary in sequence to the template DNA. The lower strand mutagenic primer, primer D, was ordered as oligo#1, and again the mutation mis-match codon was central, flanked on its 5' side by 8 bases and by 16 bases on its 3' side anti-sense in sequence to the template DNA.

By employing a primer design and sequence alignment software program, DNA Star, all primer sequences were studied for dimer and hairpin formation. The 27 nucleotide primer was sufficient to ensure site-specific annealing of the primer to the template DNA. All primers were produced in the 5' to 3' orientation, allowing them to bind to the DNA template produced in the cloning vector.

3.2.2.2 - *Pfu* DNA polymerase

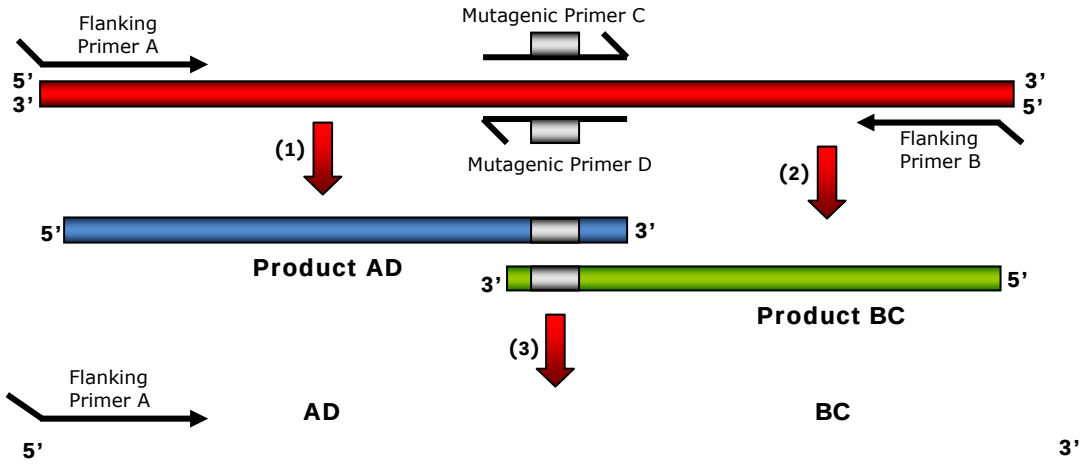
Pfu DNA polymerase is a thermostable enzyme found in the hyperthermophilic *Pyrococcus furiosus* [Lundberg *et al*., 1991], where it functions *in vivo* to replicate the organism's DNA. *Pfu* DNA polymerase was employed during this research to catalyse the polymerisation of nucleotides into double-stranded DNA, in the 5' to 3' direction, during each extension step in all PCR performed.

The main difference between *Pfu* and alternative thermostable DNA polymerases, such as *Taq* [Eckert and Kunkel, 1990], is *Pfu's* thermostability and 'proofreading' properties. Unlike *Taq*, *Pfu* possesses 3' to 5' exonuclease proofreading activity, which corrects nucleotide-misincorporation errors as it works its way along the DNA from the 3' to the 5' end. This means that *Pfu* DNA polymerase-generated PCR fragments comprise of fewer errors than *Taq*-generated PCR inserts [Cline *et al*., 1996]. It is recommended for use in PCR and primer extension reactions that require high fidelity.

3.2.2.3 - SOE-ing PCR

SOE-ing involves two successive series of PCR. The first PCR involves two parallel reactions, PCR 1 and 2, which leads to the initial amplification of two over-lapping segments of DNA containing an introduced mutation. The two mutagenic primers incorporate the desired nucleotide changes at the mutation site, and the generated fragments have a 5'-terminal overlap, as illustrated in figure 3.3.

The product fragments, AD and BC, may vary in size depending on the position of the mutation in the gene, and are purified to remove the original primers. The purified amplified fragments are then combined in a subsequent 1:1 (mol:mol) ratio fusion reaction, PCR 3, in which the overlapping ends are denatured and allowed to reanneal. This allows the DNA polymerase *Pfu* to extend the 3' end of each strand with recessed ends to serve as a primer for the 3' extension of the complementary strand. Thereafter, the resulting full length fusion product with the introduced mutation in a central segment can be further amplified by PCR using the outer primers A and B only [Ho *et al*., 1989].



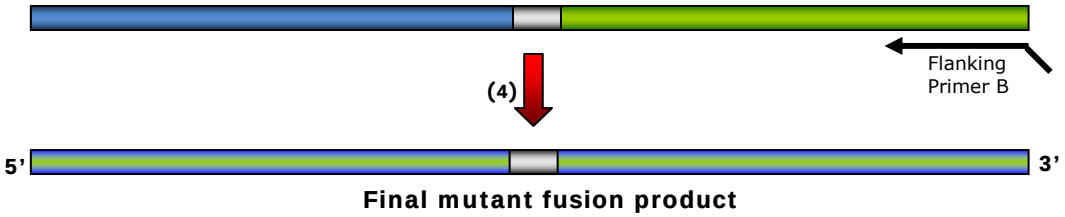


Figure 3.3 - An adapted schematic diagram of PCR-based site-directed mutagenesis by overlap extension

Modified two sided splicing by overlap extension SOE-ing method of site-directed mutagenesis is presented. (1) Primers A and D are used to amplify the template to produce the product AD. (2) Primers B and C are used to produce product BC. (3) AD and BC are combined in a new PCR with primers A and B. (4) A final fragment is produced, which has the desired mutations encoded in primers C and D. The initial double-stranded DNA is represented by a red rectangle; synthetic primers are represented by lines with arrows indicating the 5' to 3' orientation; the site of mutagenesis is represented by a small grey rectangle; and PCR products are represented by blue and green rectangles [Ho et al., 1989].

3.3 - Construction of the gene for SpG_{C2}-1

3.3.1 - Gene manipulation

A synthetic gene encoding a single Ig-binding domain was constructed from chemically synthesised oligonucleotides and placed in an appropriate plasmid vector, *PKK223-3*, to direct expression in *E.coli* [Walker, 1994]. This 208 bp Ig-binding domain gene was originally based on the sequence for the C2 domain [Goward *et al*., 1990], which is a hybrid of the C-terminus of the C1 domain and the N-terminus of the C3 domain created by gene duplication, mutation and recombination [Olsson *et al*., 1987]. Alterations were made to the 3' end of the SpG_{C2} gene, which included three stop codons to ensure effective termination of transcription at the end of the gene, an additional six residues (residues MTPAVS) were incorporated at the N-terminus to facilitate cloning and expression and a C-terminal cysteine was introduced for possible immobilisation onto Sepharose to create an IgG purification column [Walker, 1994]. Figures 3.5 and 3.7 present the entire nucleotide sequence for the SpG_{C2}-1 gene, together with the amino acid sequence it encodes.

3.3.2 - Expression vectors of SpG_{C2}-1 gene

3.3.2.1 - Expression vector *PKK223-3*

The expression vector *PKK223-3*, presented in figure 3.4, was initially chosen for the expression of the SpG_{C2}-1 gene. It provides transcriptional control via the strong *tac* promoter, which is upstream from the multiple cloning site (MCS), and the strong *rrn*B ribosomal terminator downstream for control of protein expression. *rrn*B transcription terminators stabilise the plasmid by inhibiting read-through transcription initiated from the *tac* promoter in the parent plasmid. The *tac* promoter is a hybrid of the *lac* and *trp* promoter sequences, and is stronger than either. This sequence is recognised and bound by the *lac* repressor, which prevents the initiation of downstream DNA by *E.coli*

polymerase. Repression is alleviated by a non-hydrolysable analogue, such as isopropyl-β-D-thiogalactoside (IPTG), which binds to the repressor preventing it binding to the promoter [Bell and Lewis, 2000]. In order for the synthetic gene to be expressed directly, translation initiation elements, that were originally provided by the DNaseI gene sequence [Worrall and Connolly, 1990] as with the expression of other Ig-binding proteins, had to be introduced.

Figure 3.4 – A schematic representation of expression plasmid vector *PKK223-3*

The expression plasmid vector, PKK223-3, was used to over express SpG from a cloned gene in E.coli. This is achieved by the presence of a strong tac promoter upstream from the MCS and the strong rrnB ribosomal terminator downstream for control of protein expression.

Direct expression was achieved using synthetic linkers constructed from the Shine-Dalgarno site from the DNaseI sequence. This sequence has already been incorporated into the *PKK223-3* vector in an AP07-2 construct for the direct expression of SpA [Popplewell, 1991], and was adapted for the SpG gene. A linker sequence was designed to encode six amino acids before the start of the C2 domain as found in the native linker of SpG. The sequence chosen is adapted from that used by Lian [Lian *et al*., 1992], and exchanges a methionine for the native leucine as the first amino acid, providing the codon required for the initiation of translation. The synthetic gene for the single Ig-binding domain of SpG was renamed D-SpG$_{C2}$-1.

```
            EcoR1                                  SacI    NcoI(blocked)
    PKK vector ------  ** *** **                 ------------- Linker
5'  gaa aca gaa ttc gat ttc tag gag gtg agc tcc atg act ccg gct gtt
3'  ctt tgt ctt aag cta aag atc ctc cac tcg agg tac tga ggc cga caa
                                                  Met Thr Pro Ala Val
     KpnI(blocked)                                 1   2   3   4   5
     ------ Domain 1
    agt acc tac aaa ctg gtt atc aac ggt aaa acc ctg aaa ggt gaa acc
    tca tgg atg ttt gac caa tag ttg cca ttt tgg gac ttt cca ctt tgg
    Ser Thr Tyr Lys Leu Val Ile Asn Gly Lys Thr Leu Lys Gly Glu Thr
     6   7   8   9  10  11  12  13  14  15  16  17  18  19  20   21

    acc acc gaa gct gtt gac gct gct acc gct gaa aaa gtt ttc aaa cag
```

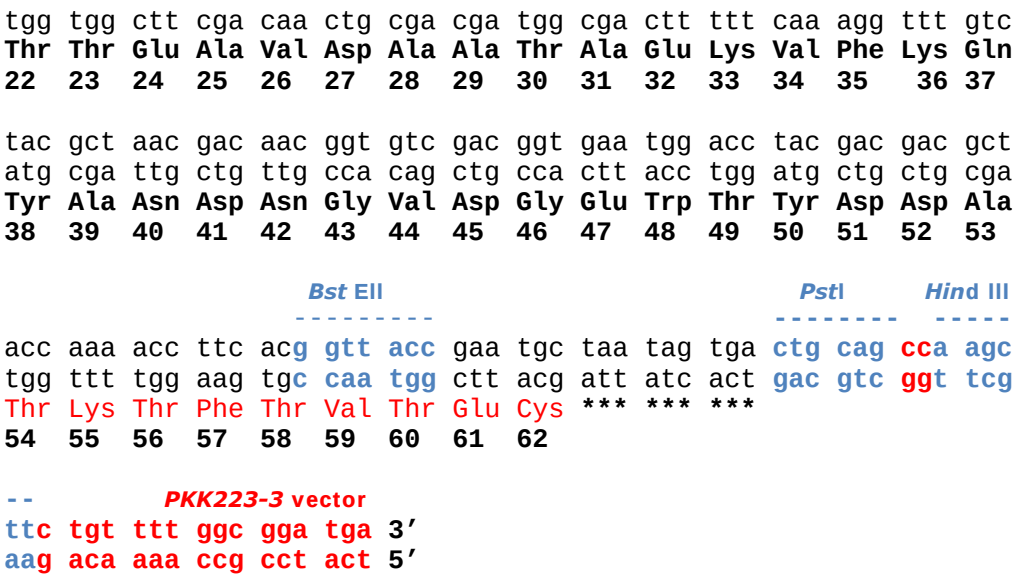


Figure 3.5 - Nucleotide sequence for the wild-type synthetic gene D-SpG$_{C2}$-1 coding for a single Ig-binding domain based on SpG in *PKK223-3* vector

The entire nucleotide sequence for wild-type D-SpG$_{C2}$-1 gene, together with the amino acid sequence it encodes is presented in PKK223-3 vector. Red bases denote the vector PKK223-3 sequence; black bases signify the wild-type D-SpG$_{C2}$-1 gene sequence; blue bases indicate the position of restriction sites within the sequence; pink bases denote the linker sequence; purple bases signify blocked cloning site; *** *** represents the Shine-Dalgarno site; and *** *** represents the stop codons.

3.3.2.2 - Expression vector *pQE*-30

The successful expression and purification of the gene construct is required to effectively achieve the aims of this research. However, despite repeating the identical induction conditions stated by Walker [Walker, 1994], expression levels of D-SpG$_{C2}$-1 dropped considerably since the time that the construct was made. Many attempts were made to rectify this problem, which was eventually overcome by the successful cloning of the D-SpG$_{C2}$-1 gene into the N-terminal 6xHis tagged *pQE*-30 over-expression vector, with high levels of expression of the 6xHis-tagged D-SpG$_{C2}$-1 protein being achieved. Figure 3.6 presents an adapted schematic diagram of the *pQE*-30 over-expression vector. At pH 8.0, the His tag is small, largely uncharged, and therefore, does not affect secretion, compartmentalisation or folding of the fusion protein within the cell. Later studies show that the presence of the 6xHis tag does not interfere with the structure or function of the purified protein, and also allows the immobilisation of the protein on metal-chelating surfaces for protein interaction studies.

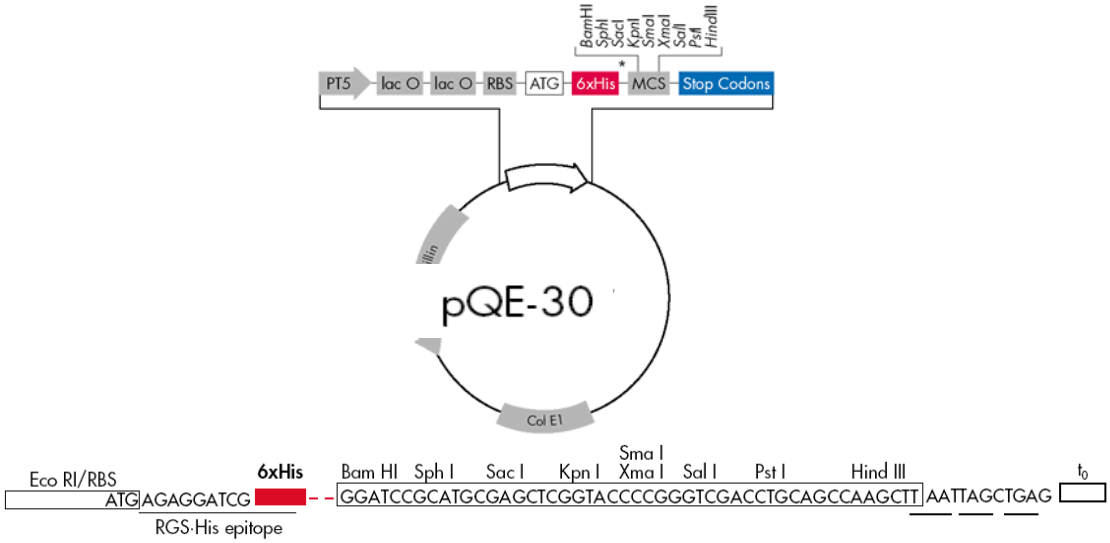


Figure 3.6 - An adapted schematic representation of the 6xHis-tagged *pQE*-30 over-expression vector

A high-level expression of 6xHis-tagged proteins can be achieved in E.coli using the QIAexpress pQE-30 vector. The vector has been N-terminally tagged with 6 consecutive histidine residues (6xHis tag). The 0SpG$_{C2}$-1 gene was cloned into the vector at the restriction sites Bam HI and Hind III, such that there are no additional residues introduced between the RGS-His tag and the start of the protein sequence [Qiagen].

3.3.2.3 - Sub-cloning D-SpG$_{C2}$-1 gene into *pQE*-30 vector

PCR and restriction enzyme digests were used to clone the gene constructs into the 6xHis tagged *pQE*-30 expression vector, as previously described in section 2.4.4.2 and 2.4.5 respectively. *Bam* HI and *Hin*d III restriction sites were introduced at the start and end of the D-SpG$_{C2}$-1 genes respectively, by PCR using specifically designed oligonucleotide primers.

Bam HI and *Hin*d III restriction enzyme digests were then performed on the PCR gene products obtained and the 6xHis tagged *pQE*-30 vector, as shown in figure 3.8, and the appropriate fragments excised and extracted from agarose gel (section 2.4.2.3). The *pQE*-30 vector was dephosphorylated to prevent self-ligation (section 2.4.6). Various ligation mixtures were set up, with appropriate controls, to give final vector: insert µl ratios of 1:1, 1:3 and 1:7. The ligation control grew no colonies, and the ligation ratios grew between 50 to 100 colonies for all gene constructs. This ratio difference was high enough to be certain that the ligation had been successful. Ligation products were transformed into *E.coli* JM103 cells (section 2.4.7) and used to produce plasmid DNA.

```
   pQE-30 vector          * *** ***                                  -------
5' aca gaa ttc att aaa gag gag aaa tta act atg aga gga tcg cat cac
3' tgt ctt aag taa aat tga ctc ttt aat tga tac tct cct agc gta gtg
                                          Met Arg Gly Ser His His

       6X HIS TAG      BamH 1
```

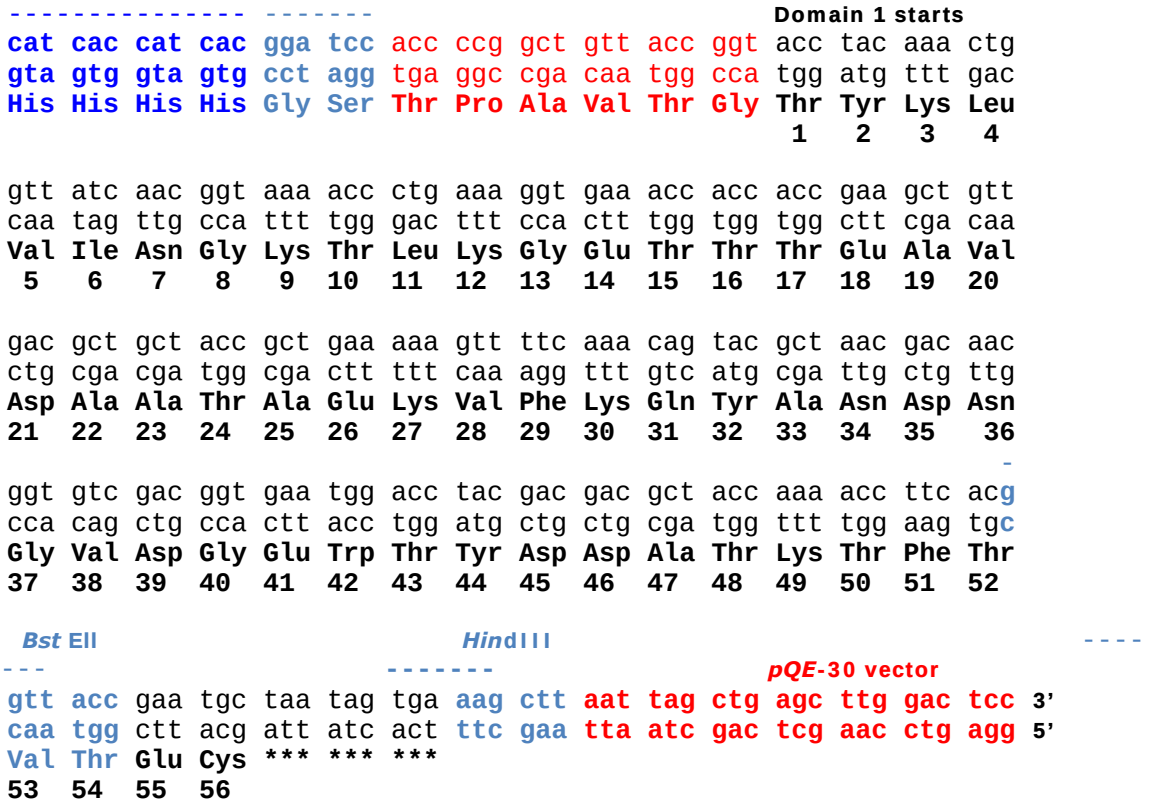


Figure 3.7 - Nucleotide sequence for the wild-type synthetic gene D-SpG$_{C2}$-1 coding for a single Ig-binding domain based on SpG in *pQE*-30 vector

The entire nucleotide sequence for wild-type D-SpG$_{c2}$-1 gene, together with the amino acid sequence it encodes is presented. Red bases denote the vector pQE30-1 sequence; black bases signify the wild-type D-SpG$_{c2}$-1 gene sequence; blue bases indicate the position of restriction sites within the sequence; * *** *** represents the Shine-Dalgarno site; and *** *** *** represents the stop codons.

The D-SpG$_{C2}$-1 constructs were then sequenced to check that the sequences have been correctly inserted into the *pQE*-30 vector at the designated restriction sites. To be certain of the insertion of the genes into the *pQE*-30 plasmid vector, an analytical restriction enzyme digest was also performed on three samples of plasmid DNA for each gene construct, prepared as described in section 2.4.5. Figure 3.8 presents 2% agarose gels illustrating the 208 bp gene fragment successfully released from the *pQE*-30 plasmid vector, indicating that the D-SpG$_{C2}$-1 genes had been successfully inserted into the 6xHis tagged *pQE*-30 vector.

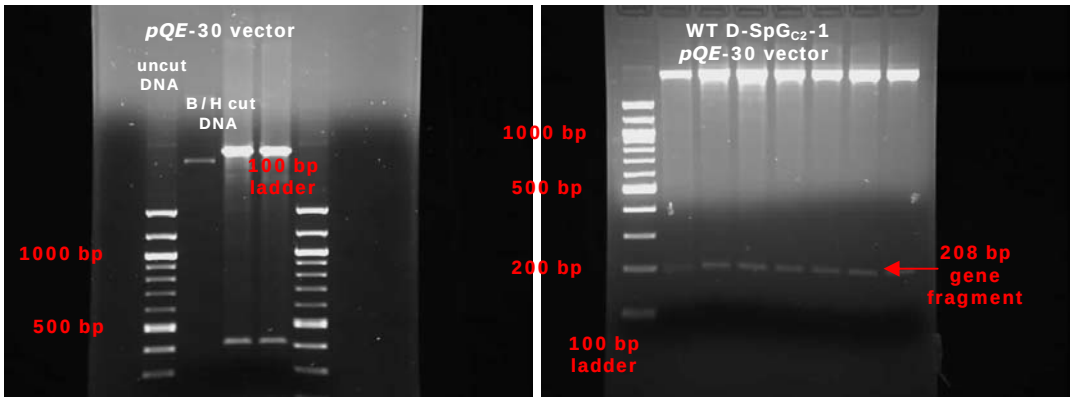


Figure 3.8 - Two 2% agarose gels illustrating the successful cloning of wild-type D-SpG$_{C2}$-1 genes into *pQE*-30 vector

These two 2% agarose gels illustrate Bam HI and Hind III restriction enzyme digests performed on both the 6xHis tagged pQE-30 vector (first gel) and wild-type D-SpG$_{C2}$-1 pQE-30 DNA (second gel).

3.3.2.4 - Gene nomenclature

Several groups have reported complete or partial DNA sequences for the gene for SpG in their research [Guss *et al*., 1986; Fahnestock *et al*., 1986; Olsson *et al*., 1987; Goward *et al*., 1990; Lian *et al*., 1991 and 1994; Gronenborn *et al.,* 1993 and 1996; Sauer-Eriksson *et al.,* 1995; Derrick *et al*., 1999], all of which employ various numbering systems. The exclusive tryptophan residue in the amino acid sequence is used to compare these different nomenclatures.

The work of Walker [Walker, 1994] on SpG, uses the same nomenclature employed by Lian *et al* [Lian *et al.*, 1991 and 1994] and Derrick *et al* [Derrick *et al*., 1999] in which the tryptophan residue is number 48 (W48) in the sequence. This is also the same nomenclature presented in figure 3.5, which shows the SpG sequence cloned into the *PKK223-3* vector that was initially chosen for the expression of the D-SpG$_{C2}$-1 gene.

Gronenborn *et al.,* [1993 and 1996], who studied the X-ray structure of the B1 domain in complex with Fc, and Sauer-Eriksson *et al*., [1995] who identified key residues of the C2 domain in complex with Fc, have both published research using another nomenclature, where the tryptophan residue is number 43 (W43) in the sequence.

However, in comparison, the gene sequences previously published by Fahnstock *et al.,* [1986] who isolated a two IgG-binding domain SpG and labelled the domains B1 and B2, and Guss *et al.,* [1986] who isolated a three IgG-binding domain SpG and labelled the domains C1, C2 and C3, employ a different numbering system. A comparison of the nucleotide sequences for the three Ig-binding C domains of SpG isolated by Guss *et al* is

presented in figure 3.9.

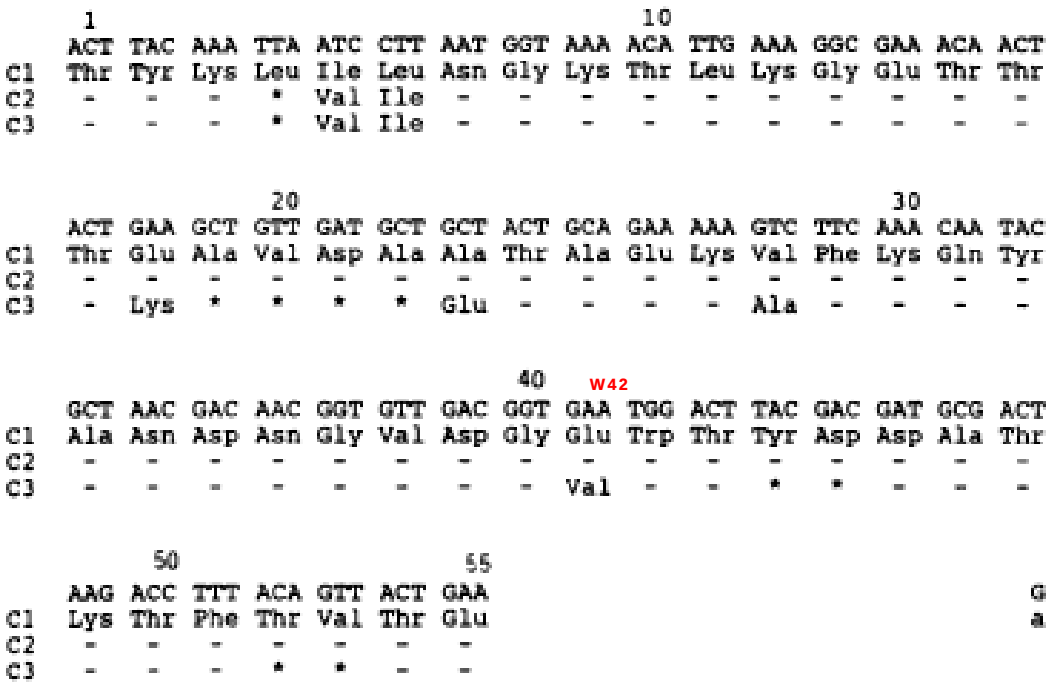


Figure 3.9 - A comparison of nucleotide sequences for the three Ig-binding C domains of wild-type SpG

A comparison of the nucleotide sequences for the three Ig-binding C domains of wild-type SpG, together with the amino acid sequence it encodes is presented. The sequences are aligned to achieve maximum homology and the comparison is based on the C1 domain. Identical triplets are indicated with a bar and altered nucleotide triplets, which do not alter the amino acid, are marked with an asterisk. Tryptophan residue number 42 is highlighted in red [Guss et al., 1986].

The C2 domain of SpG is a hybrid of the C1 and C3 domains, and the gene sequence of this domain, cloned into the *pQE*-30 expression vector, was used throughout this research. This particular sequence of amino acids, presented in figure 3.7, corresponds to the same nomenclature proposed by Guss *et al* and Fahnestock *et al*, where the tryptophan residue is number 42 (W42) in the sequence.

3.4 - Construction of D-SpG$_{C2}$-1 mutants

3.4.1 - Rationale of choosing specific D-SpG$_{C2}$-1 Fc and Fab mutants

The C domain of SpG has two main attributes. Firstly, its ability to bind IgG non-immunogenically via the interaction with both Fc and Fab regions, and secondly, it's unique and stable folded structure. Studies determining any differences in both binding

and stability were initially achieved by site-directed mutagenesis. By introducing site-directed changes at specific positions in the protein sequence, it is possible to test hypotheses about protein stability and function. The amino acid residues that were chosen for mutation were selected for their possible involvement in each of SpG's characteristics.

3.4.1.1 - Studying the effect of mutagenesis on the binding of SpG

A single Ig-binding domain of SpG has been shown to bind to the Fc portion of IgG [Gronenborn *et al.*, 1993; Sauer-Eriksson *et al.*, 1995], and also to the Fab portion [Derrick and Wigley, 1992; Lian *et al*, 1994; Derrick and Wigley, 1994]. It is possible to make a direct comparison of the binding of SpG to these fragments of IgG. In the complex with Fab, the second strand of the β-sheet plays a major role in the interaction, together with the loop at the loop at the carboxyl-terminal end of the helix. In contrast, the regions most affected by the binding of Fc are the helix, the third strand of the β-sheet, and the intervening loop, illustrated in figure 3.10.

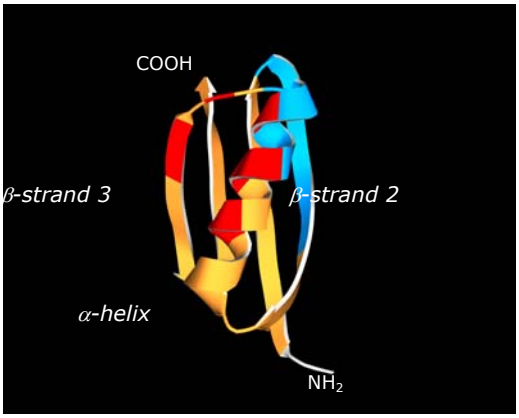


Figure 3.10 – **A ribbon representation of SpG indicating both Fc and Fab binding regions**
Specific amino-acid residues that contribute to both Fc and Fab binding interactions have been identified. The Fc binding region, α-helix and β-strand 3, is shown in red; the Fab binding region, α-helix and β-strand 2, is shown in blue.

It is therefore clear, that SpG binds quite differently to the constant domains in Fab and Fc, and binds edge-on to the C_H1 domain of Fab [Derrick and Wigley, 1992] rather than in a cleft between C_H2 and C_H3 of Fc [Diesenhofer, 1981]. It is particularly notable that a small protein of just 56 residues is able to recognise two different protein surfaces specifically. This is achieved by the two almost completely non-overlapping sets of residues on its surface.

Mutagenesis of the key residues in D-SpG_{C2}-1, that are critical for both interactions, will identify the contribution of a specific amino acid to the biological function of this domain.

3.4.1.2 - Studying the effect of mutagenesis on the stability of SpG

The principal stabilising element in all globular proteins is the hydrophobic core, the

group of hydrophobic residues that is shielded from solvent in the native state structure [Northey *et al*., 2002]. The role and packing of residues in the hydrophobic core has been shown to be important for the structure and stability of a variety of proteins [Kellis *et al*., 1989; Eriksson *et al*., 1992; Richards and Lim, 1993; Munson *et al*., 1996].

The C2 Ig-binding domain of SpG, upon which the D-SpG$_{C2}$-1 gene is based, is a very stable, small protein, with a large proportion of its residues (95%) being involved in its ordered secondary structure or forming part of the hydrophobic core. The involvement of nearly all of the residues in the regular secondary structure ensures that a large number of hydrogen bonds stabilise the protein, with over 70% of amide nitrogen and carbonyl oxygen atoms participating in hydrogen bonds. This large number of hydrogen bonds includes 44 backbone to backbone, 9 backbone to side chain and 5 side-chains to side-chain interactions [Achari *et al*., 1992]. When SpG is complexed with Fc and Fab, the hydrogen bond network is continued, and stabilises these interactions.

The stability of the C2 domain of SpG can also be attributed to its unique folded structure, which renders it unusually stable without disulphide bridges, and the hydrophobic core, which is formed by the four-stranded β-sheet and the four-turn α-helix. This structure results in a very tightly packed hydrophobic core, which comprises of 16 amino acid side chains. These include Tyr2, Leu4, Ile6, Gly8, Leu11, Thr15, Thr17, Val25, Phe29, Tyr32, Ala33, Val38, Trp42, Tyr44, Phe51 and Val53, with six of these residues possessing hydrogen-bonding capabilities [Achari *et al*., 1992]. This is illustrated in figure 3.11.
Specific amino acid residues that have been shown to contribute to the stability of SpG's interaction with Fc and Fab via an extensive network of hydrogen bonds were also studied by a programme of site-directed mutagenesis.

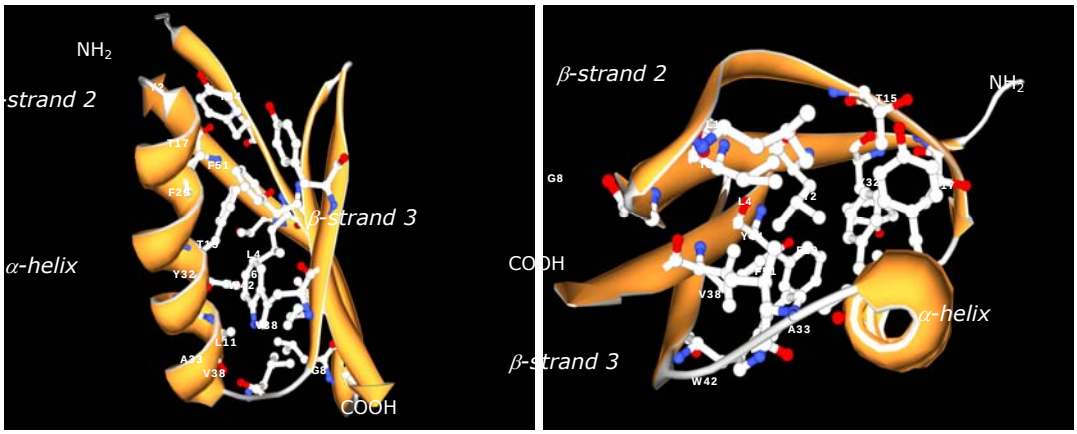


Figure 3.11 – <u>Ribbon representations of the hydrophobic core of an Ig-binding domain of SpG</u>

The three-dimensional structure of a single Ig-binding domain of SpG is presented, illustrating the very tightly packed hydrophobic core. The hydrophobic core lies between

the β-sheet and the α-helix, and includes 16 amino acid side chains [Achari et al., 1992].

3.4.2 - Molecular modelling

By employing molecular modelling, via DeepView – Swiss Protein Data Bank (PDB) Viewer, the proposed D-SpG$_{C2}$-1 mutants implicated in both Fc and Fab interactions were studied. This was to minimise potentially disruptive amino acid substitutions, particularly those that could affect the secondary structure and conformational stability of SpG. Ray-traced quality images of the molecular structures, generated from DeepView - Swiss-PDBViewer, were achieved using Persistence Vision POV-Ray.

When creating a mutant protein, the effect on stability should be monitored, as changes in affinity and activity can result from the deformation of secondary and tertiary structure, which is not always due to the direct involvement from the substituted amino acid side chain. When performing site-directed mutagenesis, it is important to consider these structural implications, together with the size and polarity of the amino acid side-chains. Substitutions are normally made as conservative as possible, and involve introducing small amino acid side-chains as they are more easily accommodated into the structure [Chou *et al*., 1974].

It should also be noted that when mutating a small protein such as SpG, residues forming the hydrophobic core should be packed tightly together with no steric overlaps or spaces. Studies have revealed that amino acids may have a particular propensity for secondary structure over another [Chou and Fasman, 1974]. This consideration must be taken into account when considering the substitution, as well as size and polarity of the side-chains.

3.4.3 - Sequencing of D-SpG$_{C2}$-1 mutant constructs

All D-SpG$_{C2}$-1 mutant constructs were checked to ensure that the sequences contained the mutated amino acids and had been correctly inserted into the *pQE*-30 vector at the designated restriction sites, and this was accomplished by dideoxynucleotide sequencing, performed by MWG-Biotech.

This chain termination technique utilises fluorescently labelled 2',3'-dideoxynucleotide triphosphates (ddNTPs - ddATP, ddCTP, ddGTP and ddTTP), which are added to the reaction in addition to deoxynucleotides (dNTPs). DNA extension will occur as normal, as in PCR, until the DNA polymerase randomly incorporates a fluorescently labelled ddNTP (see figure 3.12). Therefore, once a ddNTP is incorporated into the newly synthesised DNA strand, it can no longer be elongated due to the absence of the 3'-OH group. This

leads to the termination of the reaction because it cannot form a phosphodiester bond with the next deoxynucleotide.

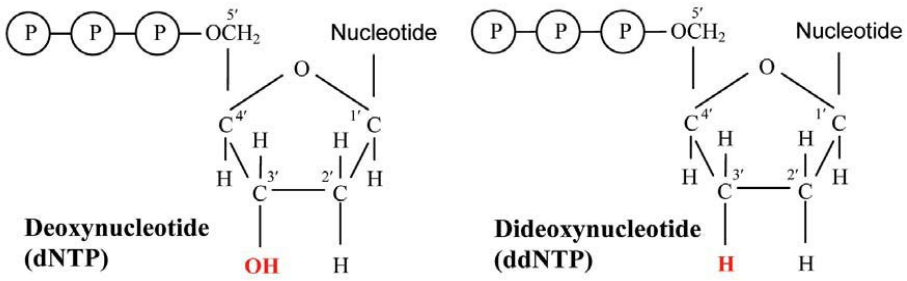


Figure 3.12 - An adapted illustration of the formula for a dNTP and a ddNTP

The formula for a deoxynucleotide (dNTP) is shown on the left and the formula for a dideoxynucleotide (ddNTP) on the right. The ddNTP has a hydrogen atom at the 3' carbon in the place of the -OH group as with the dNTP [Lander et al., 2001].

The sequences of the samples can then be ascertained by separating out the DNA fragments by size on a polyacrylamide gel, and establishing the last ddNTP that was incorporated into the strand, as each fragment being one base longer than the previous [Lander *et al*., 2001]. Sequencing software is then applied to analyse the DNA sequence.

3.5 - Construction of D-SpG$_{C2}$-1 Fc binding mutants

In order to study the various amino acids involved in the interaction of D-SpG$_{C2}$-1 with Fc, and any changes in the structure of SpG that may occur on formation of the complex, key amino acid residues, identified in both the major and minor regions of contact with the Fc C_H2 and C_H3 domains, were chosen for mutation. Mutations in corresponding positions to these residues and neighbouring residues were also considered. This complex is stabilised through a network of hydrogen bonds, and the importance and effect of these bonds to the stability and nature of the complex by removing these key residues was to be investigated.

3.5.1 - Selection of a suitable template for Fc mutants

The selection of a suitable template DNA to be used in the PCR for the proposed mutants was considered. To enable Fc binding to SpG to be studied, the wild type D-SpG$_{C2}$-1 *pQE*-30 template DNA was used. The wild-type template contains a unique tryptophan at

position 42 on β-strand 3, which was employed to report binding, as shown in figure 3.13.

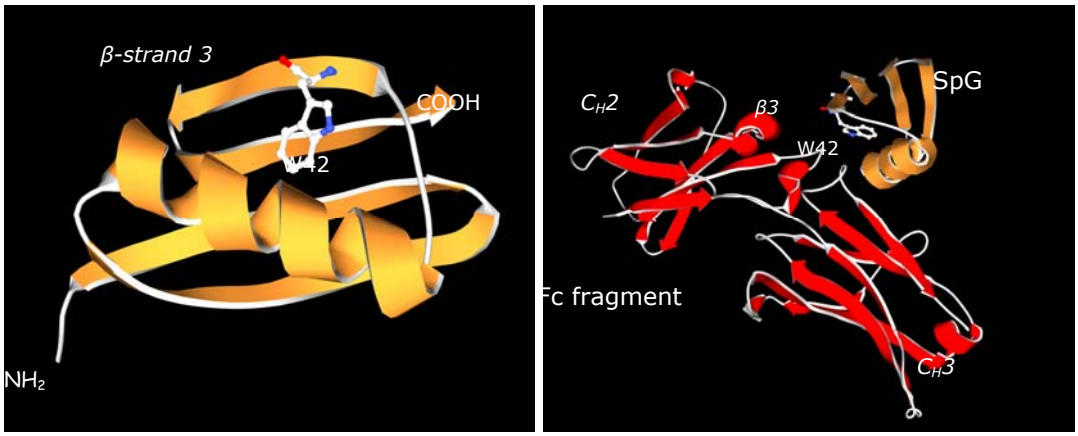


Figure 3.13 – Ribbon representations of an Ig-binding domain SpG illustrating the position of W42 residue and its involvement in the interaction with Fc

The three-dimensional structure of a single Ig-binding domain of SpG is presented in the first figure, illustrating the position of tryptophan 42 (W42) on β-strand 3. In complex with Fc, presented in the second figure, SpG binds the hinge region at the C_H2-C_H3 interface of the Fc fragment. The interaction site predominantly involves β-strand 3 of SpG, with important contact residues including tryptophan 42 (W42), which was employed to report binding [Walker et al., 1995].

3.5.2 - Selection of Fc contact residues for mutation

Various IgG-binding domains of SpG have been shown to bind Fc at the hinge region that connects the C_H2 and C_H3 domains, and the contact regions of this complex have been identified [Gronenborn *et al.*, 1991; Achari *et al*., 1992; Sauer-Eriksson *et al.*, 1995]. Results show that the interaction site with Fc lies within C-terminal part of the α-helix, the N-terminal part of β-strand 3, and the extended loop region connecting these two structures.

There are three residues of the C_H2 domain of Fc, which are involved in the interfacial interactions, namely Ile253, Ser254 and Gln311. In the C_H3 domain, there are two areas which contribute to the interface, namely Glu380 and Glu382 and the residues His433 to Gln438 [Sauer-Eriksson *et al.*, 1995]. These residues that interact with SpG are situated within loop regions of Fc. The residues in SpG that interact with Fc have been identified as Glu26, Lys27, Lys30, Gln31, Asn34, Asp39, Glu41 and Trp42. The nature of these interfacial interactions between Fc and SpG consists of a network of hydrogen bonds and

salt links, involving mainly charged and polar residues. It was decided to mutate a selection of these residues from both the major and minor contact regions, namely residues Glu26, Lys27, Gln31, Asn34, Glu41 and Trp42. An overview of the position of these Fc contact residues in SpG is shown below in figure 3.14.

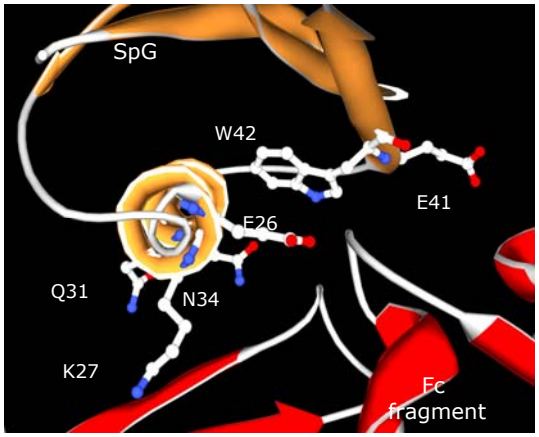


Figure 3.14 – A ribbon representation of the proposed residues involved in the interaction between D-SpG$_{C2}$-1 and Fc, selected for mutation

The proposed residues of D-SpG$_{C2}$-1 selected for mutation are presented, namely Glu26, Lys27, Gln31, Asn34, Glu41 and Trp42. These residues are located on β-strand 3 of SpG (shown in gold), and have been implicated in the interaction with the Fc (shown in red).

3.5.2.1 - Mutation of glutamate 26 to alanine (E26A)

One of two glutamate residues implicated in the major interfacial interaction of SpG with Fc, is situated on the C-terminal end of the α-helix, namely Glu26. This residue forms hydrogen bonds with the amide hydrogen of Fc residues Ile253 and Ser254. These Fc residues have been shown to be involved in this interaction with SpG, and are both located on the loop region between β-strand 2 and 3 of the C_H2 domain. Glu26 also forms part of a stabilising hydrogen network, with the carbonyl oxygen forming a hydrogen bond with the amide hydrogen of the side-chain of D-SpG_{C2}-1 residue lysine, Lys30.

Figure 3.15 illustrates the proposed E26A mutation and its position in complex with Fc, also demonstrating the effect that this mutation will have. It is proposed that by replacing glutamate 26 with alanine, the hydrogen bonds with the Fc residues and D-SpG_{C2}-1 residue will be lost, thus disrupting the hydrogen bond network and weakening the complex.

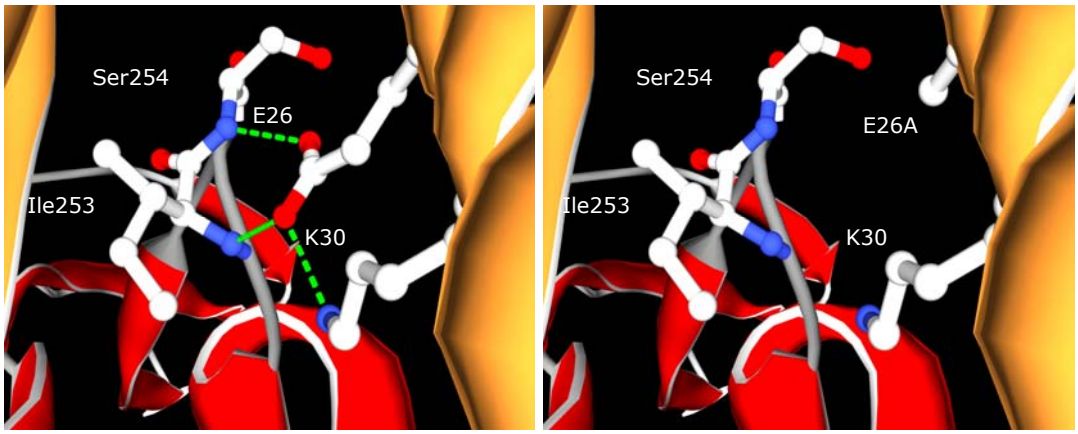


Figure 3.15 – Ribbon representations of the amino acid side-chains involved in the proposed mutant E26A

The proposed residue of D-SpG_{C2}-1 (shown in gold) selected for mutation is presented in the first figure, namely Glu26. This residue forms part of a hydrogen network, illustrated by green dotted lines, with Fc (shown in red) residues Ser254 and Ile253, and D-SpG_{C2}-1 residue Lys30. The second figure shows the effect of mutating the glutamate residue to an alanine.

3.5.2.2 - Mutation of lysine 27 to alanine (K27A)

The only lysine residue to be implicated in the interfacial interaction of SpG with Fc, is also situated on the C-terminal end of the α-helix next to Glu26, namely Lys27. This residue forms a single hydrogen bond with the carbonyl oxygen on the side-chain of the Fc residues Glu380, which is located on β-strand 3 of the C_H3 domain.

Figure 3.16 illustrates the proposed K27A mutation and its position in complex with Fc, also demonstrating the effect that this mutation will have on the local environment. It is proposed that by replacing lysine 27 with alanine, the hydrogen bond with the Fc residue will be lost.

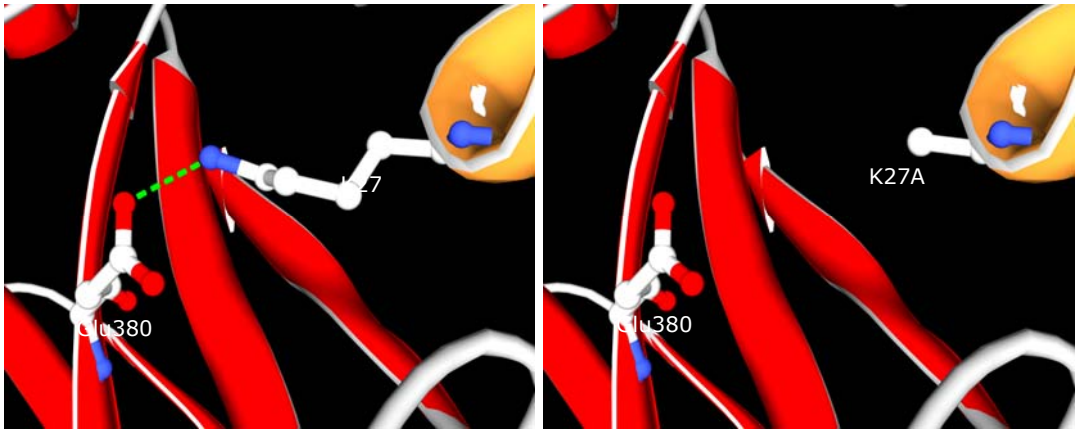


Figure 3.16 – Ribbon representations of the amino acid side-chains involved in the proposed mutant K27A

The proposed residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Lys27. This residue forms part of a hydrogen network, illustrated by green dotted lines, with Fc (shown in red) residue Glu380. The second figure shows the effect of mutating the lysine residue to an alanine.

3.5.2.3 - Mutation of glutamine 31 to alanine (Q31A)

Glutamine is also implicated in the major interfacial interaction of SpG with Fc, and is located around the middle of the α-helix, namely Gln31. The nitrogen atom of this residue's side-chain forms a hydrogen bond with the carbonyl oxygen of the side-chain of Fc residue Gln438. This Fc residue has been shown to be involved in this interaction with D-SpG$_{C2}$-1, and is located on β-strand 9 of the C$_H$3 domain. Gln31 also forms part of a stabilising hydrogen network, with its carbonyl oxygen forming a hydrogen bond with the carbonyl oxygen of the side-chain of D-SpG$_{C2}$-1 residue aspartate, Asp35.

Figure 3.17 illustrates the proposed Q31A mutation and its position in complex with Fc, also demonstrating the effect that this mutation will have. It is proposed that by replacing glutamine 31 with alanine, the hydrogen bonds with both the Fc residue and D-SpG$_{C2}$-1 residue will be lost.

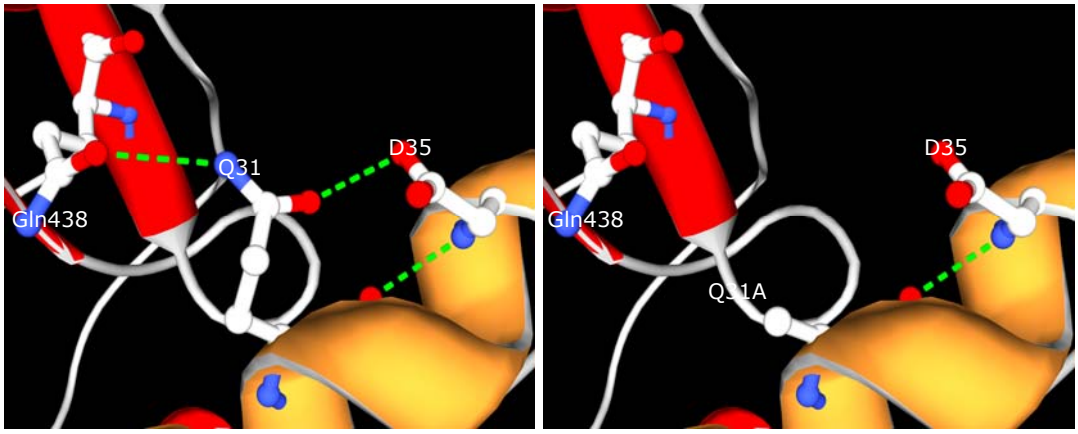


Figure 3.17 – Ribbon representations of the amino acid side-chains involved in the proposed mutant Q31A

The proposed residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Gln31. This residue forms part of a stabilising hydrogen network, illustrated by green dotted lines, with Fc (shown in red) residue Gln438, and D-SpG$_{C2}$-1 residue Asp35. The second figure shows the effect of mutating the glutamine residue to an alanine.

3.5.2.4 - Mutation of glutamine 31 to glutamate (Q31E)

It was also decided to mutate glutamine at position 31 to a glutamate residue.

Figure 3.18 illustrates the proposed Q31E mutation and its position in complex with Fc, also demonstrating the effect that this mutation will have. It is proposed that by replacing glutamine 31 with a glutamate residue, the hydrogen bond with the Fc residue will be lost, but the one with the D-SpG$_{C2}$-1 residue will remain intact.

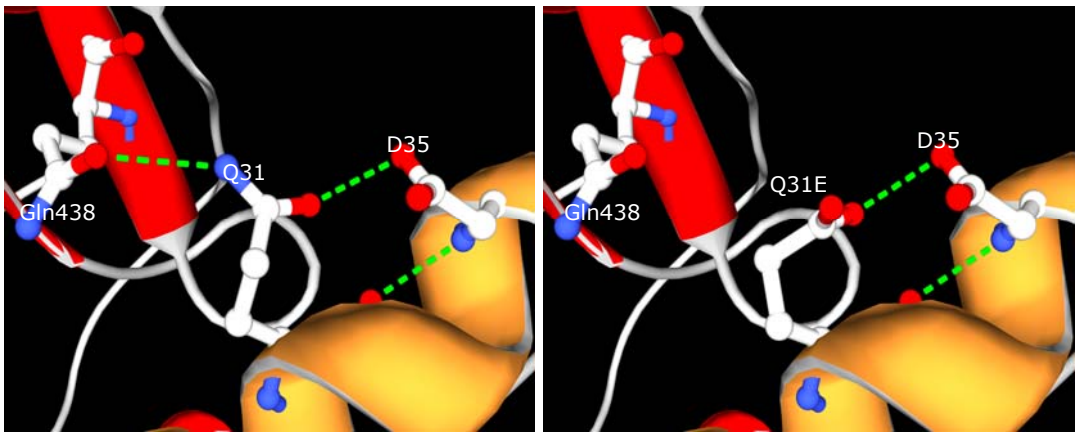


Figure 3.18 – Ribbon representations of the amino acid side-chains involved in the proposed mutant Q31E

The proposed residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Gln31. This residue forms part of a stabilising hydrogen network, illustrated by green dotted lines, with Fc (shown in red) residue Gln438, and D-SpG$_{C2}$-1 residue Asp35. The second figure shows the effect of mutating the glutamine residue to a glutamate.

3.5.2.5 - Mutation of glutamine 31 to histidine (Q31H)

It was also decided to mutate glutamine at position 31 to a histidine residue. Histidine was used to substitute glutamine in this mutant construct as it is also a polar amino acid, which tends to favour the formation of a helical structure. The residue often serves a role in stabilising the folded structures of proteins [Park *et al*., 1997].

Figure 3.19 illustrates the proposed Q31H mutation and its position in complex with Fc, also demonstrating the effect that this mutation will have. It is proposed that by replacing glutamine 31 with a histidine residue, the hydrogen bond with the Fc residue will be lost, but the one with the D-SpG$_{C2}$-1 residue will remain intact.

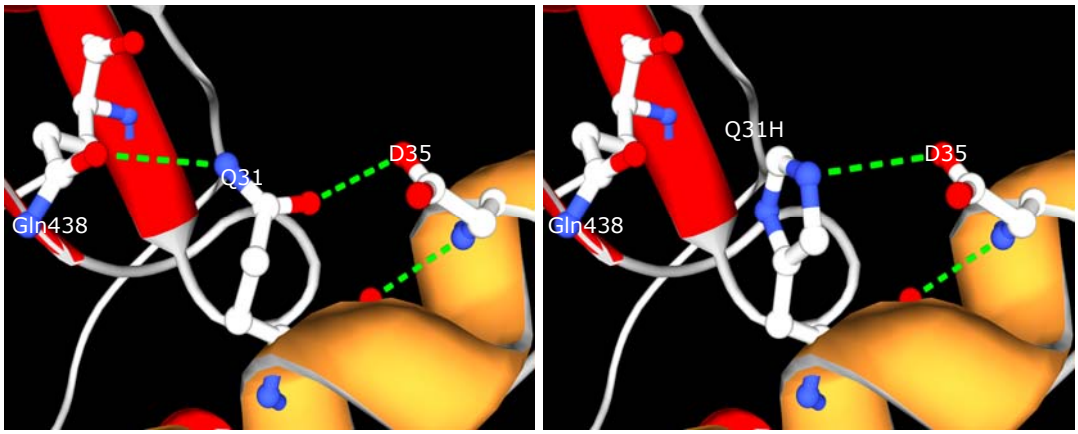


Figure 3.19 – Ribbon representations of the amino acid side-chains involved in the proposed mutant Q31H

The proposed residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Gln31. This residue forms part of a stabilising hydrogen network, illustrated by green dotted lines, with Fc (shown in red) residue Gln438, and D-SpG$_{C2}$-1 residue Asp35. The second figure shows the effect of mutating the glutamine residue to a histidine.

3.5.2.6 - Mutation of asparagine 34 to alanine (N34A)

Asparagine is implicated in the major interfacial interaction of SpG with Fc, and is located at the C-terminal end of the α-helix, namely Asn34. The oxygen atom of this residue's side-chain forms a hydrogen bond with the amide hydrogen of the side-chain of Fc residue Asn434. This Fc residue has been shown to be involved in this interaction with SpG, and is located on the loop region between β-strand 2 and 3 of the C_H3 domain.

Asn34 also forms part of a vast stabilising hydrogen network, with the same amide hydrogen forming a hydrogen bond with the carbonyl oxygen of the side-chain of D-SpG_{C2}-1 residue valine, Val38. Asparagine 34 continues this vast hydrogen bond network by forming hydrogen bonds with two other Fc residues. This time its amide hydrogen forms a hydrogen bond with the carbonyl oxygen of the backbone of Fc residue Tyr436, and also the carbonyl oxygen of the backbone of Fc residue His433. It was decided to mutate this asparagine residue to an alanine to disrupt the binding of the asparagine residue with its corresponding Fc residues through the hydrogen bond network.

Figure 3.20 illustrates the proposed N34A mutation and its position in complex with Fc, also demonstrating the effect that this mutation will have. It is proposed that by replacing asparagine 34 with an alanine residue, the hydrogen bonds with the three Fc residues will be lost.

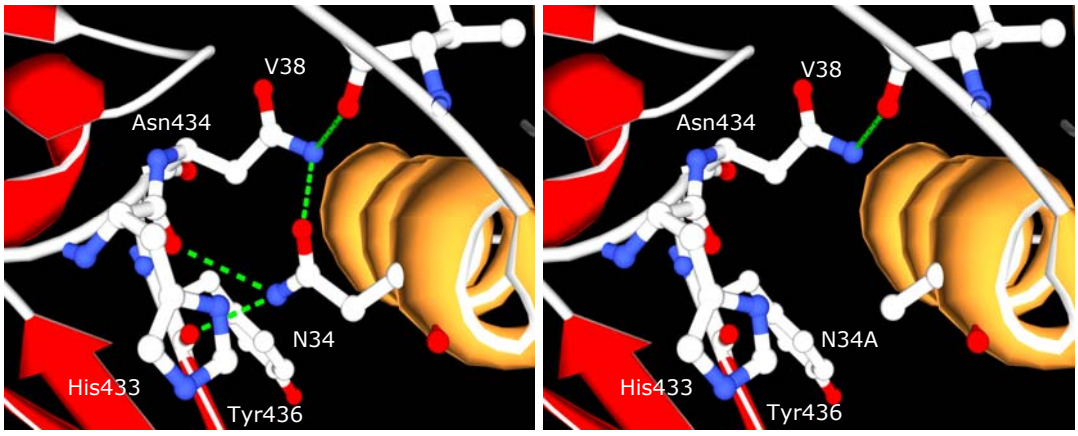


Figure 3.20 – Ribbon representations of the amino acid side-chains involved in the proposed mutant N34A

The proposed amino acid residue of D-SpG_{C2}-1 (shown in gold) selected for mutation is presented in the first figure, namely Asn34. This residue forms part of a stabilising hydrogen network, illustrated by green dotted lines, with Fc (shown in red) residues His433, Asn434 and Tyr436, and D-SpG_{C2}-1 residue Val38. The second figure shows the effect of mutating the asparagine residue to an alanine.

3.5.2.7 - Mutation of asparagine 34 to aspartate (N34D)

It was also decided to mutate asparagine at position 34 to an aspartate residue, as their structures are similar. The aim of this mutation was to affect the binding of the asparagine residue with its corresponding Fc residues through the hydrogen bond network.

Figure 3.21 illustrates the proposed N34D mutation and its position in complex with Fc, also demonstrating the effect that this mutation will have. It is proposed that by replacing asparagine 34 with an aspartate residue, the hydrogen bonds with Fc residues Asn434 and Tyr436 will be lost. The hydrogen bond with His433 is maintained, but it moves from the oxygen atom on the backbone to the nitrogen atom on the imidazole side-chain ring.

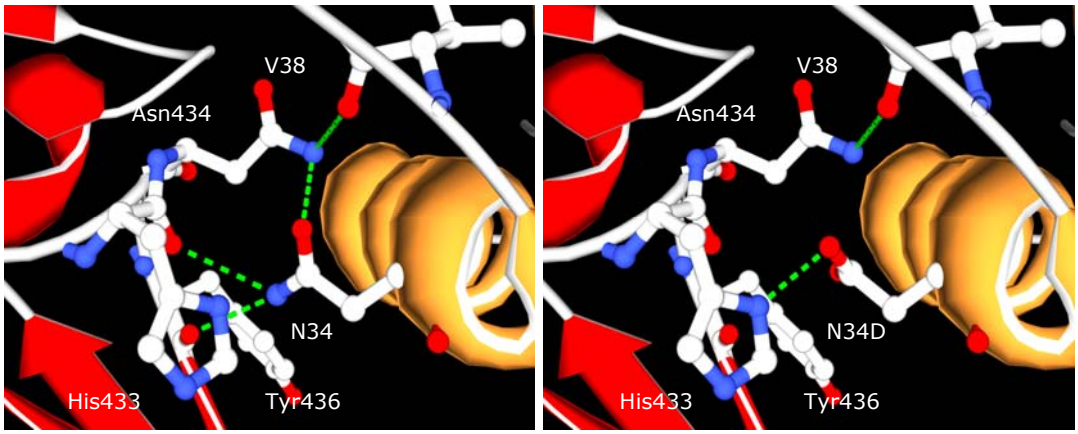


Figure 3.21 – Ribbon representations of the amino acid side-chains involved in the proposed mutant N34D

The proposed amino acid residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Asn34. This residue forms a hydrogen bond, illustrated by green dotted lines, with Fc (shown in red) residues His433, Asn434 and Tyr436, and D-SpG$_{C2}$-1 residue Val38. The second figure shows the effect of mutating the asparagine residue to an aspartate.

3.5.2.8 - Mutation of glutamate 41 to alanine (E41A)

The other glutamate residue implicated in the major interfacial interaction of SpG with Fc, is situated on the N-terminal end of β-strand 3, namely Glu41. This residue forms a hydrogen bond with the amide hydrogen of Fc residue Gln311. This Fc residue has been shown to be involved in this interaction with SpG, and is located on the loop region between β-strand 8 of the C_H2 domain. Alanine was chosen again to substitute the Glu41, as with Glu26 residue, to disrupt the binding of the glutamate residue with its corresponding Fc residue through the hydrogen bond.

Figure 3.22 illustrates the proposed E41A mutation and its position in complex with Fc, also demonstrating the effect that this mutation will have. It is proposed that by replacing glutamate 41 with an alanine residue, the hydrogen bond with the Fc residue is lost.

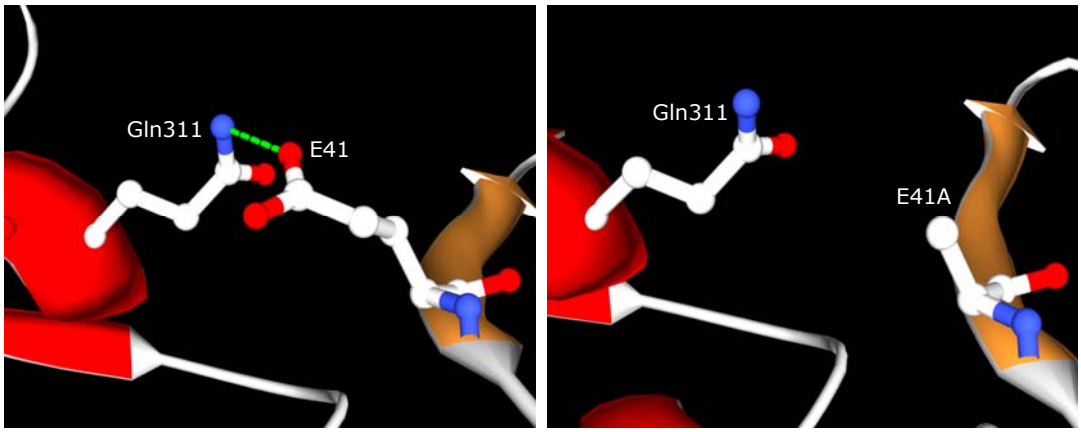


Figure 3.22 – Ribbon representations of the amino acid side-chains involved in the proposed mutant E41A

The proposed residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Glu41. This residue forms part a hydrogen bond, illustrated by green dotted lines, with Fc (shown in red) residue Gln311. The second figure shows the effect of mutating the glutamate residue to an alanine.

3.5.2.9 - Mutation of glutamate 41 to glutamine (E41Q)

It was also decided to mutate glutamate at position 41 to a glutamine residue, as their structures are similar. The aim of this mutation was to affect the binding of the glutamate residue with its corresponding Fc residue through its hydrogen bond.

Figure 3.23 illustrates the proposed E41Q mutation and its position in complex with Fc, also demonstrating the effect that this mutation will have. It is proposed that by replacing glutamate 41 with a glutamine residue, the hydrogen bond with the Fc residue will be lost.

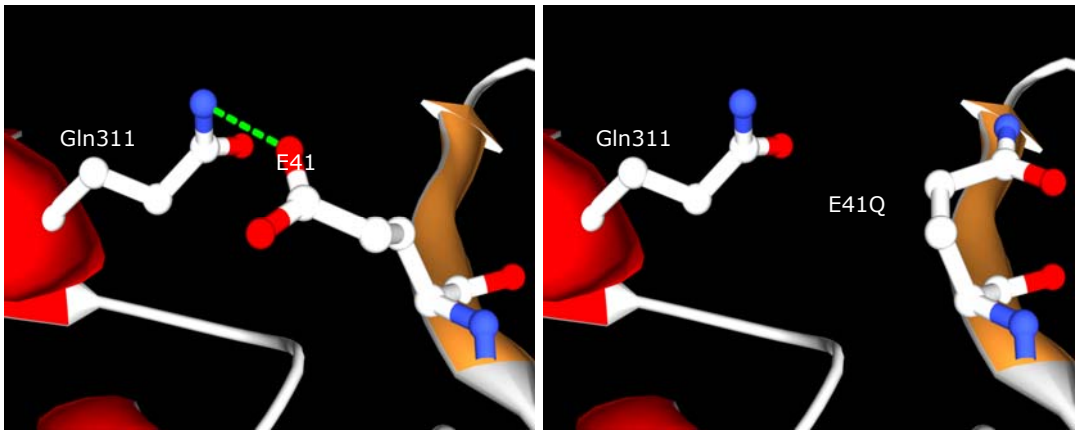


Figure 3.23 – Ribbon representations of the amino acid side-chains involved in the proposed mutant E41Q

The proposed residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Glu41. This residue forms a hydrogen bond, illustrated by green dotted lines, with Fc (shown in red) residue Gln311. The second figure shows the effect of mutating the glutamate residue to a glutamine residue.

3.5.2.10 - Mutation of tryptophan 42 to phenylalanine (W42F)

The unique tryptophan residue in the D-SpG$_{C2}$-1 gene has been implicated in the interfacial interaction with Fc is situated on the N-terminal end of β-strand 3, namely Trp42 [Walker *et al*., 1995]. However, this tryptophan residue does not appear to form any hydrogen bonds with neighbouring residues. Conversely, it is in close proximity with the hydrogen network centralised around asparagine residue 34. The aim of this W42F mutation was to generate a domain devoid of tryptophan and therefore, virtually non-fluorescent.

Figure 3.24 illustrates the proposed W42F mutation and its position in complex with Fc, also demonstrating the effect that this mutation will have. It is proposed that by replacing tryptophan 42 with a phenylalanine residue, the hydrogen bond network between Asn34 and the three Fc residues will not be affected. The interaction involves β-strand 3 of D-SpG$_{C2}$-1, with important contact residues including tryptophan 42 (W42), which has been clearly implicated in binding by the studies of Walker [1994]. Furthermore, the loss of the natural fluorophore permits the placement of other unique tryptophan residues that may be permit other binding reactions to be studied (see section 3.6.1).

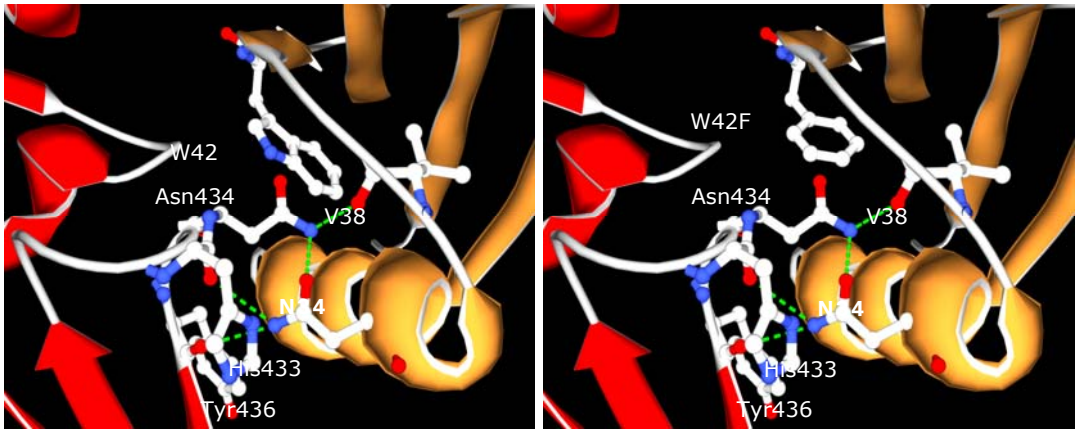


Figure 3.24 – Ribbon representations of the amino acid side-chains involved in the proposed mutant W42F

The proposed residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Trp42. This residue is in close proximity to a stabilising hydrogen network, illustrated by green dotted lines, with Fc (shown in red) residues His433, Asn434 and Tyr436, and D-SpG$_{C2}$-1 residue Val38. The second figure shows the effect of mutating the tryptophan residue to a phenylalanine.

3.5.2.11 - Mutation of tryptophan 42 to tyrosine (W42Y)

It was also decided to mutate tryptophan at position 42 to a tyrosine residue. It is already known that tryptophan's substitution with phenylalanine dramatically reduces the binding interaction of D-SpG$_{C2}$-1 with Fc [Walker, 1994]. Therefore, the aim of this W42Y mutation was to see if the replacement by the more polar tyrosine had the same effect.

Figure 3.25 illustrates the proposed W42Y mutation and its position in complex with Fc, also demonstrating the effect that this mutation will have. It is proposed that by replacing tryptophan 42 with a tyrosine residue, the surrounding hydrogen bond network with the three Fc residues is not affected. Modelling studies suggest that when tryptophan is mutated to a tyrosine a hydrogen bond is formed with Asn434 with the possibility that this interaction may restore Fc binding. Also, the fluorescence properties of the protein will be changed.

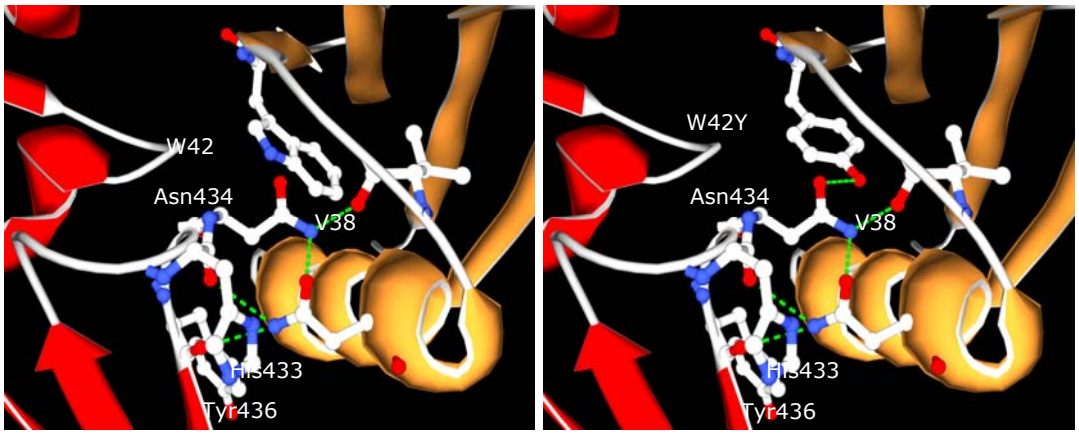


Figure 3.25 – Ribbon representations of the amino acid side-chains involved in the proposed mutant W42Y

The proposed residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Trp42. This residue is in close proximity to a stabilising hydrogen network, illustrated by green dotted lines, with Fc (shown in red) residues His433, Asn434 and Tyr436, and D-SpG$_{C2}$-1 residue Val38. The second figure shows the effect of mutating the tryptophan residue to a tyrosine.

3.5.2.12 - Mutation of cysteine 56 to alanine (C56A)

Cysteine has not been implicated in any interfacial interactions between SpG and Fc or Fab. However, it was decided to mutate this unique cysteine to an alanine to prevent cross-linking between D-SpG$_{C2}$-1 domains and also with Fab and Fc molecules. The unique cysteine residue is situated on the C-terminal end of β-strand 4, namely Cys56.

Figure 3.26 illustrates the position of the cysteine residue on D-SpG$_{C2}$-1.

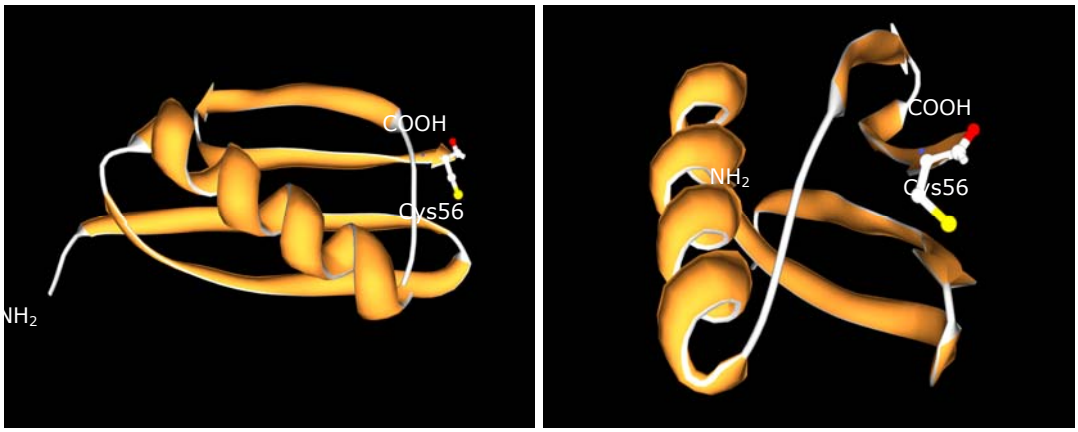


Figure 3.26 – Ribbon representations of D-SpG$_{C2}$-1, illustrating the position of the cysteine residue

The proposed amino acid residue of D-SpG$_{C2}$-1 selected for mutation is presented in the first figure, namely Cys56. This residue is located at the C-terminal end of the β-strand 4 of D-SpG$_{C2}$-1 (shown in gold). The second figure shows a different orientation of D-SpG$_{C2}$-1.

3.5.3 - Mis-match primers for Fc contact residue mutants

Mutagenesis of the key residues in D-SpG$_{C2}$-1, critical for the interaction with Fc, will identify the contribution of a specific amino acid to the biological function of SpG. For each mutant construct, two oligonucleotide primers were designed that were capable of effecting these desired mutations and complied with criteria stated in section 3.2.2.1. The sequence of these primers was complementary to the gene sequence in the region to be mutated, but with a single codon difference at the intended mutation site. All primers were sufficiently long enough to ensure site-specific annealing of the primer to the template DNA, and were produced in the 5’ to 3’ orientation, allowing them to bind to the DNA template produced in the cloning vector. In addition, codons preferred by *E.coli* were used wherever possible for the substituted amino acids. The final design for each oligonucleotide primer is presented in figure 3.27.

E26A – oligo#1: **5' –** GCT GCT ACC GCT **GCT** AAA GTT TTC AAA **– 3'**
E26A – oligo#2: **5' –** TTT GAA AAC TTT **AGC** AGC GGT AGC AGC **– 3'**

K33A – oligo#1: **5' –** CT ACC GCT GAA **GCT** GTT TTC AAA CAG TAC G **– 3'**
K33A – oligo#2: **5' –** C GTA CTG TTT GAA AAC **AGC** TTC AGC GG **– 3'**

Q31A – oligo#1: **5' –** A GTT TTC AAA **GCT** TAC GCT AAC GAC AA **– 3'**
Q31A – oligo#2: **5' –** ACC GTT GTC GTT AGC GTA **AGC** TTT GAA AA **– 3'**

Q31E – oligo#1: **5' –** T GAA AAA GTT TTC AAA **GAA** TAC GCT AA **– 3'**
Q31E – oligo#2: **5' –** ACC GTT GTC GTT AGC GTA **TTC** TTT GGA **– 3'**

Q31H – oligo#1: **5' –** CT GAA AAA GTT TTC AAA **CAT** TAC GCT A **– 3'**
Q31H – oligo#2: **5' –** ACC GTT GTC GTT AGC GTA **ATG** TTT GAA **– 3'**

N34A – oligo#1: **5' –** TC AAA CAG TAC GCT **GCT** GAC AAC GGT G **– 3'**
N34A – oligo#2: **5' –** C ACC GTT GTC **AGC** AGC GTA CTG TTT GA **– 3'**

N34D – oligo#1: **5' –** AAA CAG TAC GCT **GAC** GAC AAC GGT GTC **– 3'**
N34D – oligo#2: **5' –** GAC ACC GTT GTC **GTC** AGC GTA CTG TTT **– 3'**

E41A – oligo#1: **5' –** C GGT GTC GAC GGT **GCT** TGG ACC TAC G **– 3'**
E41A – oligo#2: **5' –** GTC GTA GGT CCA **AGC** ACC GTC GAC ACC **– 3'**

E41Q – oligo#1: **5' –** GTC GAC GGT **CAG** TGG ACC TAC GAC GAC GCT **– 3'**
E41Q – oligo#2: **5' –** GTC GTA GGT CCA **CTG** ACC GTC GAC ACC **– 3'**

W42F – oligo#1: **5' –** GTC GAC GGT GAA **TTC** ACC TAC GAC GAC **– 3'**
W42F – oligo#2: **5' –** GTC GTC GTA GGT **GAA** TTC ACC GTC GAC **– 3'**

W42Y – oligo#1: **5' –** C GAC GGT GAA **TAC** ACC TAC GAC GAC GCT **– 3'**
W42Y – oligo#2: **5' –** GTC GTC GTA GGT **GTA** TTC ACC GTC GAC **– 3'**

C56A – oligo#1: **5' –** ACG GTT ACC GAA **GCT** TAA TAG TGA AA **– 3'**
C56A – oligo#2: **5' –** AG CTT TCA CTA TTA **AGC** TTC GGT AAC C **– 3'**

Figure 3.27 – Mis-match primers for the production of D-SpG$_{C2}$-1 Fc mutants
For each mutant construct, two oligonucleotide primers were designed. Mutant codons are emboldened in red.

3.6 - Construction of D-SpG$_{C2}$-1 Fab binding mutants

In order to study the various amino acids involved in the interaction of D-SpG$_{C2}$-1 with Fab, and any changes in the structure of D-SpG$_{C2}$-1 that may occur on formation of the complex, key amino acid residues, identified in both the major and minor regions of contact with the Fab C_H1 domain, were chosen for mutation. Mutations in corresponding positions to these residues and neighbouring residues were also considered. This complex is stabilised through a network of hydrogen bonds, and the importance and effect of these bonds to the stability and nature of the complex by removing these key residues was to be investigated.

3.6.1 - Selection of a suitable template for Fab mutants

The selection of a suitable template DNA to be used in PCR mutagenesis for the proposed Fab mutants was considered. As previously mentioned, D-SpG$_{C2}$-1 possess a unique tryptophan residue at position 42, located on β-strand 3, which is known for its important role in Fc binding. Its substitution with phenylalanine dramatically reduces the binding interaction of D-SpG$_{C2}$-1 with Fc, with a 300-fold decrease in affinity [Walker *et al.*, 1995]. Therefore, by using the existing W42F D-SpG$_{C2}$-1 construct, Fab binding can be studied without the intercession of any Fc interactions. However, the removal of the tryptophan in the W42F mutation created a non-fluorescent protein, meaning Fab binding studies would not be possible using fluorescent spectroscopy.

The re-introduction of a tryptophan residue at an alternative position, however, would allow the effect of Fab mutations to be probed in fluorescent studies. Residue Glu14 is located on β-strand 2 of D-SpG$_{C2}$-1 and is adjacent to residues Thr15 and Thr16, forming two hydrogen bonds with Lys212 from Fab. The side-chain of Glu14 can be seen to be pointing away from the complex, so it was decided to incorporate a tryptophan residue at this position, as shown in figure 3.28, which is close to the proposed Fab binding site.

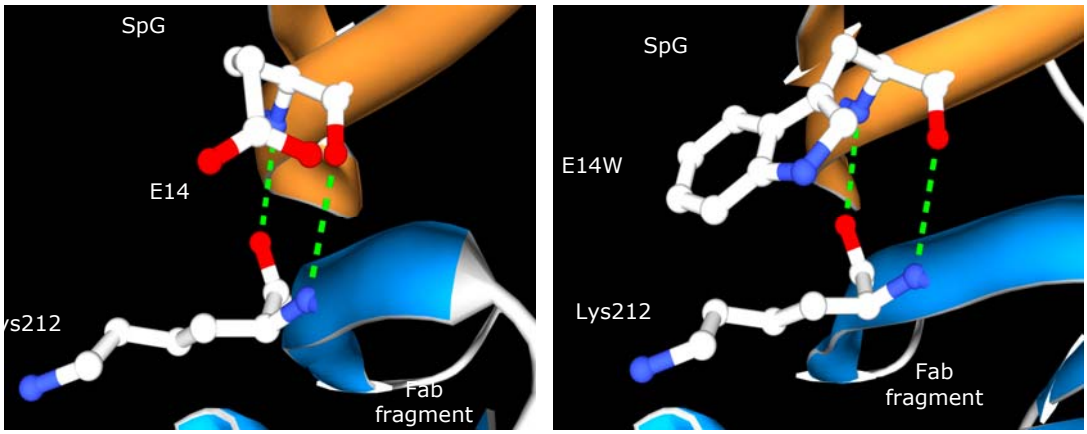


Figure 3.28 – Ribbon representations of D-SpG$_{C2}$-1 in complex with the C$_H$1 domain of Fab, illustrating the position of E14W mutant, as it acts as a reporter for Fab binding

In these complexes, the Ig-binding domain of D-SpG$_{C2}$-1 binds the C$_H$1 of the Fab fragment. As shown, E14W mutant is located on β-strand 2 of D-SpG$_{C2}$-1 (shown in gold). It forms part of a stabilising hydrogen network, illustrated by green dotted lines, with Lys212 on β-strand 7 of the C$_H$1 domain of Fab (shown in blue).

The Fc mutant W42F D-SpG$_{C2}$-1 *pQE*-30 was initially used as template DNA to make the E14W mutant, and the successive double mutant, E14W-W42F D-SpG$_{C2}$-1 *pQE*-30, was then employed as a template for subsequent Fab mutants. This mutant allows the binding of Fab to D-SpG$_{C2}$-1 to be studied exclusively, as its ability to bind Fc has been removed with the W42F mutant [Walker, 1994]. Subsequent mutations will be used to reduce the Fab binding affinity to D-SpG$_{C2}$-1, which can be reported back by

fluorescence studies from the tryptophan E14W residue, illustrated in figure 3.29.

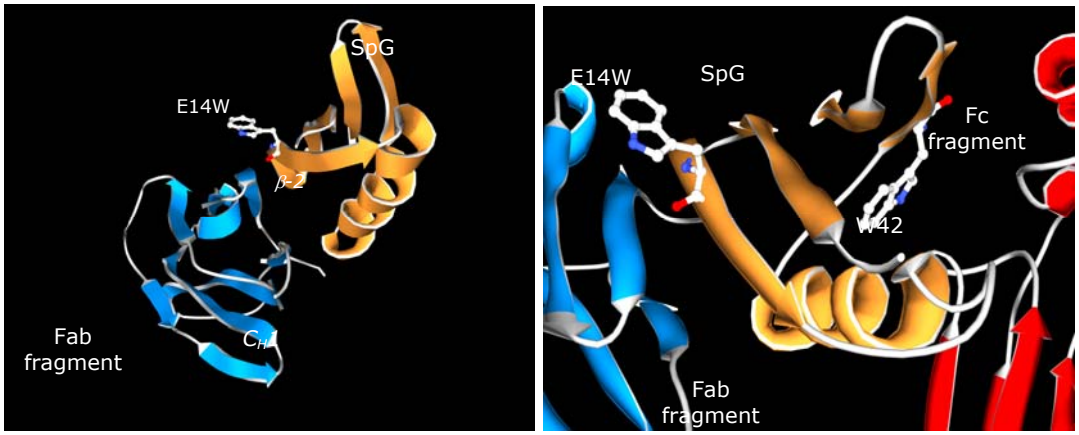


Figure 3.29 – Ribbon representations of D-SpG$_{C2}$-1 in complex with the C$_H$1 domain of Fab and Fc illustrating the position of tryptophan residues

In the first complex, the Ig-binding domain of D-SpG$_{C2}$-1 binds the C$_H$1 of the Fab fragment (shown in blue). E14W mutant is located on β-strand 2 of D-SpG$_{C2}$-1 (shown in gold) and can act as a reporter group solely for Fab binding. The second complex shows how D-SpG$_{C2}$-1 can bind both the C$_H$1 of the Fab fragment and Fc (shown in red) at the C$_H$2-C$_H$3 interface. Both tryptophan residues, E14W mutant and W42, can act as reporter groups for Fab and Fc binding, respectively.

Wild-type D-SpG$_{C2}$-1 *pQE*-30 DNA was also used as a template for the single mutant E14W, which was designed to be employed as a control for all experiments undertaken in this study. This E14W-W42 construct comprises two tryptophan residues and the second image in figure 3.40 illustrates their positions in D-SpG$_{C2}$-1. It is proposed that tryptophan at position 14 (E14W) may act as a reporter for the binding of D-SpG$_{C2}$-1 to the C$_H$1 domain of Fab, and that tryptophan at position 42 (W42) may act as a reporter for the binding of D-SpG$_{C2}$-1 to the C$_H$2-C$_H$3 interface of Fc.

3.6.2 - Selection of Fab contact residues for mutation

Single Ig-binding domains of SpG has been shown to bind to Fab by forming an anti-parallel interaction with the last β-strand of the C$_H$1 domain of Fab [Derrick and Wigley, 1992; Lian *et al.*, 1994; Derrick and Wigley, 1994; Sauer-Eriksson *et al.*, 1995]. SpG's interaction with Fab lies predominantly within β-strand 2, but also involves a minor region at the C-terminal end of the α-helix [Derrick and Wigley, 1994]. The amino acid residues that form the major contact regions have been identified, and the site of interaction spans residues Lys9 to Thr16 from β-strand 2 on SpG and Ser209 to Lys216 from the 7th β-strand of the C$_H$1 domain of Fab. In addition, residues Tyr32 to Gly37 from the C-terminal end of the α-helix of SpG and Pro125 to Tyr129 on the 1st β-strand of the C$_H$1 domain of Fab constitute the minor contact region [Derrick and Wigley, 1994].

The positions of these residues are illustrated in figure 3.30. The first figure shows how

these two sites of contact are stacked together. The atoms involved in these interactions can be principally found on the backbone of the C_H1 domain of the IgG, but include both backbone and side-chain atoms from the binding domain of SpG, involving hydrogen bonds and van der Waal forces, but no salt bridges. The complex is stabilised through a network of hydrogen bonds contributed by the last β-strand of the C_H1 domain to the second β-strand in SpG, and is further stabilised by the association of exposed non-polar residues from Fab and SpG, providing a continuous hydrophobic core, which is shielded from solvent [Derrick and Wigley, 1992].

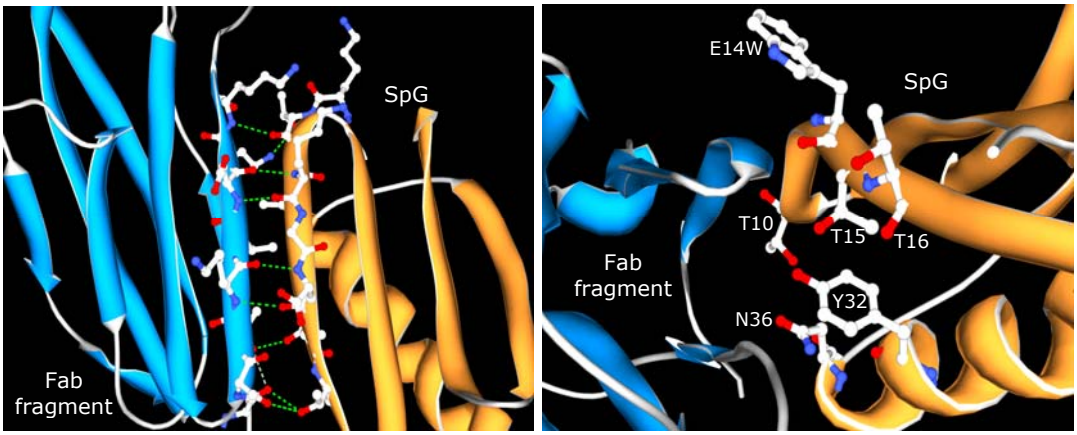


Figure 3.30 – Ribbon representations of the contact residues involved in the interaction between SpG and Fab

SpG has been shown to bind to Fab by forming an anti-parallel interaction with the last β-strand of the C_H1 domain of Fab. The site of interaction spans residues Lys9 to Thr16 from β-strand 2 on SpG and Ser209 to Lys216 from the 7th β-strand of the C_H1 domain of Fab. In addition, residues Tyr32 to Gly37 from the C-terminal end of the α-helix of SpG and Pro125 to Tyr129 on the 1st β-strand of the C_H1 domain of Fab constitute the minor contact region [Derrick and Wigley, 1994].

Using this information, six residues were selected to be mutated and used to study and characterise the binding affinity of Fab to D-SpG$_{C2}$-1, which can be reported back by fluorescence studies from the tryptophan E14W residue. A selection of residues from both the major and minor contact regions, namely residues Thr10, Glu14, Thr15, Thr16, Tyr32 and Asn36 were chosen for their side-chain involvement in the interaction with Fab. An overview of the position of these Fab contact residues in D-SpG$_{C2}$-1 is shown in the second image in figure 3.30.

3.6.2.1 - Mutation of threonine 10, 15 and 16 to alanine (T10A, T15A and T16A)

Three threonine β-strand 2 amino acid residues were selected to be mutated, namely Thr10, Thr15 and Thr16. These threonine amino acids appear to play important roles in the stabilising hydrogen bond network and have been shown to constitute the major contact region. Threonine 10 forms a hydrogen bond with Fab residue Lys215 located on the 7th β-strand on the C_H1 domain. Threonine 15 forms a hydrogen bond with Ser209

and Thr211, with threonine 16 also forming hydrogen bonds with Ser209 and in addition, with Ser210. It was decided to mutate each of these residues to alanine.

The aim of this mutation was to disrupt the extensive hydrogen bond network involved in the binding of the three threonine residues with their corresponding Fab residues, illustrated in figure 3.31. This may disrupt the high stability of D-SpG$_{C2}$-1, which is thought to be linked to such extensive interactions [Bouvignies *et al*., 2005]. The binding interface, which usually consists of protein surfaces as with D-SpG$_{C2}$-1 binding with either Fc or Fab, is generally more hydrophilic than the interior, and studies have shown that these binding interfaces tend to form more hydrogen bonds than protein interiors [Xu *et al*., 1997].

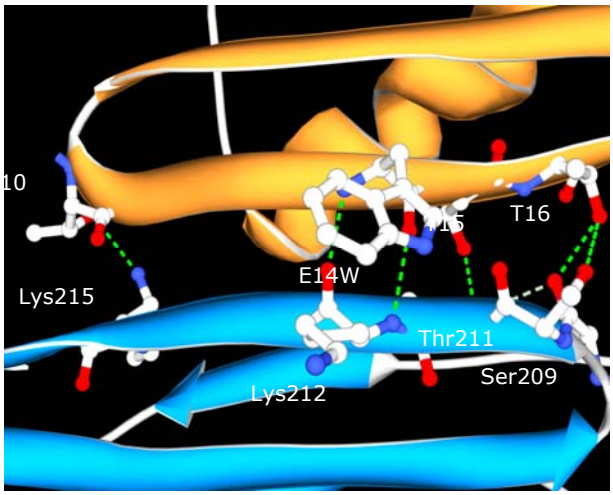


Figure 3.31 – Ribbon representation of the three threonine residues involved in the interaction between D-SpG$_{C2}$-1 and Fab

The three threonine residues of D-SpG$_{C2}$-1 selected for mutation is presented, namely Thr10, T15 and T16. These residues are located on β-strand 2 of D-SpG$_{C2}$-1 (shown in gold), and forms part of a hydrogen network, illustrated by green dotted lines, with Lys215, Thr211 and Ser209 respectively. These are located on β-strand 7 of the C$_H$1 domain of Fab (shown in blue).

Alanine was chosen to substitute all three threonine residues to disrupt the hydrogen bond network. Successive figures 3.32, 3.33 and 3.34 illustrate the proposed T10A, T15A and T16A mutations and their positions in the complex with Fab respectively. The figures also demonstrate that these substitutions will remove hydrogen bonds to Fab.

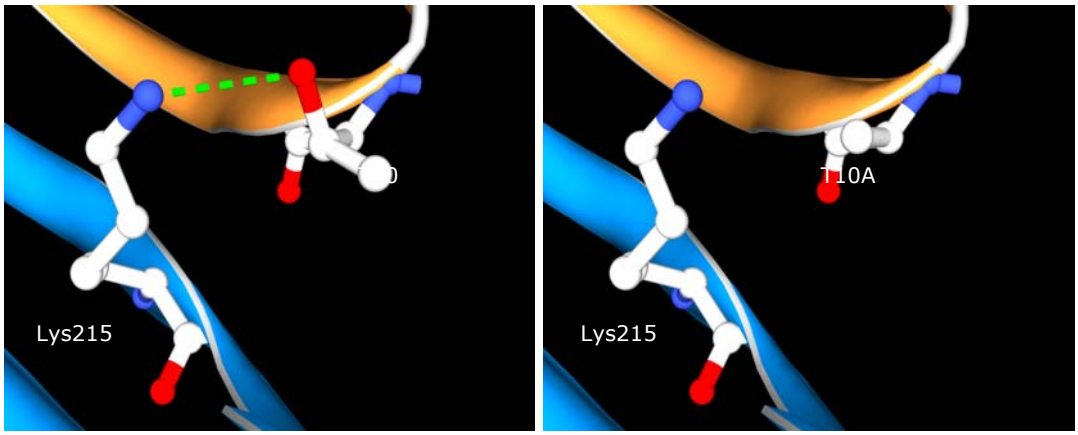


Figure 3.32 – Ribbon representations of the amino acid side-chains involved in the proposed mutant T10A

The proposed amino acid residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Thr10. This residue forms a hydrogen bond, illustrated by a green dotted line, with Fab (shown in blue) residue Lys215. The second figure shows the effect of mutating the threonine residue to an alanine.

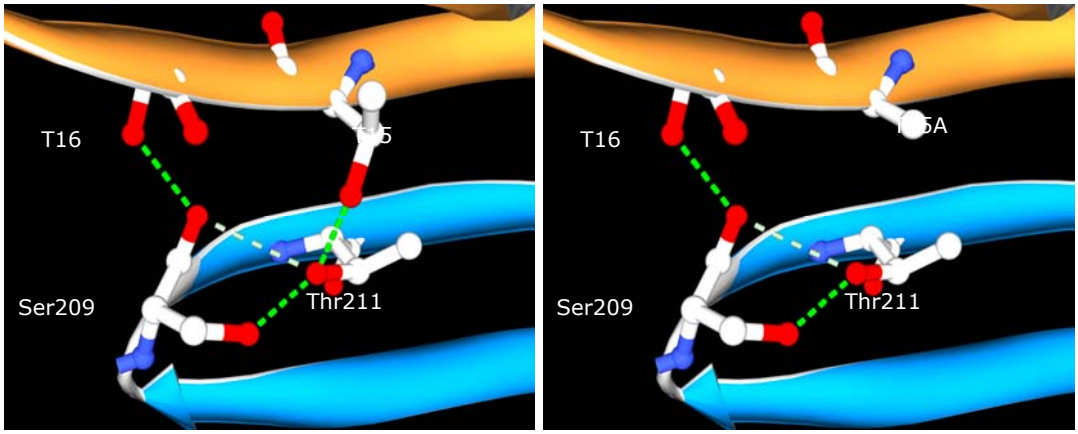


Figure 3.33 – Ribbon representations of the amino acid side-chains involved in the proposed mutant T15A

The proposed amino acid residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Thr15. This residue forms part of a hydrogen network, illustrated by green dotted lines, with Fab (shown in blue) residue. The second figure shows the effect of mutating the threonine residue to an alanine.

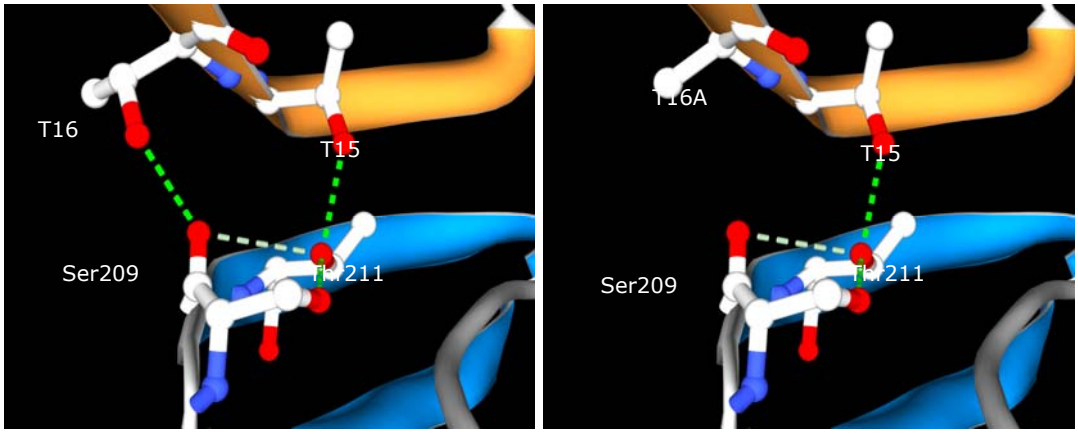


Figure 3.34 – Ribbon representations of the amino acid side-chains involved in the proposed mutant T16A

The proposed amino acid residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Thr16. This residue forms part of a hydrogen network, illustrated by green dotted lines, with Fab (shown in blue) residue Ser209. The second figure shows the effect of mutating the threonine residue to an alanine.

Figure 3.35 illustrates the proposed double mutant T15A-T16A mutation and its position in the complex with Fab respectively. This double mutation will remove two hydrogen bonds to Fab simultaneously.

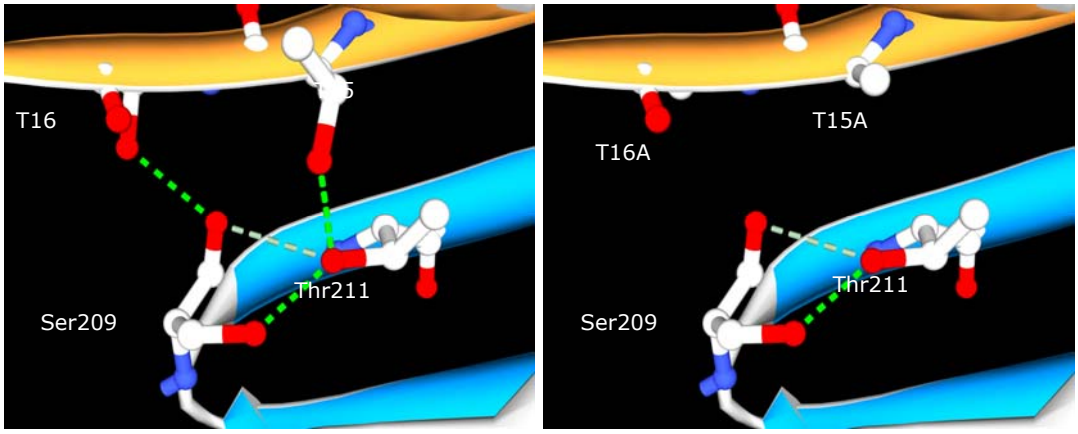


Figure 3.35 – Ribbon representations of the amino acid side-chains involved in the proposed double mutant T15A-T16A

The proposed amino acid residues of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Thr15 and Thr16. These residues form part of a hydrogen network, illustrated by green dotted lines, with Fab (shown in blue) residues Ser209 and Thr211. The second figure shows the effect of mutating both threonine residues to alanine.

3.6.2.2 - Mutation of threonine 10, 15 and 16 to serine (T10S, T15S and T16S)

It was also decided to mutate these three threonine residues at positions 10, 15 and 16 to serine residues, to effectively reduce the length of the side chain by one methyl group. Successive figures 3.36, 3.37 and 3.38 illustrate the proposed T10S, T15S and T16S mutations and their positions in the complex with Fab respectively, also demonstrating the effect that these mutations will have. It is possible that by replacing threonine 10 with serine, the hydrogen bond with its corresponding Fab residue may remain intact, but may be in a different orientation.

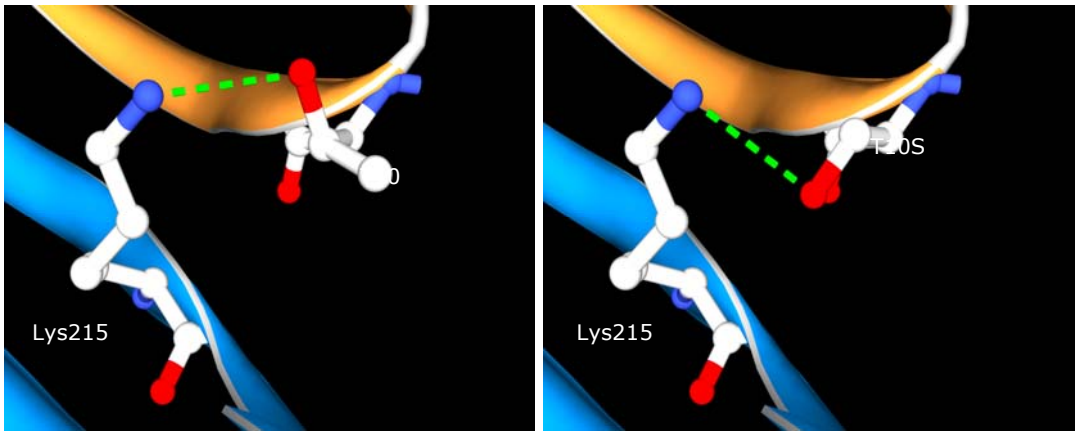


Figure 3.36 – Ribbon representations of the amino acid side-chains involved in the proposed mutant T10S

The proposed amino acid residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Thr10. This residue forms a hydrogen bond, illustrated by a green dotted line, with Fab (shown in blue) residue Lys215. The second figure shows the effect of mutating the threonine residue to serine.

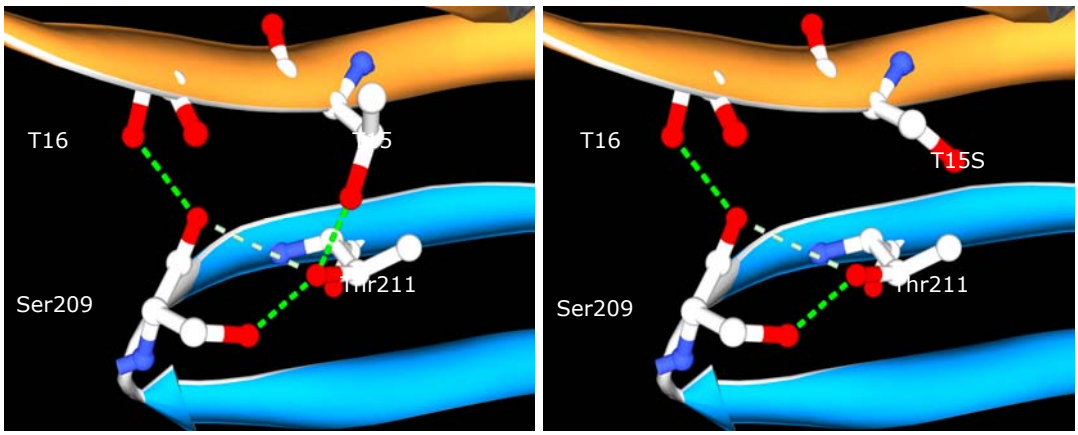


Figure 3.37 – Ribbon representations of the amino acid side-chains involved in the proposed mutant T15S

The proposed amino acid residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Thr15. This residue forms a hydrogen bond, illustrated by a green dotted line, with Fab (shown in blue) residue Lys215. The second figure shows the effect of mutating the threonine residue to serine.

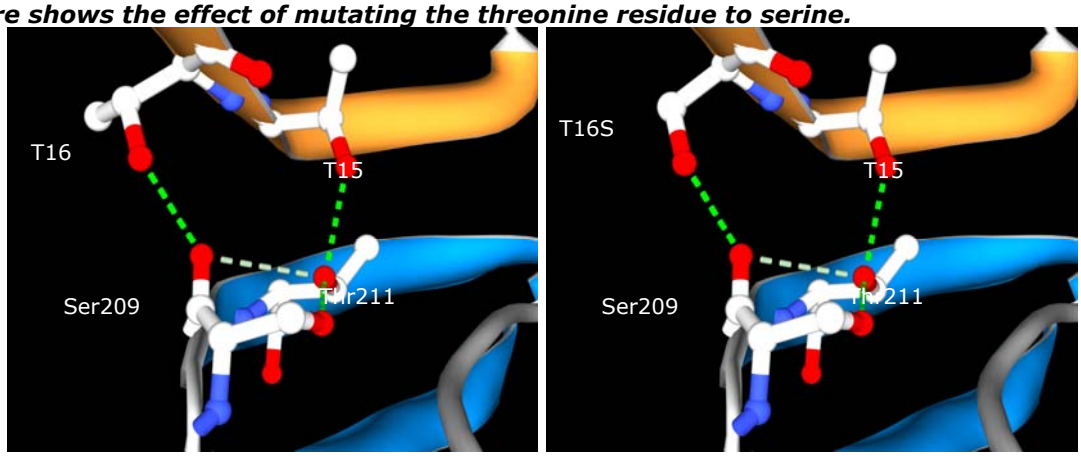


Figure 3.38 – Ribbon representations of the amino acid side-chains involved in the proposed mutant T16S

The proposed amino acid residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is

presented in the first figure, namely Thr16. This residue forms a hydrogen bond, illustrated by a green dotted line, with Fab (shown in blue) residue Lys215. The second figure shows the effect of mutating the threonine residue to serine.

3.6.2.3 - Mutation of tyrosine 32 to phenylalanine (Y32F)

The presence of the polar hydroxyl group in the side-chain enables tyrosine to be always involved in hydrogen bonds when it is found in the interior of proteins [Chen *et al*., 2005]. This also applies to D-SpG$_{C2}$-1, as all three tyrosine residues are involved in the hydrophobic core. Tyr32 forms part of a stabilising hydrogen network, with its backbone oxygen atom forming hydrogen bonds with both the amide hydrogen atom of the backbone and the side chain amide hydrogen of D-SpG$_{C2}$-1 residue Asn36.

This asparagine residue in turn forms a hydrogen bond with the nitrogen atom on the backbone of Fab residue Val128. The hydrogen bond network surrounding this residue is presented in figure 3.39.

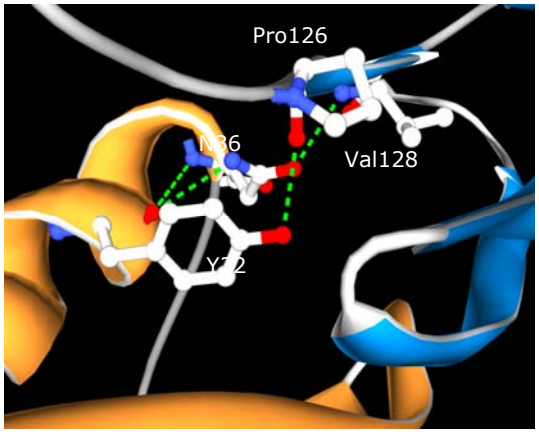


Figure 3.39 – A ribbon representation of the amino acid side-chains and hydrogen bond network surrounding the proposed mutant Y32F

The proposed amino acid residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented, namely Tyr32. This residue is located at the C-terminal end of the α-helix of D-SpG$_{C2}$-1, and forms part of an extensive hydrogen bond network, illustrated by green dotted lines, with Pro126. Tyr32 also forms backbone hydrogen bonds with Asn36, which in turn forms hydrogen bonds with Val128. Both Fab residues are located on β-strand 1 of the C$_H$1 domain (shown in blue).

Tyrosine 32 is also implicated in the minor interfacial interaction of SpG with Fab, and is located at the C-terminal end of the α-helix. The tyrosine's side-chain forms a hydrogen bond with the oxygen atom of the backbone of Fab residue Pro126 located on the first β-strand of the C$_H$1 domain. It was therefore decided to mutate this tyrosine residue to a phenylalanine to explore the consequence on the binding interaction.

Figure 3.40 illustrates the proposed Y32F mutation and its position in complex with Fab.

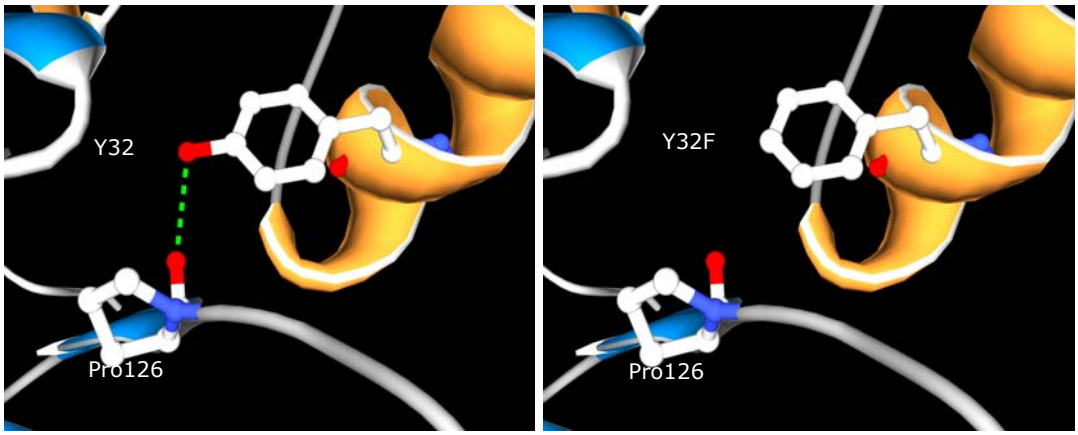


Figure 3.40 – Ribbon representations of the amino acid side-chains involved in the proposed mutant Y32F

The proposed amino acid residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Try32. This forms part of a hydrogen bond network, illustrated by a green dotted line, with Fab (shown in blue) residue Pro126. The second figure shows the effect of mutating the tyrosine residue to phenylalanine.

3.6.2.4 - Mutation of asparagine 36 to alanine (N36A)

Asparagine 36 has been implicated in the minor interfacial interaction of SpG with Fab, and is situated on the C-terminal end of the α-helix. The oxygen atom of this residue's side-chain forms a hydrogen bond with the amide hydrogen atom on the backbone of Fab residue Val128. This Fab residue is located on the first β-strand of the C_H1 domain. Asn36 also forms part of a stabilising hydrogen network, with its side-chain amide and its backbone amide each forming a hydrogen bond with the oxygen atom of the backbone of D-SpG_{C2}-1 residue Tyr32. This tyrosine residue in turn completes the hydrogen network, with its side-chain forming a hydrogen bond with Fab residue, Pro126. It was decided to mutate this asparagine residue to an alanine.

Figure 3.41 illustrates the proposed N36A mutation and its position in complex with Fab, also demonstrating the effect that this mutation will have.

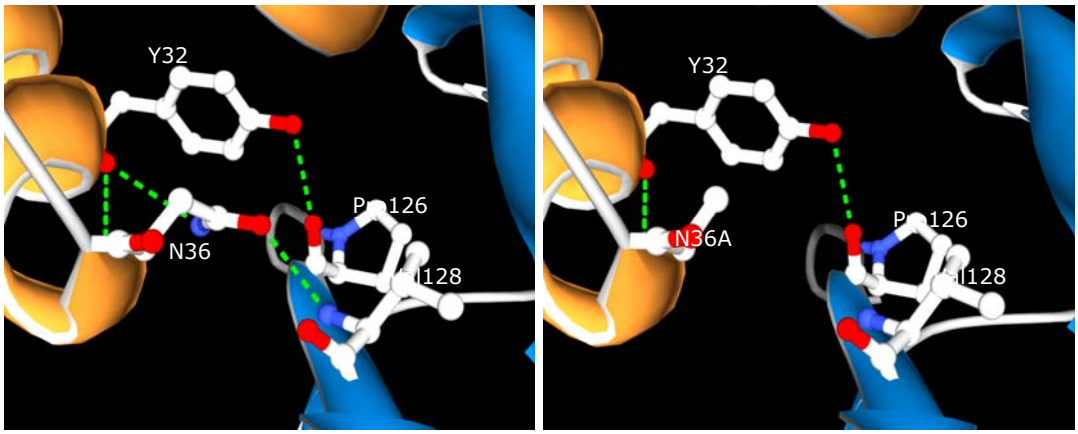


Figure 3.41 – Ribbon representations of the amino acid side-chains involved in the proposed mutant N36A

The proposed amino acid residue of D-SpG_{C2}-1 (shown in gold) selected for mutation is presented in the first figure, namely Asn36. This residue forms part of a hydrogen bond network, illustrated by green dotted lines, with Fab (shown in blue) residue Val128 and D-SpG_{C2}-1 residue Tyr32. The second figure shows the effect of mutating the asparagine residue to an alanine.

3.6.2.5 - Mutation of asparagine 36 to aspartic acid (N36D)

It was also decided to mutate asparagine 36 to the structurally conservative aspartate residue. Figure 3.42 illustrates the proposed N36D mutation and its position in complex with Fab, also demonstrating the effect that this mutation will have. In this case it is proposed that the mutation will not affect the hydrogen bond with its corresponding Fab residue.

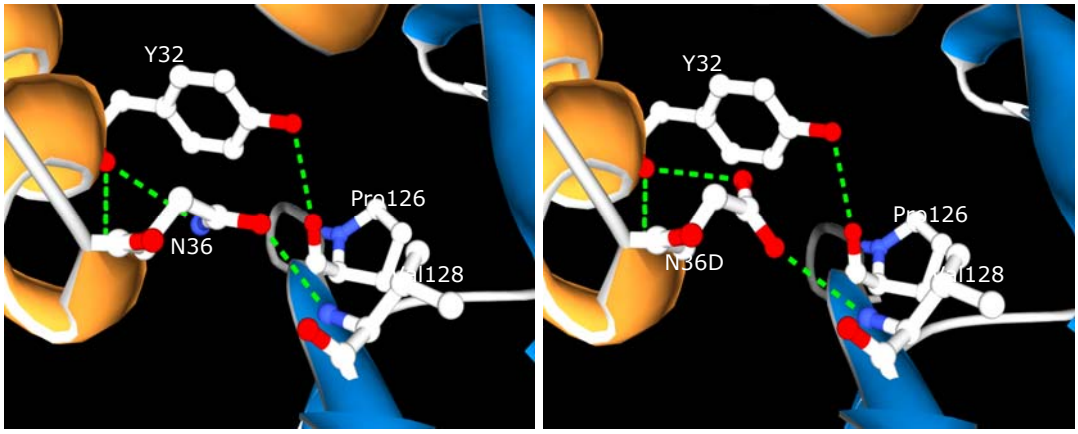


Figure 3.42 – Ribbon representations of the amino acid side-chains involved in the proposed mutant N36D

The proposed amino acid residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Asn36. This residue forms part of a hydrogen bond network, illustrated by green dotted lines, with Fab (shown in blue) residue Val128 and D-SpG$_{C2}$-1 residue Tyr32. The second figure shows the effect of mutating the asparagine residue to an aspartate.

3.6.2.6 - Mutation of asparagine 36 to glutamine (N36Q)

It was also decided to mutate asparagine at position 36 to a glutamine residue which retains an amide group on the side-chain although the latter has one more methylene group in it. Figure 3.43 illustrates the proposed N36Q mutation and its position in complex with Fab, also demonstrating the effect that this mutation will have. It is proposed that by replacing asparagine 36 with glutamine, the hydrogen bond with its corresponding Fab residue will be lost.

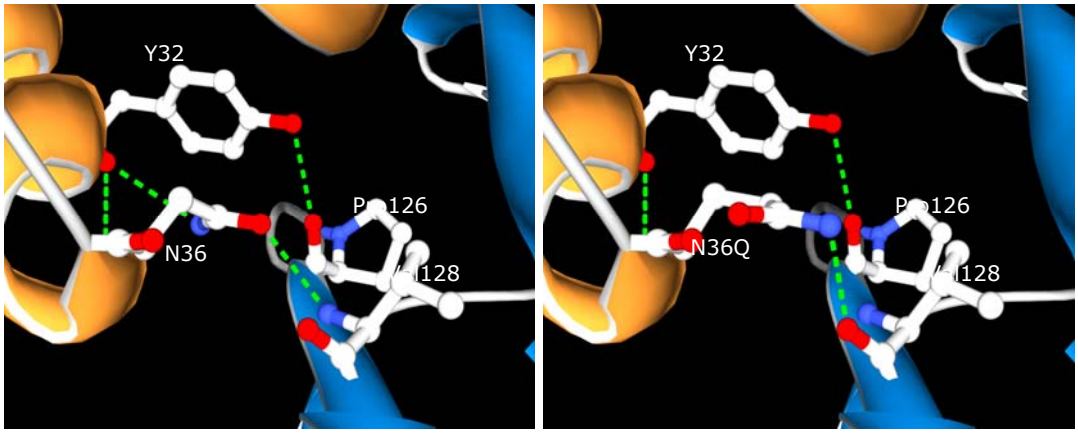


Figure 3.43 – Ribbon representations of the amino acid side-chains involved in the proposed mutant N36Q

The proposed amino acid residue of D-SpG$_{C2}$-1 (shown in gold) selected for mutation is presented in the first figure, namely Asn36. This residue forms part of a hydrogen bond network, illustrated by green dotted lines, with Fab (shown in blue) residue Val128 and D-SpG$_{C2}$-1 residue Tyr32. The second figure shows the effect of mutating the asparagine residue to a glutamine.

3.6.3 - Mis-match primers for Fab mutant D-SpG$_{C2}$-1 constructs

Two oligonucleotide primers were designed that were capable of effecting these desired mutations and complied with criteria stated in section 3.2.2.1. The final design for each oligonucleotide primer is presented in figure 3.44.

T10A – oligo#1: **5' –** ATC AAC GGT AAA **GCT** CTG AAA GGT **TGG – 3'**
T10A – oligo#2: **5' –** GGT GGT **CCA** ACC TTT CAG **AGC** TTT ACC **– 3'**

T10S – oligo#1: **5' –** ATC AAC GGT AAA **AGT** CTG AAA GGT **TGG – 3'**
T10S – oligo#2: **5' –** GGT GGT **CCA** ACC TTT CAG **ACT** TTT ACC **– 3'**

E14W – oligo#1: **5' –** TG AAA GGT **TGG** ACC ACC ACC GAA GCT G **– 3'**
E14W – oligo#2: **5' –** GT GGT GGT **CCA** ACC TTT CAG GGT TTT A **– 3'**

T15A – oligo#1: **5' –** AA GGT **TGG GCT** ACC ACC GAA GCT GTT G **– 3'**
T15A – oligo#2: **5' –** TC GGT GGT **AGC CCA** ACC TTT CAG GGT T **– 3'**

T15S – oligo#1: **5' –** AA GGT **TGG AGT** ACC ACC GAA GCT GTT G **– 3'**
T15S – oligo#2: **5' –** TC GGT GGT **ACT CCA** ACC TTT CAG GGT T **– 3'**

T16A – oligo#1: **5' –** GT **TGG** ACC **GCT** ACC GAA GCT GTT GAC G **– 3'**
T16A – oligo#2: **5' –** GC TTC GGT **AGC** GGT **CCA** ACC TTT CAG G **– 3'**

T16S – oligo#1: **5' –** GT **TGG** ACC **AGT** ACC GAA GCT GTT GAC G **– 3'**
T16S – oligo#2: **5' –** GC TTC GGT **ACT** GGT **CCA** ACC TTT CAG G **– 3'**

T15A-T16A – oligo#1: **5' –** AA GGT **TGG GCT GCT** ACC GAA GCT GTT G **- 3'**
T15A-T16A – oligo#2: **5' –** GC TTC GGT **AGC AGC CCA** ACC TTT CAG G **– 3'**

Y32F – oligo#1: **5' –** TC AAA CAG **TTC** GCT AAC GAC AAC GGT G **– 3'**
Y32F – oligo#2: **5' –** TC GTT AGC **GAA** CTG TTT GAA AAC TTT T **– 3'**

N36A – oligo#1: **5' –** CT AAC GAC **GCT** GGT GTC GAC GGT GAA T **– 3'**
N36A – oligo#2: **5' –** TC GAC ACC **AGC** GTC GTT AGC GTA CTG T **– 3'**

N36D – oligo#1: **5' –** TAC GCT AAC GAC **GAC** GGT GTC GAC GGT GAA T **– 3'**
N36D – oligo#2: **5' –** ACC GTC GAC ACC **GTC** GTC GTT AGC GTA CTG **– 3'**

N36Q – oligo#1: **5' –** TAC GCT AAC GAC **CAG** GGT GTC GAC GGT GAA T **– 3'**
N36Q – oligo#2: **5' –** ACC GTC GAC ACC **CTG** GTC GTT AGC GTA **– 3'**

Figure 3.44 – Mis-match primers for the production of D-SpG$_{C2}$-1 Fab mutants

For each mutant construct, two oligonucleotide primers were designed. Mutant codons are emboldened in red and the E14W mutation present in all mutations, as the reporter group for fluorescence studies, is in bold.

3.7 - Production of D-SpG$_{C2}$-1 Fc and Fab mutant constructs

PCR-based site-directed mutagenesis of the wild-type direct expression construct gene, D-SpG$_{C2}$-1 *pQE*-30, was performed using two-sided Splicing by Overlap Extension (SOE-ing) for the production of all D-SpG$_{C2}$-1 mutations. This was achieved by using the specifically designed oligonucleotides, presented in figures 3.26 and 3.44, and two flanking primers *pQE*-30 forward and *pQE*-30 reverse. The final design for each of the flanking oligonucleotide primer is presented below in figure 3.45.

***pQE*-30 forward oligo:** **5' –** AAA ATA GGC GTA TCA CGA GGC **– 3'**

***pQE*-30 reverse oligo:** **5' –** AGC TGA ACG GTC TGG TTA TAG GTA CA **– 3'**

Figure 3.45 – Flanking primers used for the production of D-SpG$_{C2}$-1 mutants

For each mutant construct, two flanking oligonucleotide primers were designed. All primers were sufficiently long enough to ensure annealing of the primer to the cloning vector DNA, and were produced in the 5' to 3' orientation.

Two PCR reactions, A and B, were set up according to table 2.1, with one reaction mix prepared for each desired mutant. PCR amplification products, A and B, were analysed by 2% agarose gel electrophoresis (section 2.4.3). Examples of PCR products A and B are presented in figure 3.46, illustrating how each product can vary in size depending on the position of the point mutation. Fragments corresponding to the correct size were excised and purified using a QIAquick gel extraction kit (section 2.4.2.3).

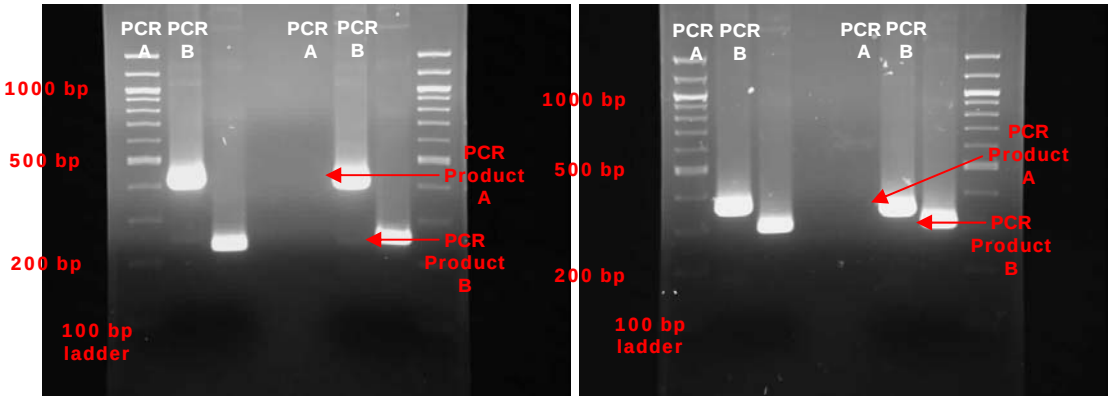


Figure 3.46 - Two 2% agarose gels illustrating the successful PCR-based site-directed mutagenesis for the production of D-SpG$_{C2}$-1 Fc and Fab mutant constructs

These two 2% agarose gels illustrate examples of PCR products A and B produced by PCR-based site-directed mutagenesis, performed on D-SpG$_{C2}$-1 pQE-30 template DNA for all D-SpG$_{C2}$-1 Fc and Fab mutant constructs.

PCR products A and B were then used to produce the final mutagenic product PCR C, according to the reaction mix shown in table 2.2. Examples of PCR gene product C are presented in figure 3.47, illustrating how each product is always the same size, 654 bp.

Fragments corresponding to the correct size were excised and purified using a QIAquick gel extraction kit (section 2.4.2.3).

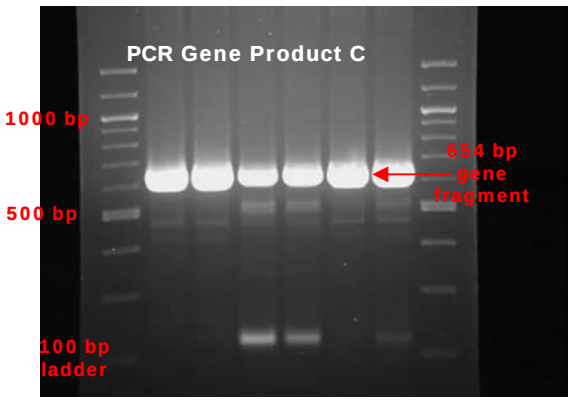


Figure 3.47 - A 2% agarose gel illustrating the successful PCR-based site-directed mutagenesis for the production of D-SpG$_{C2}$-1 Fc and Fab mutant constructs

This 2% agarose gel illustrates examples of PCR gene product C produced by PCR-based site-directed mutagenesis, performed on PCR products A and B for all D-SpG$_{C2}$-1 Fc and Fab mutant constructs.

Bam HI and *Hin*d III restriction enzyme digests were then performed on the mutant PCR gene products obtained, as described in section 2.4.5, and the 6xHis tagged *pQE*-30 vector, as previously shown in figure 3.8. The appropriate fragments were excised and extracted from agarose gel (section 2.4.2.3) and used to clone into *Bam* HI and *Hin*d III digested 6xHis tagged *pQE*-30 expression vector. Figure 3.48 presents examples of the successful restriction digests performed on PCR product C to yield the 208 bp D-SpG$_{C2}$-1 gene fragment.

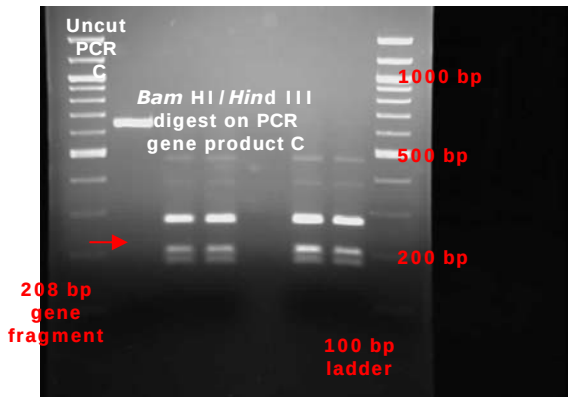


Figure 3.48 - A 2% agarose gel illustrating the successful restriction digest performed on PCR gene product C

This 2% agarose gel illustrates examples of Bam HI and Hind III restriction digests performed on PCR gene product C. The yielded 208 bp gene fragments were excised and extracted, and subsequently cloned into pQE-30 vector.

The cloning was then carried out as described in section 3.3.2.3 above. Ligation products were transformed into *E.coli* JM103 cells (section 2.4.7) and used to produce mutated plasmid DNA. All D-SpG$_{C2}$-1 mutant constructs were then checked to ensure that the sequences contained the mutated amino acids and had been correctly inserted into the

pQE-30 vector at the designated restriction sites. This was accomplished by dideoxynucleotide sequencing, performed by MWG-Biotech.

3.8 - Expression and purification of D-SpG$_{C2}$-1

The induced protein was extracted from the cytoplasm by sonication and the preparation of crude cell extract is critical in determining the total amount of desired protein available for purification. Heat treatment was then employed as an initial step in the purification of D-SpG$_{C2}$-1 from the soluble fraction, followed by anion-exchange chromatography. The very high denaturation temperatures determined for IgG-binding domains C1 and C3 [Alexander *et al*., 1992] is easily exploited for this procedure and has already been used successfully by several groups [Lian *et al*., 1992; Achari *et al*., 1992] in the purification of these domains. The denaturation temperature stated for the C1 domain, which shares the most homology to the C2 domain upon which the D-SPG$_{C2}$-1 constructs were based, is 87.5 °C. Heat treatment at 80 °C was found to be most effective temperature for purification, without the significant loss of the desired protein.

The IgG-binding domains of SpG possess an isoelectric point (p*I*) of 4.2 [Achari *et al*., 1992], so above this pH, the proteins will have a net negative charge. This means that anion-exchange chromatography was thought to be the most appropriate method to purify D-SpG$_{C2}$-1 at pH 8.0. Anion-exchange has previously been used by Lian *et al* [1992] to purify similar SpG IgG-binding domains.

3.8.1 - Expression and purification of D-SpG$_{C2}$-1 *pQE*-30 constructs

All D-SpG$_{C2}$-1 *pQE*-30 proteins were expressed and purified according to conditions stated in sections 2.5.1.1 to 2.5.1.4. Figure 3.49 presents a typical elution profile obtained from a Q-Sepharose column using a linear salt gradient (see sections 2.5.1.1. to 2.5.1.4), illustrating the absorbance peaks at 280 nm and conductivity of each fraction. Purified domain is typically eluted when the conductivity reaches between 4.5 and 22 mS at pH 8.0, which corresponds to fractions 18 to 50 taken from the larger peak and the subsequent smaller peak.

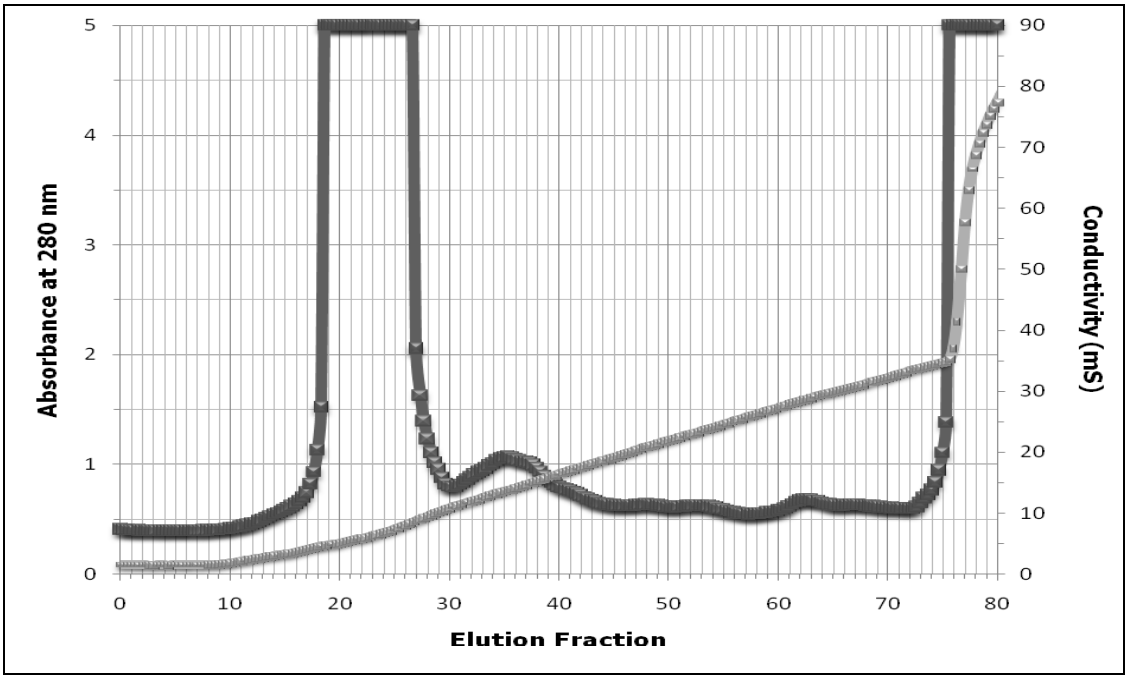


Figure 3.49 - A typical elution profile of D-SpG$_{C2}$-1 obtained from a linear gradient, illustrating the absorbance peaks at 280nm and conductivity for each fraction number

D-SpG$_{C2}$-1 pQE-30 proteins were expressed and purified on a Q-Sepharose matrix column. A typical elution profile obtained illustrates the absorbance of the collected fraction samples at 280 nm, and the corresponding conductivity. The absorbance at 280 nm is denoted by the black line, and the conductivity is denoted by the grey line. Fractions were analysed by 17.5% SDS polyacrylamide gels.

The detection of expression of all D-SpG$_{C2}$-1 *pQE*-30 constructs was performed by running fraction samples on a high-resolution separating gel containing 17.5% acrylamide under reducing conditions, as described in section 2.5.3. Figure 3.50 presents two typical 17.5% SDS polyacrylamide gels of samples taken from the collected eluate in the specified region. Fractions were pooled according to purity, and concentrated by an ammonium sulphate precipitation at the 80% saturation level (section 2.5.4).

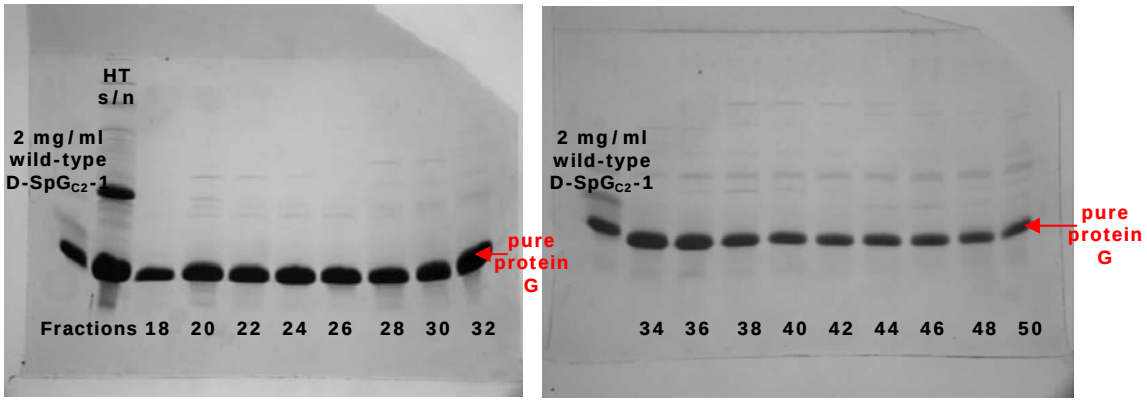


Figure 3.50 - Two typical 17.5% SDS polyacrylamide gels of fraction samples collected from the purification of D-SpG$_{C2}$-1 protein

Two high-resolution reducing 17.5% SDS gels were used to analyse the fraction samples collected from the linear gradient employed to purify all D-SpG$_{C2}$-1 protein constructs. A 2 mg/ml sample of purified wild-type SpG was used as a control, as well as a sample taken from the heat-treated supernatant (HT s/n). Samples shown above from fractions 18 to 50 inclusive were judged and pooled and stored at 4 °C as an 80% $(NH_4)_2$ SO_4 precipitate.

These SDS gels exhibit a very high level of protein expression that was representative for both wild-type and mutant 6xHis-tagged D-SpG$_{C2}$-1 constructs, and demonstrates that pure protein has been obtained. This protein can now provide a model for subsequent characterisation of SpG's interactions with Fc and Fab and for further studies into determining any differences in both binding and stability.

The expression and purification procedures employed in this study should be applicable for the production of the wild-type D-SpG$_{C2}$-1, as well as mutant constructs, as they are not IgG-binding dependent. This means that any protein that cannot be produced using these methods must be significantly structurally altered, and therefore will not provide a valid model for characterisation studies.

All of the mutants were made, expressed and purified like wild-type D-SpG$_{C2}$-1, except for the mutant K33A, occupying a position in the α-helix. This mutation involves the only lysine to be implicated in the interfacial interaction of SpG with Fc, and is situated on the C-terminal end of the α-helix next to Glu26. Lys33 forms a single hydrogen bond with the side-chain of the Fc residue Glu380, which is located on β-strand 3 of the C_H3 domain. This mutant could not be expressed, suggesting that the insertion of an alanine residue into the helix at this position disrupted the structure in some way. An incompletely folded protein is considered more susceptible to degradation than a folded one, so it is likely that the protein was degraded on formation by host *E.coli* proteases.

3.9 - Expression and purification of Fab

Fab consists of V_H–C_H1 (Fab heavy chain) and V_L–C_L (Fab light chain), in which the two

polypeptide chains are linked covalently by a C-terminal disulphide bond. To allow the interaction of Fab with D-SpG$_{C2}$-1 to be studied extensively, it was essential that enough Fab protein was available. This was achieved by employing a modified pTTOD vector [Humphreys *et al*., 2002]. The pTTOD vector, kindly supplied by UCB-Celltech, Slough, had been optimised for the expression of Fab in *E.coli,* and is shown below in figure 3.51.

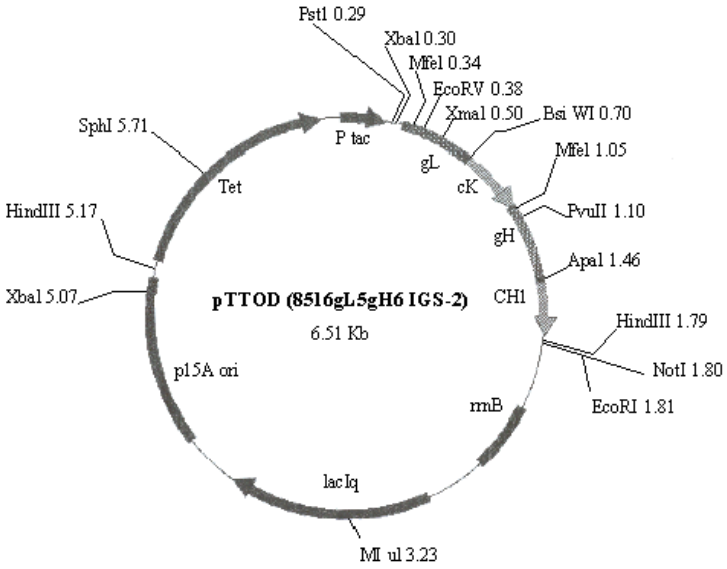


Figure 3.51 – An adapted schematic diagram of the pTTOD vector used for cloning and expression of Fab

The Fab gene is cloned in at the EcoR I and Pst I restriction sites, and is encoded in two sections, with the heavy and light chains produced separately.

The pTTOD vector was used to clone the wild-type Fab heavy and light chains genes, of known sequence, which are encoded separately in different reading frames. The amino acid sequences of the humanised Fab heavy and light chains are shown in figure 3.52.

Heavy Chain:

EVQLVESGGGLVQPGGSLRLSCAFSGFSLSTSGVGVGWVRQAPGKGLEWVAHIWDGDESYNPSLKTQFTISKDTSKNTVYLQ
MNSLRAEDTAVYYCARNRYDPPWFVDWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGA
LTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDTHTCAA

Light Chain:

DIQMTQSPSSLSASVGDRVTITCRASQDISNYLSWYQQKPGKAPKLLIYYTSKLHSGVPSRFSGSGSGTDYTLTISSLQPEDFA
TYYCQQGKMLPWTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQ
DSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

Figure 3.52 - Amino acid sequence of the heavy and light chains

The pTTOD vector encodes the heavy and light chains of humanised mouse Fab 8516 separately. Each chain has an N-terminal ompA signal sequence, which targets the translated protein to the periplasm where the chains can fold and associate [UCB-Celltech].

The expression and purification of Fab protein, was achieved in a similar way as with D-SpG$_{C2}$-1, by cell growth, induction and harvest methods. However, a lower yield of Fab was achieved using this technique, so a slightly different extraction and purification

system was devised to produce larger quantities of the recombinant Fab fragments using fermentation. The use of *E. coli* for the expression of Fab can be difficult, as each Fab may be tolerated by the bacteria to a different degree [Lin *et al*., 2006]. Therefore, the growth conditions that relate to protein expression have to be optimised for each particular construct used [Humphreys, *et al.*, 2002].

An *E.coli* fermentation system uses simple conditions and involves two methods for the expression of recombinant antibody fragments.

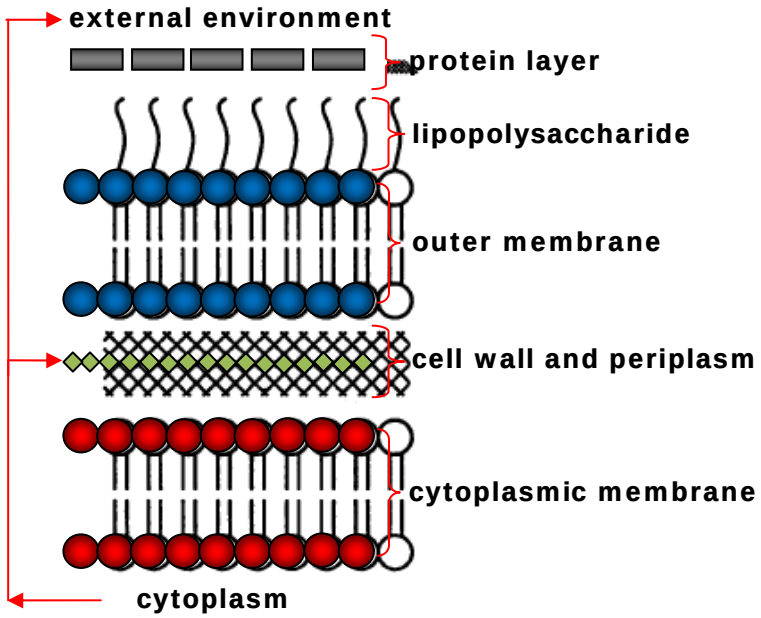


Figure 3.53 – An adapted schematic representation of the pathways for protein expression in gram-negative bacteria

The general secretory pathway in gram-negative bacteria is presented. Secreted recombinant Fab protein follows the pathway indicated by the red arrows [Pugsley, 1993].

The first method simply expresses the protein in the cytoplasm. Alternatively, if both chains of Fab possess an N-terminal ompA (outer membrane protein A) signal peptide sequence, then the protein is maintained in a loosely folded conformation bound to ompA for translocation across the cytoplasmic membrane into the periplasm, from where the ompA is inserted into the outer membrane. The second method targets the fragment to the periplasmic space between the cytoplasmic and outer membrane. Cleavage of the ompA signal sequence releases the protein into the periplasm, where the protein folds and adopts it tertiary and quaternary structure. In the periplasm there are a number of proteins that contribute to the folding, such as catalysing the formation of disulphide bonds. In the latter method, osmotic shock is used to release the folded protein from the periplasmic space into the culture medium. This expression system should therefore, produce a fully folded and functional antibody fragment [Pugsley, 1993]. The periplasm is the ideal subcellular location for Fab expression since its oxidising environment

enables the formation of intra- and inter-chain disulphide bonds [Debarbieux and Beckwith, 1998].

Bacterial expression of wild-type Fab was carried out successfully for this study, with the 48 hour fermentation of JM103 *E.coli* cells containing the pTTOD vector. The Fab protein initially expresses as separate light and heavy chains, which associate and fold in the periplasm. However, it is the heavy chains that are of particular interest in this study, as SpG has been shown to bind to Fab by forming an anti-parallel interaction with the last β-strand of the C_H1 domain. Average wet cell yields obtained from the fermentation expression system, outlined above, ranged from 70 to 280 g.

The Fab protein was then extracted from the periplasm of these cells by osmotic shock using a sucrose solution with EDTA, based on that used by Neu and Heppel [1965]. The preparation of crude cell extract is critical in determining the total amount of desired protein available for purification, and it is preferable to selectively release the periplasmic contents of the cell without disrupting the cytoplasmic membrane. This minimises the amount of contaminating proteins in the extract. Fab fragments can be purified with a simple aqueous periplasmic heat extraction that removes the bulk of the host cytoplasmic and membrane proteins. Heat treatment at 60 °C was found to be most effective temperature for purification, without the significant loss of the desired protein.

3.9.1 - Expression and purification of Fab heavy chain C_H1

Fab proteins were expressed and purified according to conditions stated in sections 2.5.2.1 to 2.5.2.4. The theoretical p*I* of Fab heavy chain is 9.1 and the light chain is 6.7 [Fujii *et al*., 2007]. Therefore, attempts were made to use cation-exchange chromatography to purify the Fab fragment from the refined supernatant obtained from the heat treatment at pH 2.2. It was also initially thought that the high p*I* of the Fab fragment would be unlikely to be similar to that of the contaminant proteins. However, the purification of Fab fragments from the periplasm of *E.coli* using cation-exchange chromatography also resulted in the co-purification of some *E.coli* host proteins, having similar functional p*I*'s. This problem was also encountered by Humphrey *et al* [Humphrey *et al*., 2004], who identified contaminating proteins, such as periplasmic phosphate binding protein, that possessed similar p*I* values. Therefore other methods had to be employed. Other workers in the laboratory had success in the use of an affinity column comprised of single B domains of protein L immobilised onto NHS-Sepharose (Cossins, 2006). This column is able to bind Fab and kappa light chain but not heavy chain. The supernatant from the heat treatment was applied to a 50 ml column of immobilised protein L domain equilibrated in 100 mM Tris buffer containing 0.5 M NaCl at pH 8.0. The

column was then washed with loading buffer until the eluate was optically clear at 280 nm and then a linear gradient of 0 to 50mM citrate buffer at pH 2.2 also containing 0.5 M NaCl. Figure 3.54 presents a typical elution profile obtained from a Protein L-Sepharose illustrating the absorbance peaks at 280 nm and conductivity for each fraction. Purified Fab is typically eluted when the conductivity reaches between 6 and 9 mS at pH 2.2, which corresponds to fractions 44 to 50 taken from the large peak.

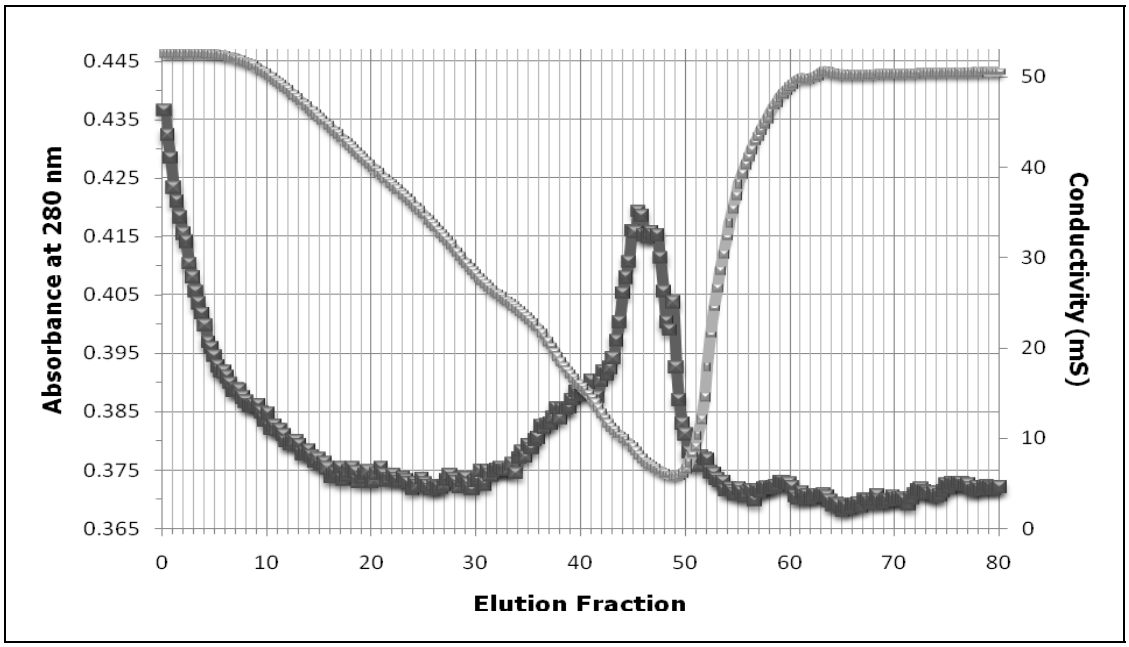


Figure 3.54 - A typical elution profile of wild-type Fab obtained from a linear gradient, illustrating the absorbance peaks at 280nm and conductivity for each fraction number

Wild-type Fab protein was expressed and purified on a NHS-Sepharose matrix column. A typical elution profile obtained illustrates the absorbance of the collected fraction samples at 280 nm, and the corresponding conductivity. The absorbance at 280 nm is denoted by the black line, and the conductivity is denoted by the grey line.

The detection of expression of the Fab fragments were performed by running fraction samples on a high-resolution non-reducing separating gel containing 17.5% acrylamide under reducing conditions, as described in section 2.5.3. Figure 3.55 presents two typical non-reducing 17.5% SDS polyacrylamide gels of samples taken from the collected elutent in the specified region. Fractions were pooled according to purity, and concentrated by an ammonium sulphate precipitation at the 80% saturation level (section 2.5.4).

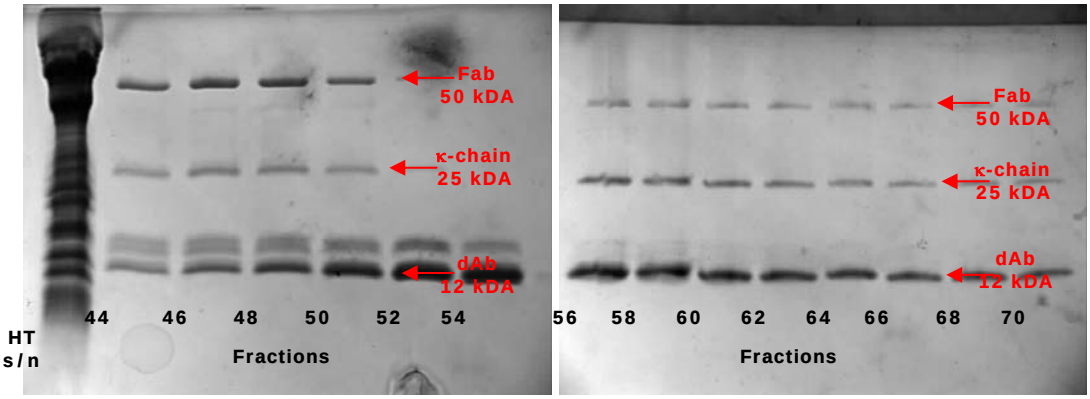


Figure 3.55 - Two typical non-reducing 17.5% SDS polyacrylamide gels of fraction samples collected from the purification of Fab protein

Two high-resolution non-reducing 17.5% SDS gels were used to analyse the fraction samples collected from the linear gradient employed to purify wild-type Fab protein. A sample taken from the heat-treated supernatant (HT s/n) is shown. All samples shown above from fractions 44 to 50 inclusive were pooled and stored at 4 °C as an 80% $(NH_4)_2 SO_4$ precipitate.

These SDS gels exhibit a standard level of protein expression that was representative for wild-type Fab protein, and demonstrates that relatively pure protein has been obtained. However, these gels also show that the use of the Fab pTTOD vector resulted in the production of several different antibody fragments as well as the desired Fab protein. The presence of kappa light chain that has not associated with heavy chain was expected, as the two chains are translated separately to associate and fold in the periplasmic space. The fractions that contained both Fab and light chain were not pure enough to be used, and were pooled and applied to a Superdex 200 Hi-Load 16/60 column gel filtration column. This additional chromatographic step is required to ensure total separation of Fab from the contaminating proteins. The other fragment that resulted from the cation-exchange chromatography purification of Fab was identified as a small κ-light chain fragment. This high expressing protein is a domain antibody (dAb) that consists of only the V_L domain of Fab and has been shown to bind to protein L [Cossins, 2006].

The wild-type Fab pTTOD protein prepared as described, together with standard Fab kindly donated by UCB-Celltech, can now be employed in the subsequent characterisation of SpG's interactions with Fab.

3.10 - Discussion

A 208 bp single Ig-binding domain gene construct of SpG, based on the wild-type SpG_{C2}-1 gene encoding the C2 domain, was successfully cloned into the *pQE*-30 vector. Specific amino acid residues that contribute to the Fc and Fab binding interactions via a network of hydrogen bonds were identified and effectively modified by a programme of two-sided SOE-ing site-directed mutagenesis. An extensive series of mutant proteins were produced and a high level of expression for the wild-type, Fab and all but one of Fc mutant constructs were achieved.

Further characterisation of these Fc and Fab mutants will allow information on the importance of these individual amino acids in the encoded protein to be obtained. Also identifying the key residues that are critical for these interactions allows the contribution of a specific amino acid to the biological function of SpG to be evaluated.

CHAPTER 4.0

CONFORMATIONAL STABILITY OF SPG CONSTRUCTS ENGINEERED TO PROBE FC AND FAB INTERACTIONS

4.0 - Introduction

All protein molecules are linear polypeptide chains of amino acids that adopt and fold into specific three-dimensional structures, known as native states. This generates a biologically active structure that is required for specific interactions, allowing proteins to play an important role in the mechanism and control of a broad spectrum of cellular processes, as well as to perform intricate biological functions. These range from the translocation of proteins across membranes to the regulation of cellular events in the cell cycle [Radford and Dobson, 1999]. Christian Anfinsen, 1973, proposed that the information needed to specify a protein's complex structure is encoded in its amino acid sequence and it has been shown that under suitable conditions, globular proteins will spontaneously fold to their native states. Since Anfinsen's famous protein refolding studies, subsequent theoretical and experimental work has provided further information and understanding into the principles involved in the mechanisms of protein folding. Complementary analytical and computational studies representing idealised models of proteins have also provided valuable insights into the prediction of protein-folding mechanisms and the three-dimensional structures of proteins [Kim and Baldwin, 1982; Kim and Baldwin, 1990; Evans and Radford, 1994].

Contributions from non-covalent interactions, such as hydrogen-bonding, van der Waal's, hydrophobic and electrostatic interactions, govern the folding of proteins to their native structures. The net contribution of these interactions in the folded state relative to the unfolded state governs the stability.

The correct three-dimensional structure is essential for the protein to function, and misfolding usually produces inactive proteins with different properties [Nilsson and Anderson, 1991]. Misfolded proteins are responsible for several diseases, such as Alzheimer's disease [Hoozemans *et al*, 2005], cystic fibrosis [Sato *et al*., 1996], alpha-1-antitrypsin deficiency [James *et al*., 1999] and prion-related illnesses including Creutzfeldt-Jakob disease (CJD) and Bovine Spongiform Encephalopathy (BSE) [Horwich and Weissman, 1997], and are all associated with the aggregation of misfolded proteins into insoluble plaques [Ramirez-Alvarado *et al*., 2003]. Therefore, by understanding the processes, which govern protein folding, medical intervention of these diseases may be achieved.

Information regarding the secondary and tertiary structure of a protein can be obtained from its spectral properties, using fluorescence and circular dichroism (CD) techniques, as well as from experiments using chaotropic reagents.

Fluorescence responds to changes in the environment of fluorophores (tryptophan and tyrosine residues in proteins) and therefore, changes in the tertiary structure. Whereas circular dichroism in the far UV range, follows changes in the symmetry of the peptide backbone and therefore, the secondary structure of proteins. The proposed study into the conformation stability due to mutagenesis of a domain of SpG will allow the contribution of a specific amino acid to the parameters associated with the native state and biological function of SpG to be evaluated. CD and the use of the chaotropic reagent, guanidine-hydrochloride (Gdn-HCl), were employed as versatile tools to probe the native structure of SpG, Fc and Fab mutant constructs, and any changes in structure or stability.

4.1 – Protein folding

Protein folding is the physical process by which an unfolded polypeptide chain becomes highly compact and folds into a specific native and functional three-dimensional, tertiary structure [Baker, 2000].

4.1.1 – Sequence, folding and structure

Every protein originates as a polypeptide, which is translated from a sequence of mRNA as a linear chain of amino acids in the peptidyl-transferase site of the ribosome from the N-terminus. The elongating polypeptide then moves across the narrow ribosome channel and emerges at the exit site from the ribosome into the cytoplasm. This polypeptide can be described as a random coil, as it does not possess any developed three-dimensional structure or long-range interactions [Dinner *et al*., 2000]. After leaving the ribosome, the polypeptide chain goes through a number of steps, known as the folding pathway, before it folds into its native protein structure.

Studies using data and propositions from both experimental [Kuwajima 1992; Ptitsyn, 1995; Baldwin, 1995, 1996 and 1999] and theoretical studies [Dill and Chan, 1997; Nyemeyer *et al*., 1998] have suggested that protein folding is divided into two stages [Arai and Kuwajima, 2000]. The first stage involves the formation of a molten globule state from the unfolded polypeptide chain, which is a flexible intermediate where the protein forms its basic structural framework, but with no specific side-chain packing. The second stage involves the formation of the native state from the molten globule state, where the specific tertiary structure is assembled with precise hydrophobic interactions and side-chain packing. This subsequent folding often requires specific molecules called molecular chaperones to help in the formation of the precise native state. However, more recent studies

have indicated that these two steps may not be as distinct from one another as previously thought, and that co-translational protein folding is a more accurate model [Basharov, 2003].

Co-translational protein folding is the folding of the polypeptide chain whilst it is being synthesised on the ribosome. Thus, while the ribosome joins residues to the elongating polypeptide, either α-helices, β-sheets or the definite conformation can be formed in any N-terminal part of the polypeptide that has emerged from the ribosome. This is thought to occur because synthesis of an entire polypeptide chain may take seconds, whereas secondary structure formation and compaction of a polypeptide chain takes less than one second [Roder and Colon, 1997]. However, although it is known that any notable conformational changes of a synthesised polypeptide are prevented by the narrow ribosome channel, 20-28 Å [Ramakrishnan, 2002], alternative studies have shown that proteins exhibit folding activity whilst still bound to the ribosome [Fedorov and Baldwin, 1997]. In addition, recent spectroscopic studies coupled with quantitative kinetic analysis, suggest that folding is facilitated by the rapid formation of compact intermediates with some native-like structural features, and as shown by Ptitsyn, formation of compact globular intermediates only takes a few seconds [Ptitsyn, 1995].

The interior of the protein's tertiary structure is tightly packed and is closer to a solid than liquid phase [Klapper, 1971]. The bonds that are formed in the interior of the protein are weak in comparison to covalent bonds. However, weak interactions have an advantage in that they are less rigid and provide the protein with more flexibility and dynamics, and the protein is compensated by their vast number. The tight packing of the core is thought to give a considerable contribution to protein stability [Pace *et al.* 2004]. In contrast, the surface of the protein consists of a majority of polar and charged side chains, which are very important for protein solubility [Malissard and Berger, 2001], and as proteins increase in size, the proportion of buried charged amino acids also increases [Kajander *et al.*, 2000].

4.2 – Thermodynamic and kinetic control of protein folding

The majority of proteins fold spontaneously into their native structure in aqueous solution, but the folding process must satisfy two conditions, one thermodynamic and one kinetic [Eaton *et al*., 2000]. The thermodynamic consideration is that a protein must adopt a single, folded conformation, whereas the kinetic requirement is that a protein must fold to the native structure on an appropriate timescale [Baker and Agard, 1994]. For most proteins, the native structure is under thermodynamic control and corresponds to a minimum Gibbs free energy, ranging from -5 to -15 $kJmol^{-1}$ [Yon, 1997].

The duration of the folding process varies depending on the size of the protein. Large proteins require minutes or hours to fold as a result of steric hindrances compared to smaller proteins, which tend to fold in milliseconds and fast protein folding reactions are completed within a few microseconds [Baldwin, 1994]. It has also been determined that α-helical proteins, which have a larger fraction of sequentially short-range contacts, fold faster than mixed α/β or β-sheet proteins, which have a larger fraction of long-range contacts [Dinner *et al*., 1996]. An unfolded protein has a vast number of accessible conformations, particular in its side-chains, and these various conformations differ only in the angle of rotation around the bonds of the backbone and amino side-chains. When a native protein folds, the side-chains in the hydrophobic core are restricted to a single conformation [Dinner *et al*., 2000].

4.3 – Conformational stability

Conformation stability generally refers to the structural arrangement of a protein and the overall conformation of a protein is independent of its stability. It can also be defined as the free energy difference between the native folded state and the unfolded state under physiological conditions [Pace *et al*., 1996]. To function as a protein, a polypeptide chain must possess a unique native state that is stable at a physiological temperature and be able to fold into its native structure within a reasonable time, milliseconds to minutes, at that temperature.

The majority of folded proteins have a hydrophobic core in which side-chain packing stabilises the folded state, together with charged side-chains on the solvent-exposed surface that interact with surrounding water molecules [Northey *et al*., 2002]. This structural uniqueness is characteristic of native proteins and is essential for them to express their biological functions. The compact native structure of a globular protein also contains bonds of different types, including

covalent bonds, ionic bonds, hydrogen bonds, sulphur bonds and van der Waals bonds. Hydrogen bonds play an important role in the stabilisation of structural elements such as α-helices and β-sheets [Myers and Pace, 1996].

However, the polarity of amino acids and hydrophobic effects between the protein and surrounding solvent are the main factors involved in driving the formation of the native structure [Calloni *et al*., 2003]. In aqueous solutions polar amino acids tend to be hydrophilic and attract polar water molecules, while non-polar residues are hydrophobic. These hydrophobic amino acids interact unfavourably with water molecules and are more inclined to associate with each other. This results in the water molecules rearranging their network to compensate for this event, leading to an increased ordering of the water [Northey *et al*., 2002]. The hydrophobic core is important for the stability of the native structure [Viguera *et al*., 2002].

4.4 – Conformational stability of SpG

The stability and folding characteristics of SpG has been extensively studied experimentally [Kuszewski *et al*., 1994; Park *et al*., 1997 and 1999; McCallister *et al*., 2000; Brown and Head-Gordon, 2004], and has proven to be an excellent model for conformation stability studies. The spacial arrangements of the secondary-structure elements in which an α-helix is positioned diagonally across a β-sheet, is very stable [Sauer-Eriksson *et al*., 1995]. Its three-dimensional structure is known [Akerstrom *et al*., 1985; Gronenborn *et al*., 1991; Gallagher *et al*., 1994], as well as its thermodynamic properties characterised [Alexander *et al*., 1992].

Previous studies have shown that SpG folds through at least two pathways. The fast pathway exhibits two-state kinetics with a transition state ensemble resulting in β-hairpin 2 [Huyghues-Despointes *et al*., 2006]. This is assisted by the α-helix and the overall folding mechanism seems to be consistent with a nucleation-condensation mechanism [Daggett and Fersht, 2003a]. The slower folding pathway involves an early intermediate [Park *et al*., 1997 and 1999] that is characterised by hydrophobic contacts involving the third β-strand interacting with β-strands 1 and 2 [Brown and Head-Gordon, 2004]. The intermediate precedes the rate-limiting barrier of folding, which separates the unfolded state from the early folding intermediate, and is lower in free energy relative to the unfolded state [Brown and Head-Gordon, 2004].

4.5 – Determination of conformational stability using circular dichroism

Circular dichroism (CD) is a form of light absorption spectroscopy that measures the difference in the absorption of right- and left-circularly polarised light when plane polarized light passes through an optical active sample solution [Woody, 1995]. This difference in absorbance is due to the fact that the extinction coefficients for the two types of polarized light are different [Fersht, 1999]. A schematic illustration of circular dichroism is presented in figure 4.1.

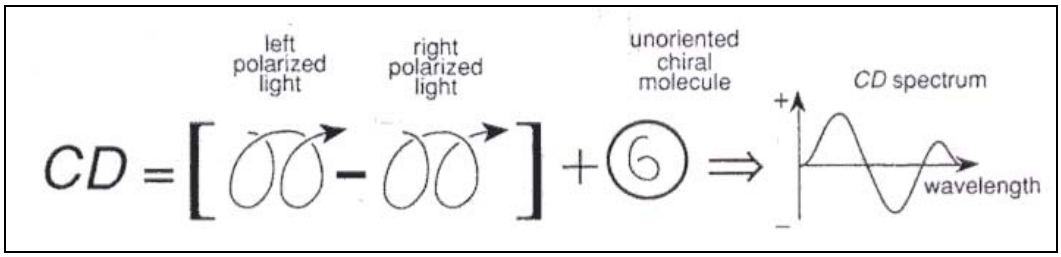


Figure 4.1 – An adapted schematic illustration of circular dichroism

In order for a molecule to produce a CD signal it must be chiral and absorb light. Plane-polarised light can be divided into its constituent equal intensity left- and right-circularly polarised light beams. When a chiral molecule absorbs plane-polarised light, left- and right-circularly polarised light will be absorbed to different extents. The differential absorbance of the two circular components by a chiral medium converts the plane-polarised light into elliptically polarised light, giving rise to the observed CD signal [Woody, 1995].

Light is a transverse electromagnetic wave consisting of an oscillating electric (E) field and a magnetic (H) field. Linear polarised light is composed of right (R) and a left (L) circularly polarized waves at right angles relative to each other and of equal amplitude and 90° difference in phase. In contrast, elliptically polarized light consists of two perpendicular waves of unequal amplitude, which differ in phase by 90° [Wen-Chuan *et al.*, 2005]. A comparison of methods for achieving polarization is shown in figure 4.2.

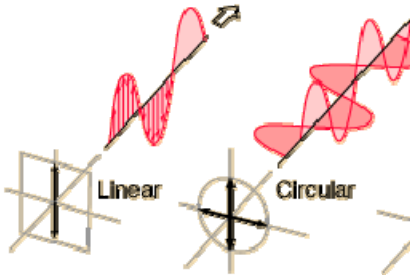


Figure 4.2 – Adapted illustration showing how right and left circularly polarized light combine under certain conditions

Diagrams showing how right and left circularly polarized light combine if (A) the two waves have the same amplitude, the result is plane-polarized light or (B) the amplitudes differ, the result is elliptically polarized light [Kelly et al., 2005].

The angle of optical rotation of an optically active sample varies with the frequency of the radiation. The rotation of the plane is due to the difference in refractive indices for R and L circularly polarised light, n_R and n_L. CD is the difference in the absorption intensity of R and L circularly polarized light (I_R and I_L) which arise due to structural asymmetry. The amplitude of the stronger absorbed component will be smaller than that of the less absorbed component. In addition, a CD spectrum is the variation of I_L - I_R with frequency of radiation. The absence of regular structure results in zero CD intensity, while an ordered structure results in a spectrum, which can contain both positive and negative signals. For CD measurements, in principle two light sources are needed, one for L and the other for R circularly polarized light, each provided with a monochromator for wavelength selection. However, L and R light may be generated from a single source [Wen-Chuan *et al*., 2004].

4.5.1 - Use of CD to study proteins

CD is a widely used technique for studying the conformation of chiral peptide and protein molecules in solution as it is uniquely sensitive to the asymmetry of the system. Chiral molecules are optically active and have the ability to rotate the plane of plane polarised light. The peptide bond is naturally chiral and therefore always optically active. CD spectra can be measured with any frequency of electromagnetic radiation at which the compound absorbs, but in practice, the majority of CD spectroscopy involves the ultraviolet-visible regions of the spectrum. For randomly oriented samples, such as solutions, a net CD signal will only be observed for chiral molecules, which are ones that cannot be superimposed on their mirror images [Sreerama and Woody, 1993].

The protein CD spectrum is divided into three regions, the far-UV (≈190-250 nm), the aromatic region of the near-UV (≈250-320 nm) and the near-UV-visible region (≈300-700 nm), and different chromophores in proteins contribute to these regions. The peptide bond is the principal element whose spectrum is detected in protein studies by CD. It exists in many conformations depending on its precise location in the protein, and the spectrum obtained is a result of an average of the various conformation parameters.

It has been shown that CD spectra are very sensitive to the secondary structures of polypeptides and proteins, and wavelengths between 260 nm and 180 nm can be analysed for the different secondary structural types, such as alpha helices, parallel and anti-parallel beta sheets, turns and random coils [Johnson, 1990].

This can then be employed to estimate the contribution of α-helices and β-sheets to the overall secondary structure. This is illustrated by the reference spectra presented in figure 4.3, which show 100% alpha-helix, beta-sheet and random coil spectra for the three different conformations of proteins.

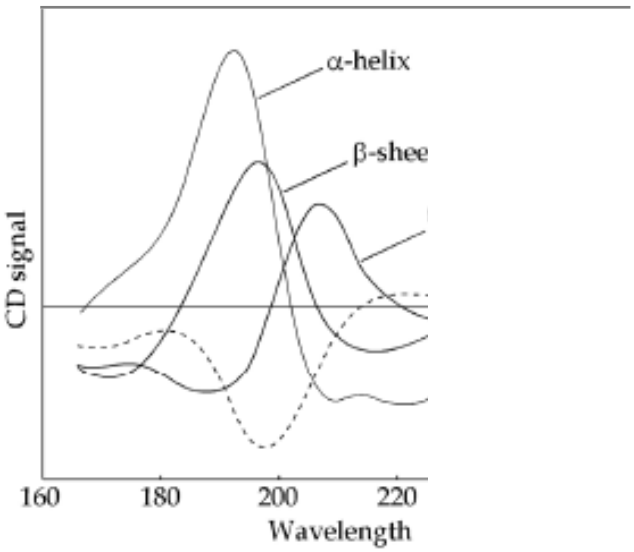


Figure 4.3 – Circular dichroism spectra of the characteristic secondary structures

Absorption in the far UV region is largely due to the peptide backbone, and is highly dependent on the secondary structure formed by the protein. The CD spectra observed for the α-helical, β-sheet and random coil are shown in black, red and green respectively. Based on spectra from http://www.jgiqc.llnl.gov

The CD signal due to α-helices is characterised by negative ellipticity showing minima at 208 nm and 222 nm and one positive peak at 192 nm. In contrast, β-sheets have a characteristic positive peak at 195 nm and a negative peak of similar magnitude around 216 nm. Random coil typically shows a negative CD signal below 200 nm, with a positive peak occurring around 218nm.

The far-UV region is reflective of the environment resulting from the contributions from side-chain chromophores, such as the aromatic amino acids, (tryptophan, tyrosine and phenylalanine), with additional contributions from cysteine residues involved in disulphide bonds. The near-UV region can provide information about the localised environment of the aromatic amino acids, thus allowing information to be obtained about the protein tertiary structure [Kahn, 1979].

4.5.2 - Determination of conformational stability of D-SpG$_{C2}$-1

All CD measurements were made in a Jasco J-720 CD spectropolarimeter at a protein concentration of 100 μM in 20 mM potassium phosphate buffer pH 8.0, at room temperature. Spectra were acquired in a 0.01 cm cell path length at a 50 nm/min scan speed. Each spectrum was measured in mdeg and expressed in

Molar Ellipticity (θ=deg.cm^2.dmol^{-1}), which eliminates inconsistencies in concentration and path length between samples, and allows a direct comparison to be made between proteins. The spectra shown are accumulations of nine spectral scans, which were averaged and baseline-corrected by subtraction of a buffer blank. All CD spectra were analysed using the manufacturer's Standard Analysis software.

4.5.3 – Far-UV CD analysis of Wt D-SpG$_{C2}$-1

Secondary structure can be determined by CD spectroscopy in the far-UV spectral region (190-250 nm). At these wavelengths the chromophore is the peptide bond, and the signal arises when it is located in a regular, folded environment. The CD signal reflects an average of the entire molecular population, and alpha-helix, beta-sheet and random coil structures each give rise to a characteristic shape and magnitude of CD spectrum.

The far-UV spectra of 100 µM Wt D-SpG$_{C2}$-1 construct and 20 mM potassium phosphate buffer, pH 8.0 were measured between 190 and 250 nm, by CD. The spectrum obtained for the 20 mM potassium phosphate buffer was then subtracted from that obtained for the wild-type protein to correct for the baseline, and is presented in figure 4.4.

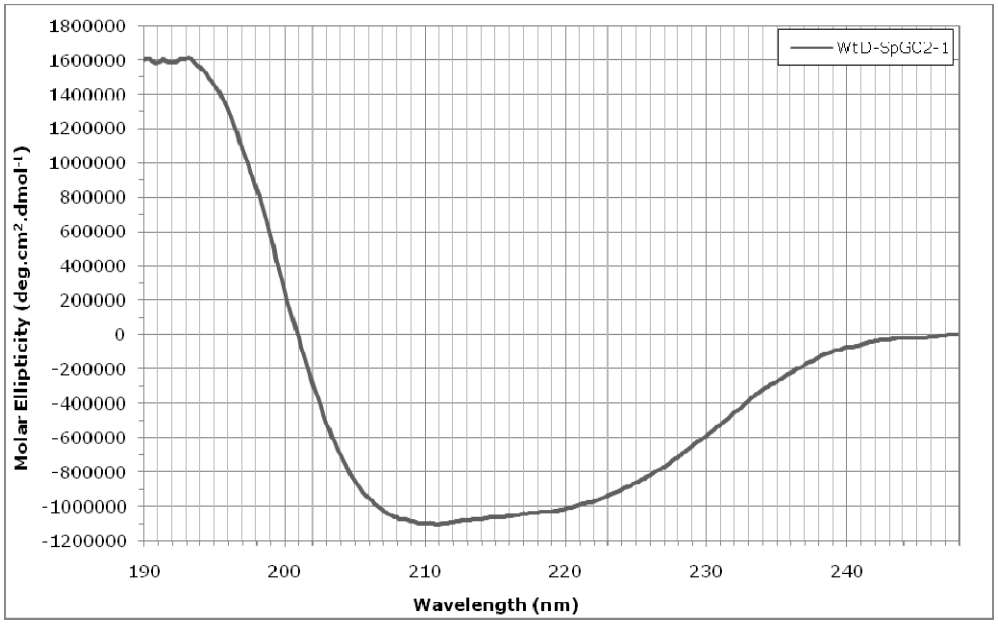


Figure 4.4 – Far-UV CD spectra of corrected Wt D-SpG$_{C2}$-1

Far-UV CD spectrum of purified 100 µM Wt D-SpG$_{C2}$-1 in 20 mM potassium phosphate buffer, pH 8.0 is presented. Measurements were taken at room temperature in a 0.01 cm path length cell. Nine scans were collected from 190 to 250 nm, averaged and baseline corrected by subtraction of blank buffer. Each spectrum was analysed as molar ellipticity (deg.cm^2.dmol^{-1}).

The far-UV spectrum obtained for wild-type D-SpG$_{C2}$-1 shows a wavelength characteristic of α-helical content, with a minima in ellipticity at 209 and 210 nm.

However, the secondary structure of SpG has a high β-strand content compared to the occurrence of just one α-helix. The expected secondary structural content for D-SpG$_{C2}$-1, calculated from the number of residues in each structural element determined by X-ray crystallography and NMR, was 46% β-strand, 27% α-helix and 27% random coil. However, the strong spectral contribution of the α-helix can obscure the true β-strand content of a protein and can give rise to errors in the estimation of the secondary structural components.

4.5.4 – Far-UV CD analysis of Fc D-SpG$_{C2}$-1 mutant constructs

Purified D-SpG$_{C2}$-1 mutant constructs were analysed by far-UV CD to determine if the mutation had affected the secondary structure of the protein. This was important as any changes in binding affinity might be attributed to the specific side-chain alterations. Far-UV spectra of engineered D-SpG$_{C2}$-1 mutant constructs are shown in figures 4.5 and 4.6, and compared to the CD spectrum obtained for the Wt D-SpG$_{C2}$-1 protein.

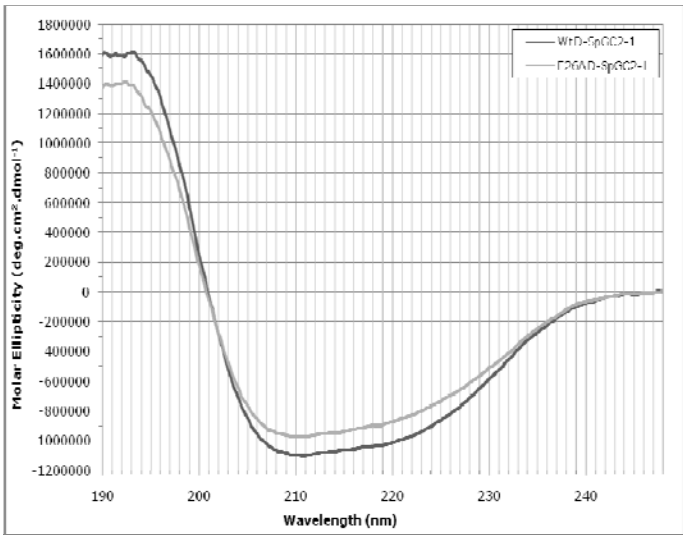
Figures 4.5 and 4.6 show that the mutations E26A, Q31A, Q31E, Q31H, N34A, N34D, E41A and E41Q have little effect on the CD spectra in comparison to that of the D-SpGC2-1 wild-type protein.

Figure 4.6b also presents a comparison between wild-type D-SpG$_{C2}$-1 with the sole cysteine mutation C56A D-SpG$_{C2}$-1. Again, the far-UV CD spectrum for this mutation is very similar to that of wild-type protein, again suggesting that the side-chain replacement of the cysteine residue by an alanine residue has not altered the structure of D-SpG$_{C2}$-1.

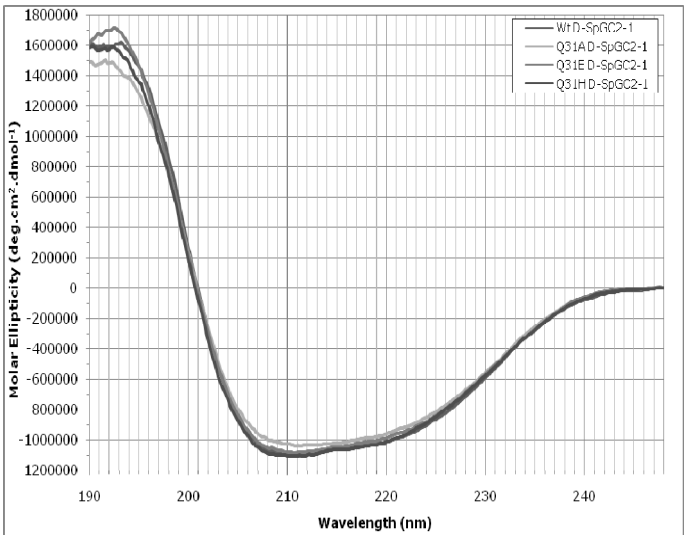
In figure 4.6c the two mutants have very similar spectra, both different to that of the wild-type protein. This is due to the loss of the tryptophan indole side-chain that contributes negative ellipticity in this region.

Although the far-UV CD spectra of proteins are generally dominated by peptide contributions, aromatic side-chains may also make significant contributions to the secondary structure analysis [Woody, 1994]. This is especially true for proteins with low α-helix content, such as immunoglobulins and their respective binding proteins, such as SpG. CD bands in the 225-235 nm region are sometimes seen, which can be attributed to tyrosine or tryptophan side-chains.

(a)



(b)



(c)

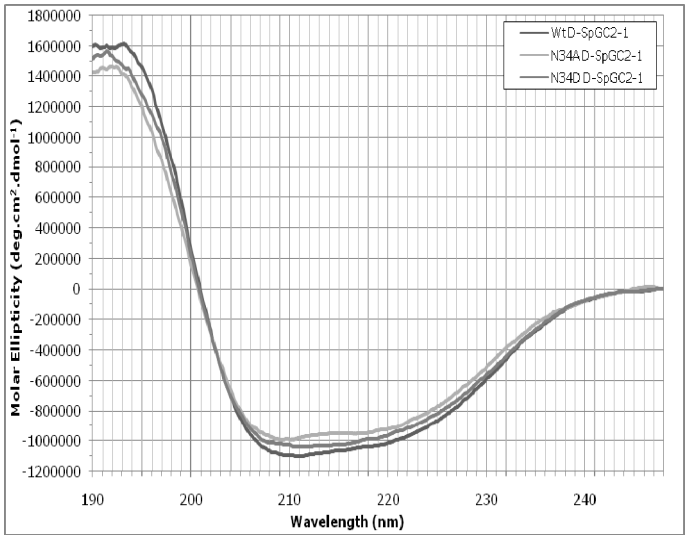
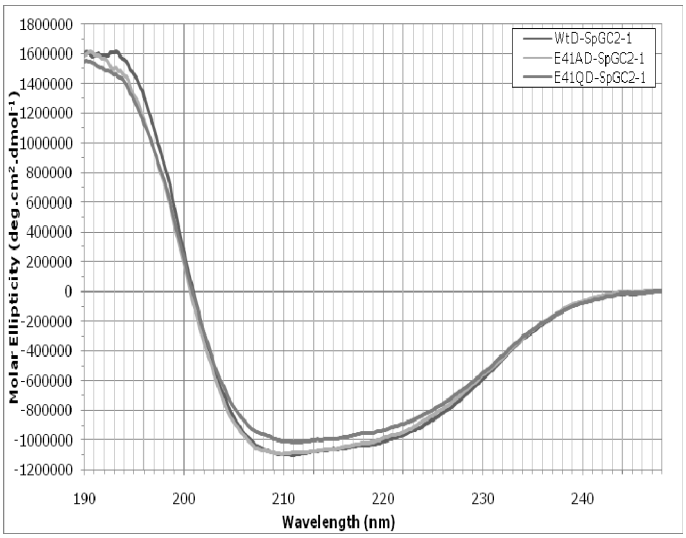


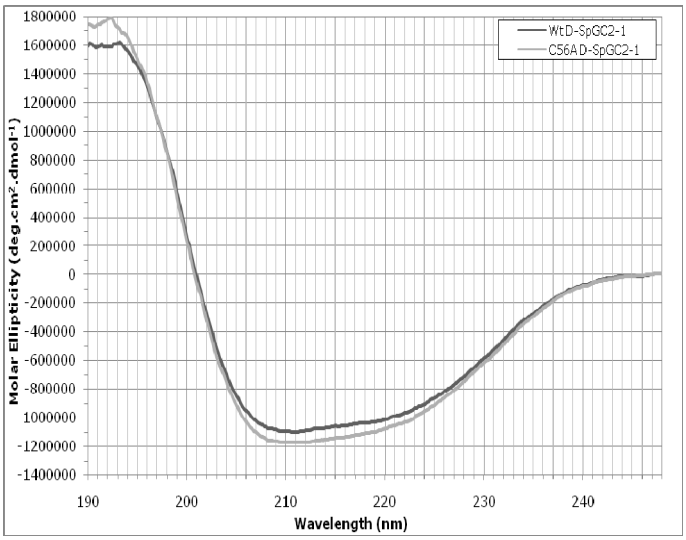
Figure 4.5 – <u>The far-UV CD spectra measured for Fc D-SpG$_{C2}$-1 mutants</u>

Far-UV CD spectra of purified 100 µM protein samples, as indicated, in 20 mM potassium phosphate buffer pH 8.0 are compared to that of wild-type. Measurements were taken at room temperature in a 0.01 cm path length cell. Nine scans were collected from 190 to 250 nm, averaged and baseline corrected by subtraction of blank buffer. Each spectrum was analysed as molar ellipticity (deg.cm^2.dmol^{-1}).

(a)



(b)



(c)

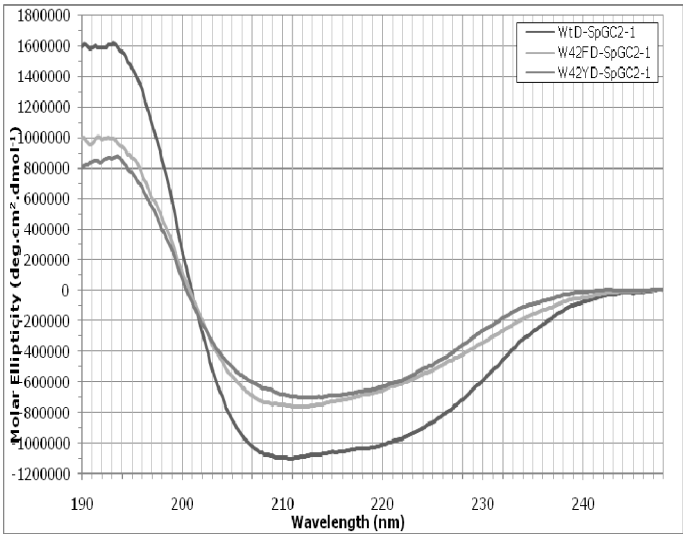


Figure 4.6 – <u>The far-UV CD spectra measured for Fc D-SpG$_{C2}$-1 mutants</u>

Far-UV CD spectra of purified 100 µM protein samples, as indicated, in 20 mM potassium phosphate buffer pH 8.0 are compared to that of wild-type. Measurements were taken at room temperature in a 0.01 cm path length cell. Nine scans were collected from 190 to 250 nm, averaged and baseline corrected by subtraction of blank buffer. Each spectrum was analysed as molar ellipticity (deg.cm^2.dmol^{-1}).

4.5.5 – Near-UV CD analysis of Wt D-SpG$_{C2}$-1

Information about a protein's tertiary structure can be obtained from the CD spectrum of a protein in the near-UV spectral region (250-350 nm). At these wavelengths the chromophores are the aromatic amino acids and disulfide bonds, and the CD signals they produce are sensitive to the overall tertiary structure of the protein.

If a protein retains secondary structure, but no side-chain packing as in an incorrectly folded or molten-globule structure, the signals in the near-UV region will be nearly zero. Conversely, the presence of significant near- UV signals is a good indication that the protein is folded into a well-defined structure. Furthermore, the near-UV CD spectrum can be sensitive to small changes in tertiary structure due to protein-protein interactions and/or changes in solvent conditions. The signal strength in the near-UV CD region is much weaker than that in the far-UV CD region due to the lower number of contributing residues.

The near-UV CD spectrum of 100 µM Wt D-SpG$_{C2}$-1 construct and 20 mM potassium phosphate buffer, pH 8.0 was measured in the 250-350 nm regions and that for 20 mM potassium phosphate buffer was subtracted from that obtained for the wild-type protein to correct for the baseline, and is presented in figure 4.7.

Signals in the region from 250-270 nm are attributable to phenylalanine residues, signals from 270-290 nm are attributable to tyrosine, and those from 280-300 nm are attributable to tryptophan. Disulfide bonds give rise to broad, weak signals throughout the near-UV spectrum [Woody, 1978]. The near-UV spectrum obtained for wild-type D-SpG$_{C2}$-1 shows wavelength contributions characteristic of both tyrosine and tryptophan content, with ellipticity signals corresponding to wavelength regions of 270-300 nm. This presence of significant near-UV signals is a good indication that the domain from wild-type D-SpG$_{C2}$-1 is folded into a well-defined structure.

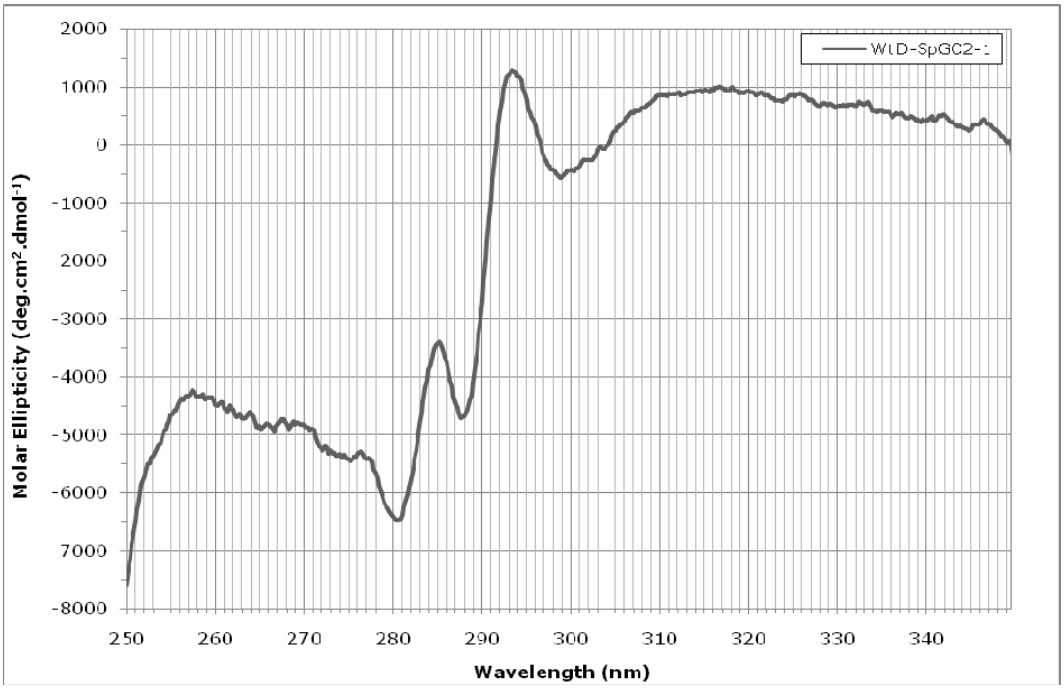


Figure 4.7 – Near-UV CD spectra of corrected Wt D-SpG$_{C2}$-1

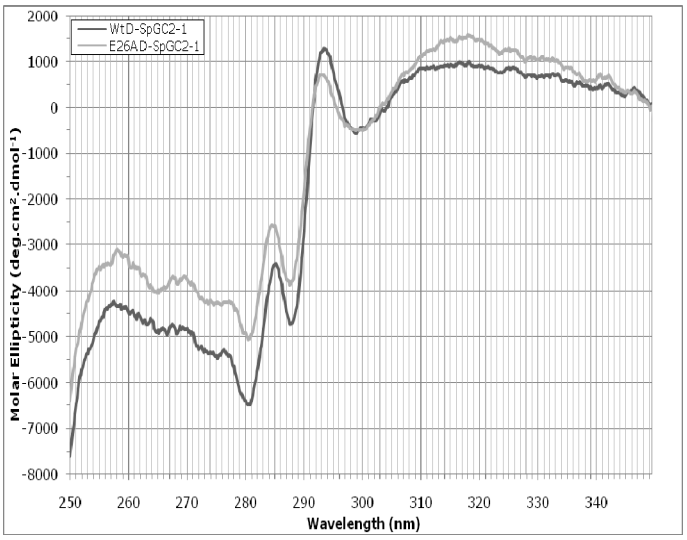
Near-UV CD spectrum of purified 100 µM Wt D-SpG$_{C2}$-1 in 20 mM potassium phosphate buffer, pH 8.0 is presented. Measurements were taken at room temperature in a 0.5 cm path length cell. Nine scans were collected from 250 to 350 nm, averaged and baseline corrected by subtraction of the spectrum of buffer. Each spectrum was analysed as molar ellipticity (deg.cm^2.dmol^{-1}).

4.5.6 – Near-UV CD analysis of D-SpG$_{C2}$-1 mutant constructs

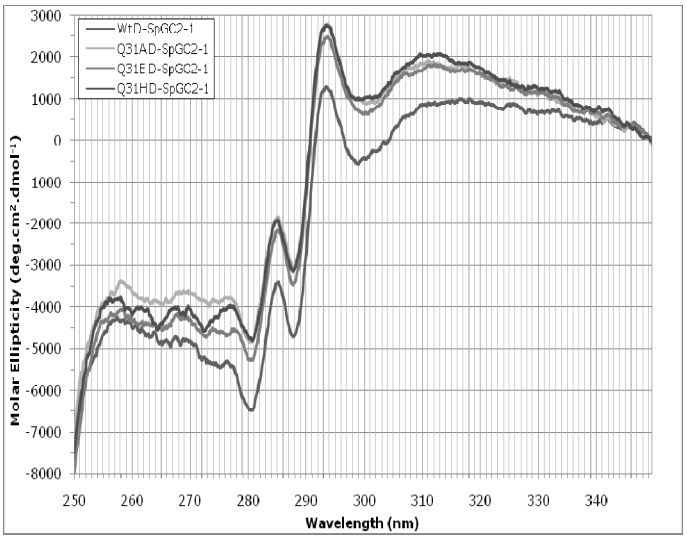
The characteristics of near-UV spectropolarimetry can be employed to detect any changes in the environment of one or more aromatic residues due to amino acid substitutions bought about by the site-directed mutagenesis [Shortle, 1989].

Near-UV spectra of engineered D-SpG$_{C2}$-1 mutant constructs are shown in figures 4.8 and 4.9, and compared to the CD spectrum obtained for the Wt D-SpG$_{C2}$-1 protein. Examination of these data reveals that the majority of the mutant constructs display spectra comparable to that of wild-type D-SpG$_{C2}$-1, with only slightly varying differences in molar ellipticity. Signals in the region from 250-270 nm are attributable to phenylalanine residues, signals from 270-290 nm are attributable to tyrosine, and those from 280-300 nm are attributable to tryptophan. Disulphide bonds give rise to broad weak signals throughout the near-UV spectrum [Woody, 1978]. The overall shape of each spectrum is very similar to that of the wild-type domain. The differences in molar ellipticity may be due to small errors in the estimation of the concentrations of proteins and/or due to fluctuation in the lamp intensity. In general, these spectra confirm that the aromatic side-chains are packed in a folded structure.

(a)



(b)



(c)

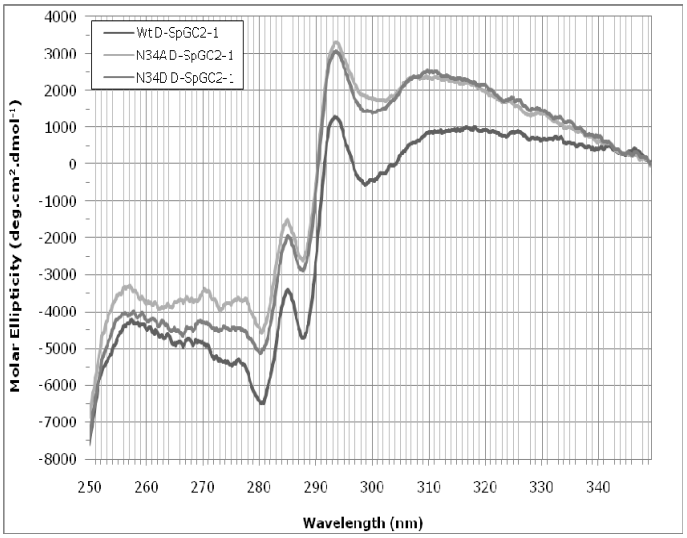


Figure 4.8 – The near-UV CD spectra measured for Fc D-SpG$_{C2}$-1 mutants

Near-UV CD spectra of purified 100 µM protein samples, as indicated, in 20 mM potassium phosphate buffer pH 8.0 are compared to that of wild-type. Measurements were taken at room temperature in a 0.5 cm path length cell. Nine scans were collected from 250 to 350 nm, averaged and baseline corrected by subtraction of the spectrum of buffer. Each spectrum was analysed as molar ellipticity (deg.cm^2.dmol^{-1}).

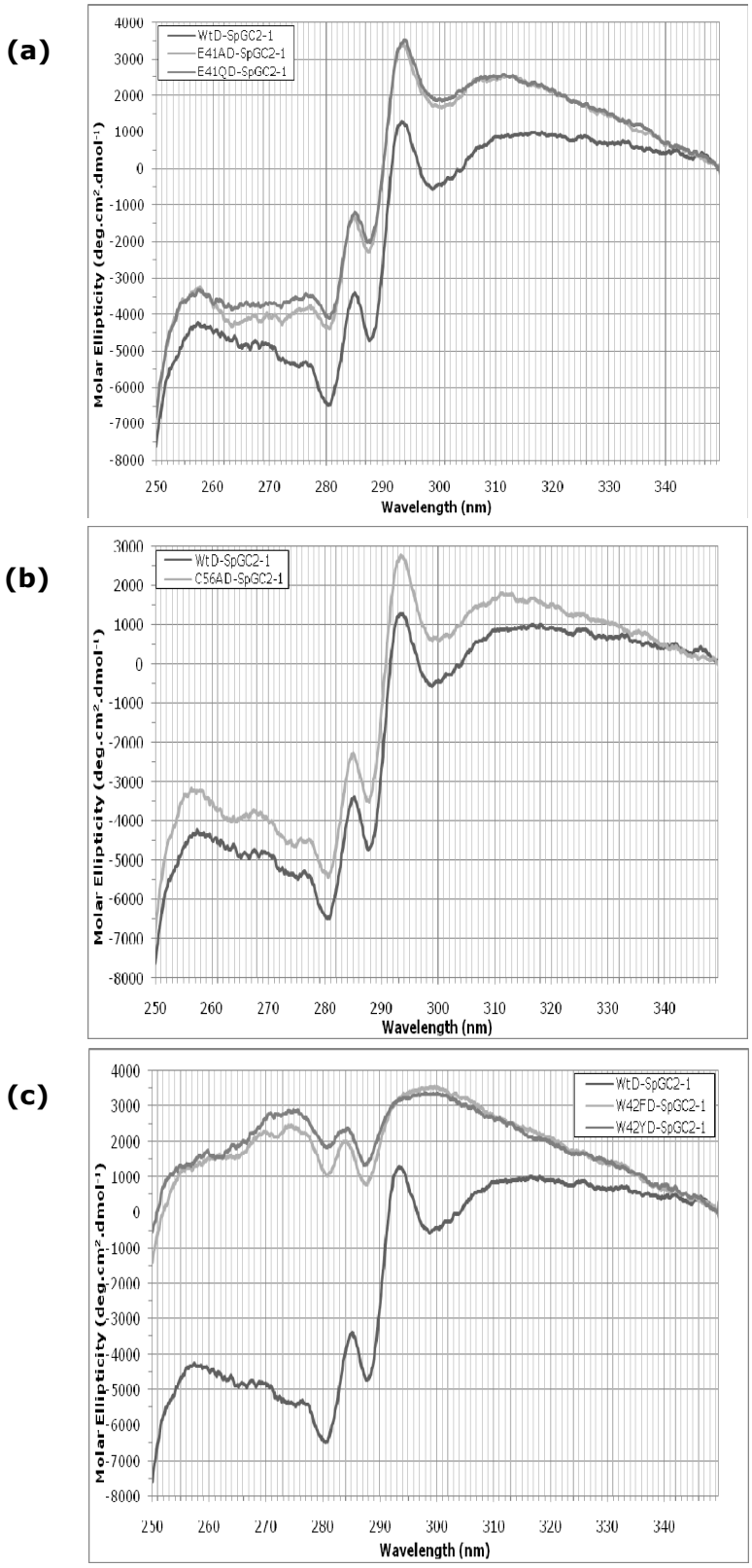


Figure 4.9 – The near-UV CD spectra measured for Fc D-SpG$_{C2}$-1 mutants

Near-UV CD spectra of purified 100 µM protein samples, as indicated, in 20 mM potassium phosphate buffer pH 8.0 are compared with wild-type. Measurements were taken at room temperature in a 0.5 cm path length cell. Nine scans were collected from 250 to 350 nm, averaged and baseline corrected by subtraction of the spectrum of buffer. Each spectrum was analysed as molar ellipticity (deg.cm^2.dmol^{-1}).

4.5.7 – Far-UV CD analysis of D-SpG$_{C2}$-1 mutant constructs designed to probe the interaction with Fab

Purified D-SpG$_{C2}$-1 mutant constructs were analysed by far-UV CD to determine if the mutation had grossly affected the secondary structure of the protein. This is important as then any changes in binding affinity could be attributed to misfolding of the proteins. Figure 4.10 presents a comparison between wild-type and E14W-W42F D-SpG$_{C2}$-1. This far-UV CD spectrum shows that the insertion of a tryptophan reporter group at the alternative position of E14W to allow the interaction between SpG and Fab to be probed, compares very well to that of wild-type D-SpG$_{C2}$-1. Both spectra are characteristic of a protein with some α-helical content, with minima in ellipticity at 210 and 220 nm. This suggests that the E14W-W42F mutation has not affected the secondary structure of the C2 domain of SpG. It is noticeable that the addition of W14 to the W42F construct restores the ellipticity between 205-230 nm to that of the wild-type protein, confirming that tryptophan contributes to this part of the spectrum.

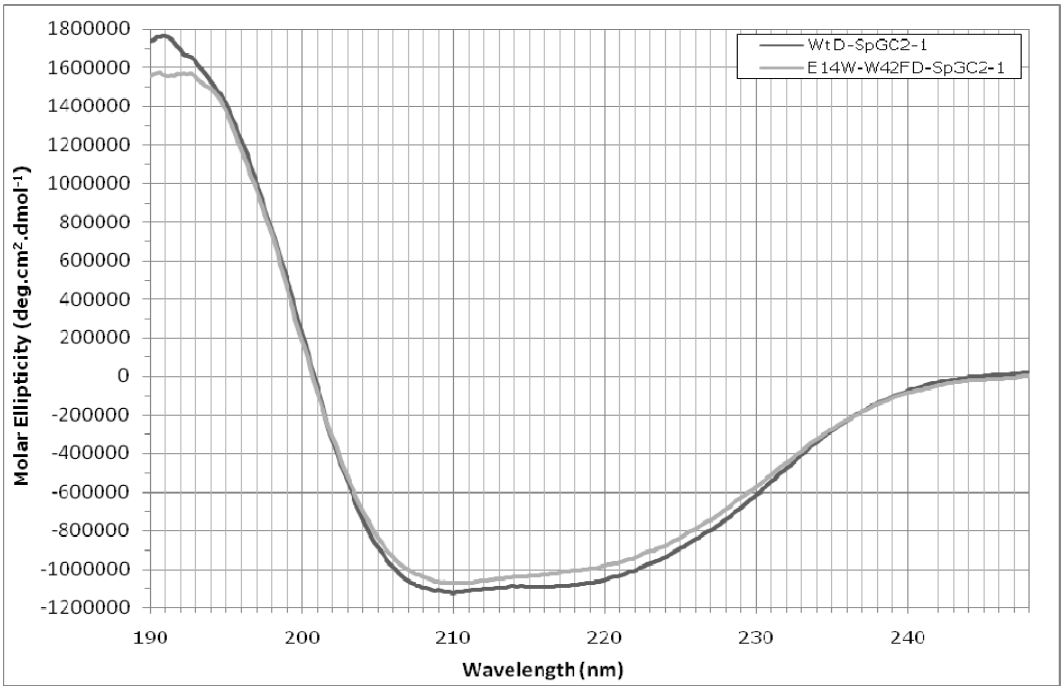
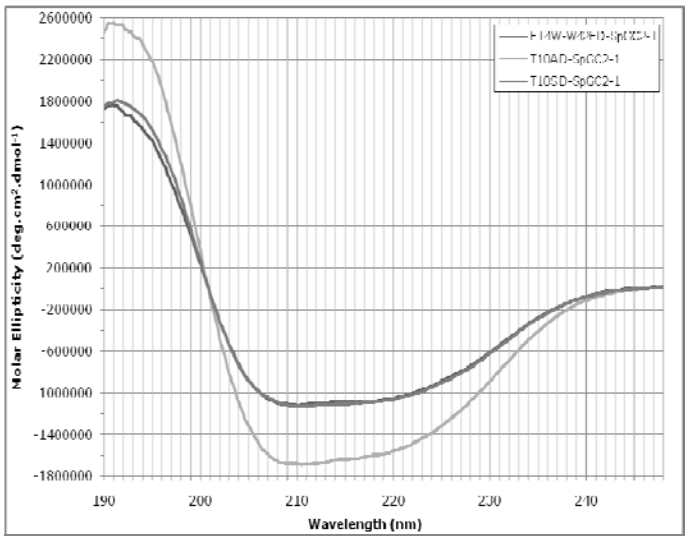


Figure 4.10 – The far-UV CD spectra measured for Wt and E14W-W42F D-SpG$_{C2}$-1

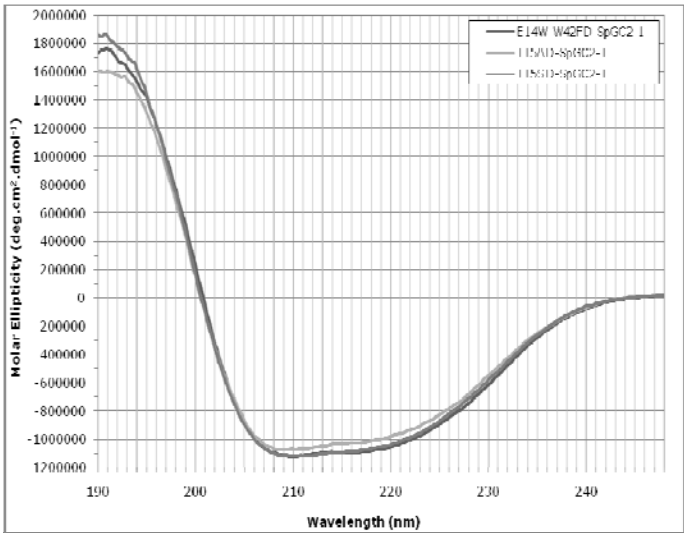
Far-UV CD spectra of purified 100 µM Wt and E14W-W42F D-SpG$_{C2}$-1 in 20 mM potassium phosphate buffer pH 8.0 is compared. Measurements were taken at room temperature in a 0.01 cm path length cell. Nine scans were collected from 190 to 250 nm, averaged and baseline corrected by subtraction of the spectrum of buffer. The spectrum was analysed as molar ellipticity (deg.cm^2.dmol^{-1}).

Therefore, further far-UV spectra of D-SpG$_{C2}$-1 mutant constructs designed to alter Fab binding are shown in figures 4.11 and 4.12, and compared to the CD spectrum obtained for the E14W-W42F D-SpG$_{C2}$-1 protein. Examination of these spectra reveals that the majority of these mutant constructs display spectra comparable to that of E14W-W42F D-SpG$_{C2}$-1, with only slight differences in molar ellipticity. All spectra are characteristic of proteins with α-helical content, suggesting that these mutations have not grossly affected the overall secondary structure of the C2 domain of SpG. The exception to this is presented in figure 4.11, which presents a comparison between E14W-W42F D-SpG$_{C2}$-1 with domains containing threonine replacements T10A and T10S. The far-UV CD spectrum for the T10S mutant overlays the spectrum of E14W-W42F perfectly. In contrast, the far-UV CD spectrum for the T10A mutant shows a significant reduction in the CD signal in the far-UV CD spectra, which may be due to folding differences.

(a)



(b)



(c)

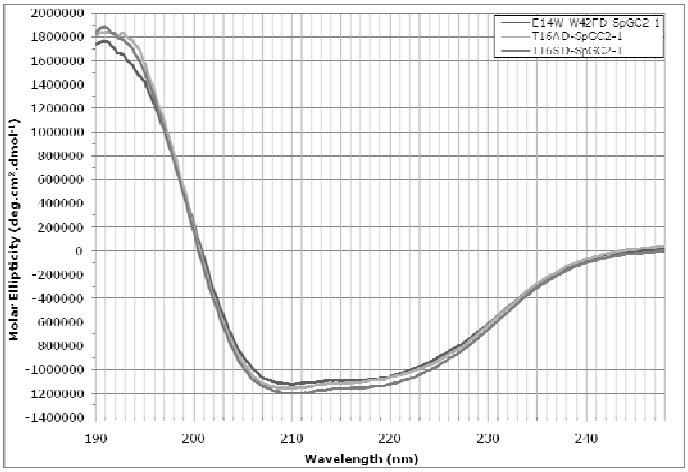
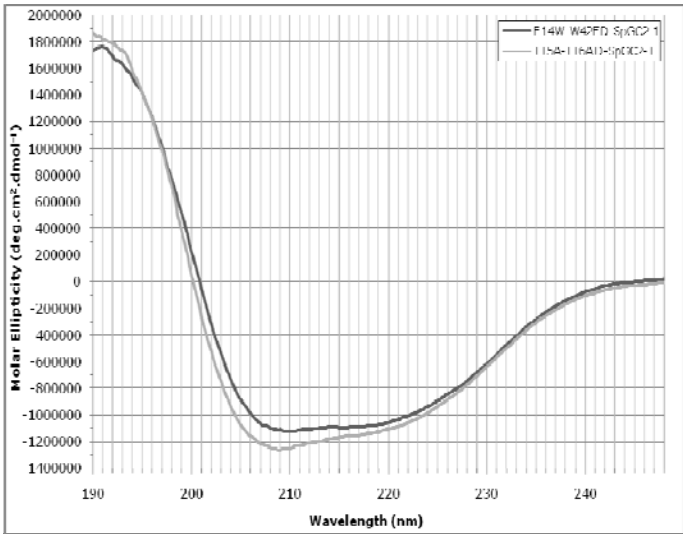


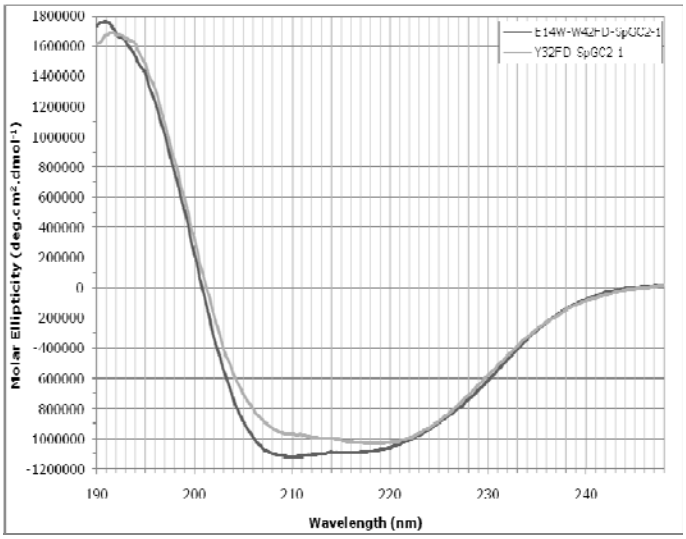
Figure 4.11 – <u>The far-UV CD spectra measured for Fab D-SpG$_{C2}$-1 mutants</u>

Far-UV CD spectra of purified 100 µM protein samples, as indicated, in 20 mM potassium phosphate buffer pH 8.0 are compared with wild-type. Measurements were taken at room temperature in a 0.01 cm path length cell. Nine scans were collected from 190 to 250 nm, averaged and baseline corrected by subtraction of the spectrum of buffer. Each spectrum was analysed as molar ellipticity (deg.cm^2.dmol^{-1}).

(a)



(b)



(c)

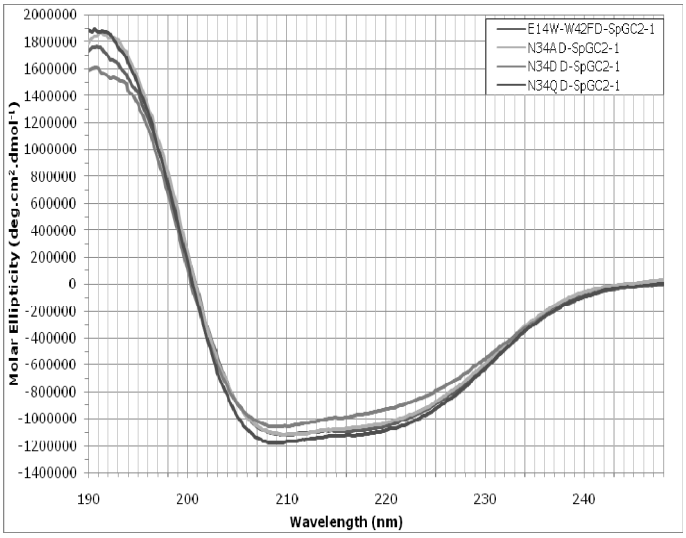


Figure 4.12 – <u>The far-UV CD spectra measured for Fab D-SpG$_{C2}$-1 mutants</u>

Far-UV CD spectra of purified 100 µM protein samples, as indicated, in 20 mM potassium phosphate buffer pH 8.0 are compared with wild-type. Measurements were taken at room temperature in a 0.01 cm path length cell. Nine scans were collected from 190 to 250 nm, averaged and baseline corrected by subtraction of the spectrum of buffer. Each spectrum was analysed as molar ellipticity (deg.cm^2.dmol^{-1}).

4.5.8 – Near-UV CD analysis of D-SpG$_{C2}$-1 mutant constructs engineered to probe the interaction with Fab

Purified D-SpG$_{C2}$-1 mutant constructs were analysed by near-UV CD to determine if the mutation had affected the secondary structure of the protein. Figure 4.13 presents a comparison between wild-type and E14W-W42F D-SpG$_{C2}$-1. This near-UV CD spectrum shows that the insertion of a tryptophan reporter group at the alternative position of E14W to allow the interaction between SpG and Fab to be probed, compares very well to that of wild-type D-SpG$_{C2}$-1. Thus, the presence of the tryptophan 14 compensates for the spectral changes noted when tryptophan 42 is substituted by phenylalanine.

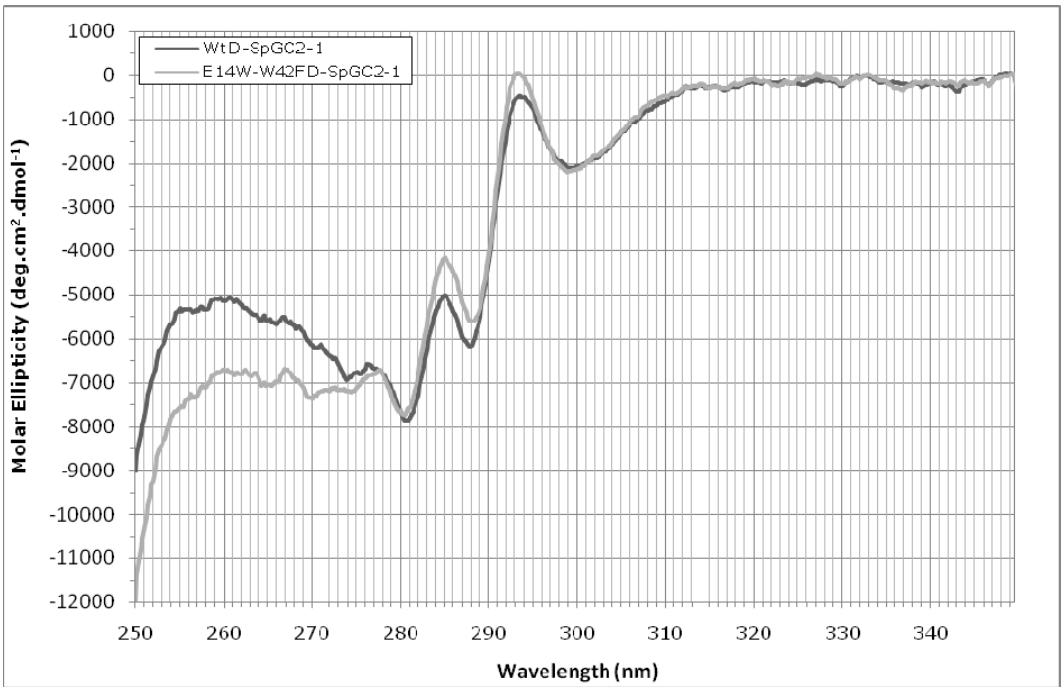


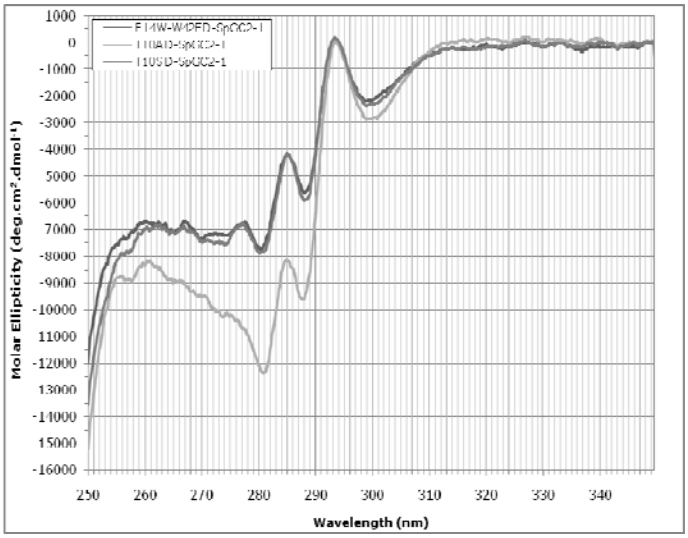
Figure 4.13 – Near-UV CD spectra measured for Wt and E14W-W42F D-SpG$_{C2}$-1

Near-UV CD spectrum of purified 100 µM Wt and E14W-W42F D-SpG$_{C2}$-1 in 20 mM potassium phosphate buffer, pH 8.0 is compared. Measurements were taken at room temperature in a 0.5 cm path length cell. Nine scans were collected from 250 to 350 nm, averaged and baseline corrected by subtraction of the spectrum of buffer. The spectrum was analysed as molar ellipticity (deg.cm^2.dmol^{-1})

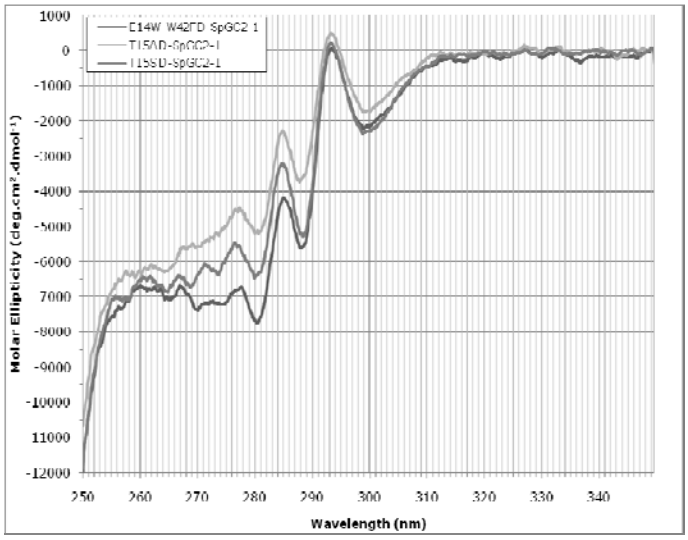
Further near-UV spectra of engineered D-SpG$_{C2}$-1 mutant constructs are shown in figures 4.14 and 4.15, and compared to the CD spectrum obtained for the E14W-W42F D-SpG$_{C2}$-1 protein.

The CD data suggest that the T10A D-SpG$_{C2}$-1 mutation may lead to serious disruption of the natural folding of the domain indicated by both the far and near UV spectral changes.

(a)



(b)



(c)

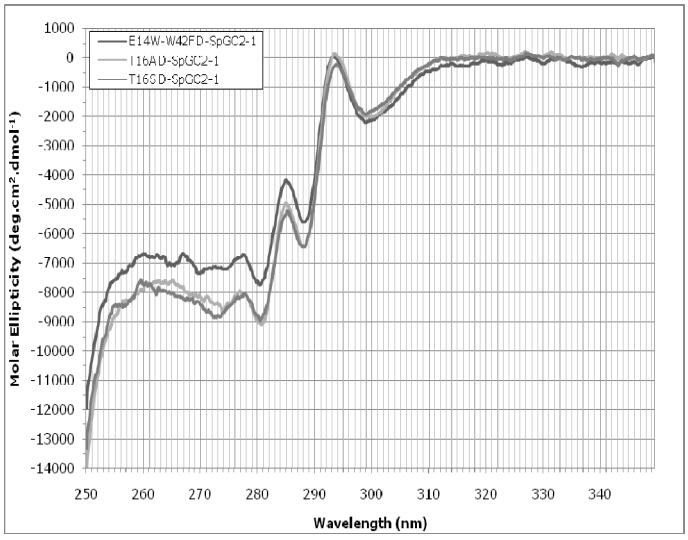
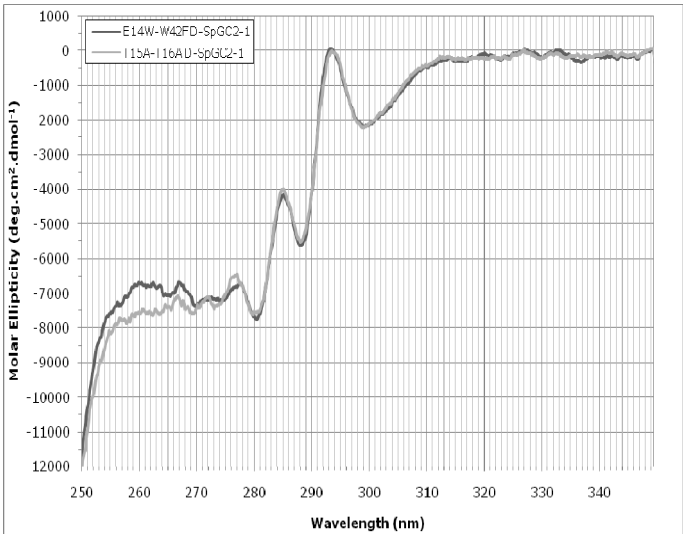


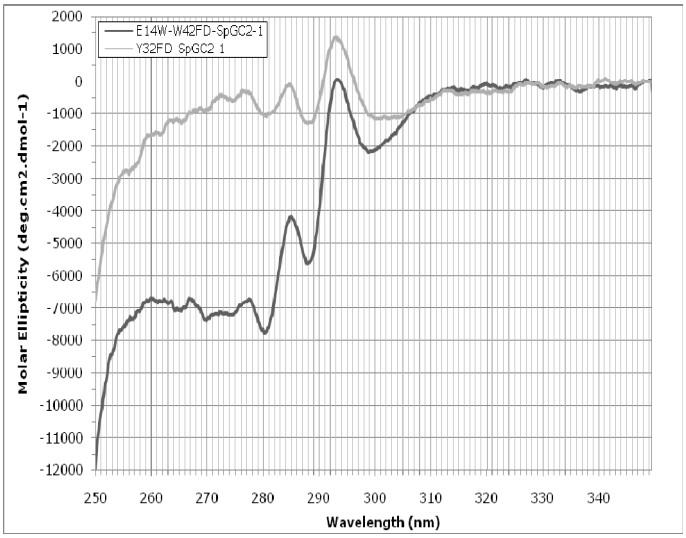
Figure 4.14 – <u>The near-UV CD spectra measured for Fab D-SpG$_{C2}$-1 mutants</u>

Near-UV CD spectra of purified 100 µM protein samples, as indicated, in 20 mM potassium phosphate buffer pH 8.0 are compared with wild-type. Measurements were taken at room temperature in a 0.5 cm path length cell. Nine scans were collected from 250 to 350 nm, averaged and baseline corrected by subtraction of the spectrum of buffer. Each spectrum was analysed as molar ellipticity (deg.cm^2.dmol^{-1}).

(a)



(b)



(c)

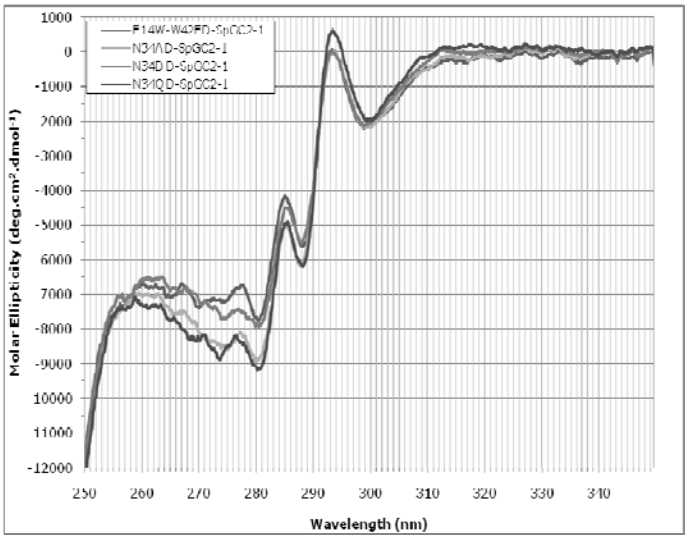


Figure 4.15 – <u>The near-UV CD spectra measured for Fab D-SpG$_{C2}$-1 mutants</u>

Near-UV CD spectra of purified 100 µM protein samples, as indicated, in 20 mM potassium phosphate buffer pH 8.0 are compared with wild-type. Measurements were taken at room temperature in a 0.5 cm path length cell. Nine scans were collected from 250 to 350 nm, averaged and baseline corrected by subtraction of the spectrum of buffer. Each spectrum was analysed as molar ellipticity (deg.cm^2.dmol^{-1}).

4.6 – Determination of conformational stability using denaturation with Gdn-HCl

The reverse of the folding process is known as protein denaturation, whereby the native structure of a protein is disrupted. The unfolded state (U) can be defined as the maximally unfolded state of a protein, in which the backbone NH groups are fully solvent exposed, and the denatured state (D) can be defined as the lowest energy non-native state under a defined set of conditions. The denatured state could also be a folding intermediate if placed in conditions that promote folding.

A fully denatured protein lacks both tertiary and secondary structure and exists as a random coil ensemble of unfolded structures. In certain solutions and under some conditions proteins will not fold into their biologically functional forms. High temperatures, high concentrations of solutes and extremes of pH will cause proteins to unfold or denature. A protein's surroundings also play an essential role in the final conformation of the protein. An aqueous environment favours hydrophilic residues on the surface, whereas non-aqueous or non-polar environment favours hydrophobic residues on the surface.

Studies of protein folding have shown that small proteins can be denatured and refolded in a reversible fashion, using denaturants such as heat, urea, or guanidine hydrochloride (Gdn-HCl) [Anfinsen, 1973; Ahmad *et al*., 1992]. Since the denatured states are less compact than the native state, the viscosity of the solution can be used as a measure of denaturation/renaturation. Likewise, the amino acid side-chains in the differing states i.e. in different environments. The wavelength of maximum fluorescence intensity and the lifetime of the fluorescence decay are very sensitive to the environment of the aromatic amino acids. Changes in the fluorescent properties of the aromatic amino acids, particularly tryptophan, can be used to measure protein folding/unfolding events. This permits the further understanding of protein folding, protein stability, as well as the understanding of the nature of the native and denatured states.

The use of chemical denaturants, such as urea and guanidinium chloride, shown in figure 4.16, has played a central role in protein-folding experiments. Almost all proteins can be unfolded and remain soluble at sufficiently high denaturant concentrations, and both denaturants favour the denatured state by increasing the solubility of the unfolded protein in an aqueous solution.

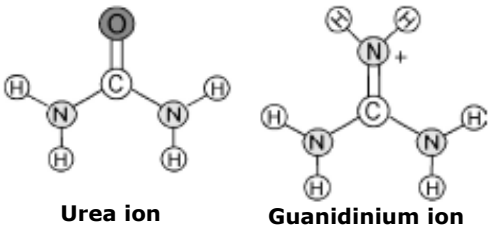


Figure 4.16 – A schematic illustration of urea and guanidinium ions

These illustrations show the two most commonly used chemical denaturants in folding studies, urea and guanidinium ions.

The chemical action of chaotropic agents is still poorly understood. In 8 M urea and in 6 M Gdn-HCl, proteins approach a random coil state, and it is thought that this is a result of the disruption to both the hydrogen bonds and hydrophobic interactions. Another theory of the mechanism can be defined by the denaturant binding model [Tanford, 1968; Pace and Vanderburg, 1979]. In this model the denaturant molecules bind to peptide groups and side-chains, creating a shell around the folded and the unfolded polypeptide chain, presented in figure 4.17. As the concentration of denaturant increases, the unfolded state will be favoured, since it has a larger surface area exposed to the solvent, where more molecules can bind. The driving force for denaturation with chemical denaturants is therefore proportional to the change in exposed surface area going from the native to the denatured state.

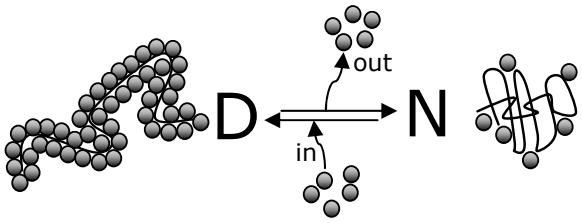


Figure 4.17 – The denaturant binding model

This denaturant binding model represents chemical denaturation. The grey dots represent the chemical denaturant and the black lines represent the various states of the protein polypeptide. The denaturant molecules bind to peptide groups and side-chains and thereby create a shell around the native (N) folded and the unfolded denatured (D) polypeptide chain.

In addition, chemical denaturation is often a reversible process. This is possible since the hydrophobic groups of the unfolded chain are shielded by the denaturants, which prevent aggregation. The fact that the process is reversible and that the protein approaches a maximum unfolded state is important to allow the conformational transitions to be measured.

4.6.1 - Calculating the conformational stability of D-SpG$_{C2}$-1 by equilibrium denaturation with Gdn-HCl

Further characterisation studies using solvent denaturation were carried out to determine the conformational stability of a domain of wild-type SpG and mutant constructs. By using spectral techniques to follow the process, equilibrium thermodynamic and kinetic measurements were made.

A two-state model of denaturation for a single domain protein is employed in these studies, and refers to the reversible disruption of protein structure where only the fully folded native and the fully unfolded denatured states are populated, as shown below.

Folded (F) = Unfolded (U)

If, for example, fluorescence spectroscopy is used to follow the unfolding and assuming the denatured protein fully refolds when the denaturant is removed, the fraction of unfolded protein, F_U, can be calculated using equation 4.1.

$$\mathbf{F_U = (y_F - y)/(y_F - y_U)} \qquad (\textit{equation 4.1})$$

and the fraction of refolded protein F_F, is given using equation 4.2 [Pace, 1986].

$$\mathbf{F_F = (y - y_U)/(y_F - y_U)} \qquad (\textit{equation 4.2})$$

Where, y is the observed fluorescence at a chosen concentration of denaturant, and y_F and y_U are the fluorescence characteristics of the folded and unfolded conformation, respectively. y_F and y_U are determined by the extrapolation of the pre- and post-transition region of the unfolding curve, occurring at low and high concentrations of denaturant, respectively.

The equilibrium, K_{eq}, for the unfolding process is given by equation 4.3 [Pace, 1986].

$$\mathbf{K_{eq} = F_U/F_F = (y_F - y)/(y - y_U)} \qquad (\textit{equation 4.3})$$

The stability of a protein's conformation is equal to the difference in free energy between the folded and unfolded denatured states, ΔG_U, and can be calculated from equation 4.4.

$$\mathbf{\Delta G_U = -RT\ ln\ K_{eq}} \qquad (\textit{equation 4.4})$$

Where K_{eq} is the equilibrium constant for unfolding in the two-state model, calculated using equation 4.3. Combining these two equations, equation 4.5 can be derived.

$$\Delta G_U = -RT \ln K_{eq} = -RT \ln [(y_F - y)/(y - y_U)] \quad (equation\ 4.5)$$

Therefore, the difference in free energy between the folded and unfolded states (ΔG_U) and hence, the conformational stability of a protein can be derived from the fraction of unfolded and folded protein (K_{eq}). There is a linear dependence of ΔG_U with the denaturation concentration along the transition region of the unfolding curves, see equation 4.6.

$$\Delta G_U = \Delta G_U^0 - m\,[\text{Gdn-HCl}] \quad (equation\ 4.6)$$

Thus, in a plot of ΔG_U against denaturant concentration, the slope of the line yields *m*, the dependence of ΔG_U on the denaturant concentration ($kJmol^{-1}M^{-1}$), and the intercept yields ΔG_U^0 ($kJmol^{-1}$), the difference in free energy between the unfolded and folded state in the absence of denaturant. In addition, the mid-point of unfolding is given by equation 4.7, which is the concentration of denaturant that corresponds to when the protein is 50% unfolded.

$$[\text{Gdn-HCl}]_{½} = \Delta G_U^0 / m \quad (equation\ 4.7)$$

Thus, the ΔG_U^0 value can also be calculated and compared with the estimate given graphically. The calculated ΔG_U^0 value is given by multiplying the *m* value by the mid-point value.

4.6.2 – Determination of conformational stability of Wt D-SpG$_{C2}$-1 by equilibrium denaturation with Gdn-HCl

Denaturation studies on Wt D-SpG$_{C2}$-1 were carried out using guanidine hydrochloride at a stock solution of 6 M solution in 20 mM potassium phosphate buffer, which was filtered prior to use. For accurate determination of guanidine hydrochloride concentrations after filtration, the refractive index of the solutions were measured and calculated using the following equation [Pace *et al*., 1986]:

$$[\text{Gdn-HCl}] = 57.147(\Delta N) + 38.68(\Delta N)^2 - 91.60\,(\Delta N)^3 \quad (equation\ 4.8)$$

Where, ΔN is the difference between the refractive index of the denaturant and the refractive index of the buffer used.

To ascertain that the assumption of the reversible denaturation of protein structure is correct, refolding studies were carried out on Wt D-SpG$_{C2}$-1. A 1 ml solution containing 10 µM Wt D-SpG$_{C2}$-1 in 6M Gdn-HCl was added to 1ml of 10 µM Wt D-SpG$_{C2}$-1 in 0 M Gdn-HCl. The 2 ml solution (3 M) was left to equilibrate and induce protein refolding before its fluorescence emission spectrum

was determined. The spectrum of the refolded protein was compared with the fluorescence emission spectrum of Wt D-SpG$_{C2}$-1 in 0 M, 3 M and 6 M Gdn-HCl. Figures 4.18 and 4.19 present fluorescence emission spectra for both the Wt and E14W-W42F D-SpG$_{C2}$-1 constructs, which also served as templates for the other mutant constructs.

These figures show that 6 M Gdn-HCl causes an approximate 75% drop in fluorescence intensity and a red shift of between 2 and 10 nm for 3 M and 6 M Gdn-HCl, respectively. This is as a result of a significant solvent change, hydrophobic to hydrophilic, that occurs around residues in the interior of the protein when a protein is denatured. The tryptophan in D-SpG$_{C2}$-1, (W42 in the wild-type construct and W14 in the E14W-W42F construct), absorbs light at 275-295 nm and re-emits at 330-340 nm. Its fluorescence is sensitive to polarity changes in the environment, and when exposed to a non-polar solution (hydrophobic core) the fluorescence intensity is normally high, and when transferred to a more polar solvent during unfolding, the signal is usually reduced [Alston *et al*., 2004]. At the same time the fluorescence maximum is shifted to a longer wavelength [Swaminathan *et al*., 1994]. With the E14W-W42F construct, the fluorescence maximum is shifted to a longer wavelength when compared to that obtained for the wild-type protein. This suggests the tryptophan at position 14 is more solvent exposed in the unfolded state.

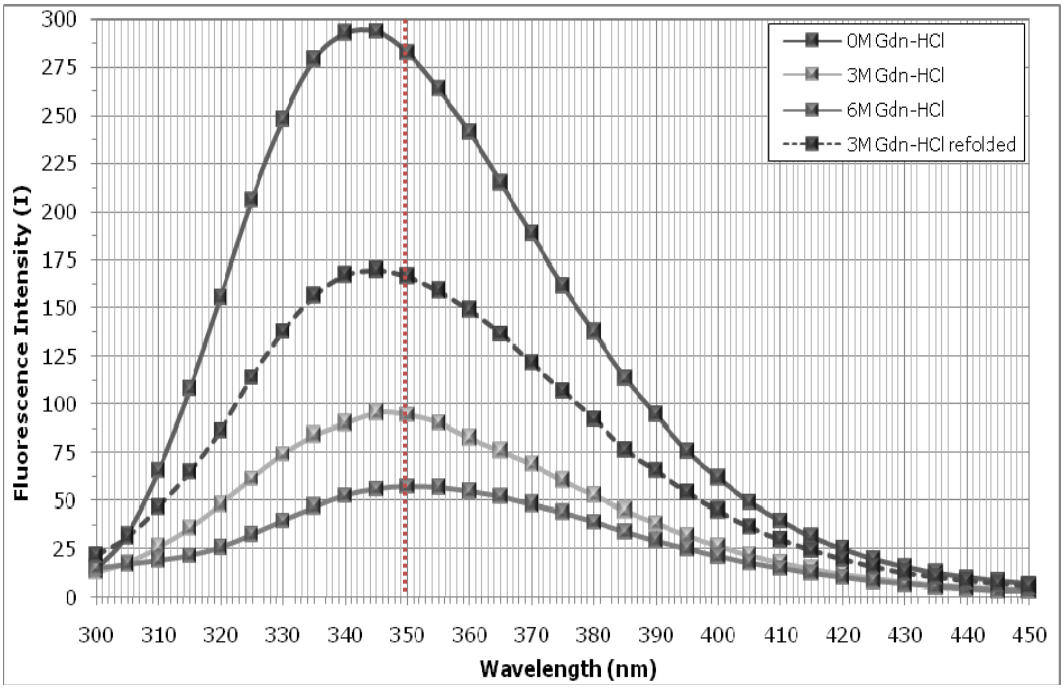


Figure 4.18 - Emission spectra for Wt D-SpG$_{C2}$-1 in Gdn-HCl

The effect of increasing the concentration of denaturant on the emission spectra, obtained for 10 µM Wt D-SpG$_{C2}$-1 in 0 M, 3 M and 6 M Gdn-HCl, and the refolding of 3 M. Protein samples were excited at 280 nm and fluorescence emissions were measured between 300 and 450 nm. All spectra were measured in triplicate and corrected for the buffer fluorescence. The red line denotes the intercept of the spectrum at 350 nm, which is used to determine a denaturation curve.

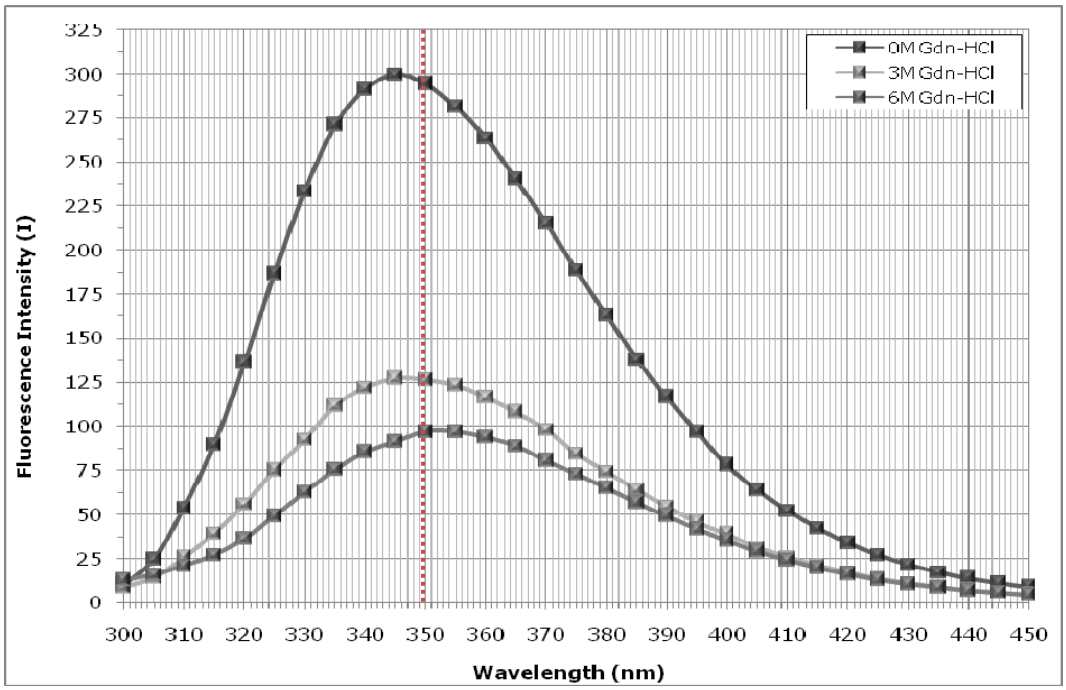


Figure 4.19 - Emission spectra for E14W-W42F D-SpG$_{C2}$-1 in Gdn-HCl

The effect of increasing the concentration of denaturant on the emission spectra, obtained for 10 µM E14W-W42F D-SpG$_{C2}$-1 in 0 M, 3M and 6 M Gdn-HCl. Protein samples were excited at 280 nm and fluorescence emissions were measured between 300 and 450 nm. All spectra were measured in triplicate and corrected for the buffer fluorescence. The red line denotes the intercept of the spectrum at 350 nm, which is used to determine a denaturation curve.

Figure 4.20 shows a typical Gdn-HCl denaturation curve for Wt D-SpG$_{C2}$-1, and figure 4.21 presents a typical plot demonstrating the linear dependence of ΔG_F against Gdn-HCl concentration. Data were analysed using equations 4.1 to 4.8.

The linear dependence of ΔG_F against Gdn-HCl concentration for wild-type D-SpG$_{C2}$-1 allows parameters of equilibrium Gdn-HCl denaturation to be obtained. The value of ΔG_U^0 was calculated from the mid-point concentration of Gdn-HCl and the *m* value, and an average value of 14.09 ±1.83 kJmol^{-1} was obtained. This gives the conformational stability of Wt D-SpG$_{C2}$-1, i.e. the difference in free energy of the unfolded state relative to that of the folded state. The intercept on the graph of ΔG_U^0 against denaturant concentration suggests a higher value of 15.5 kJmol^{-1}. The relationship of plotting ΔG_U^0 in the transition region against the concentration of Gdn-HCl was first seen by Tanford [Tanford, 1968] and gives an empirical straight line with a slope. In these studies, the proportionality factor across the transition region (*m*) was found to be -5.87 ±0.56 kJmol^{-1}M^{-1}, which reflects the ratio between the number of bound molecules in the two states, whereby the denaturant binds to the protein and favours the unfolded state. Proteins that possess high *m* values are more sensitive to denaturants and possess steeper transition slopes [Lindberg *et al*., 2002]. The concentration of Gdn-HCl at which half the protein denatures (ΔG_U=0) is 2.40 ±0.08 M. At this transition mid-point, the stability of the denatured state and the native state are the same and they are equally populated at 2.40 M Gdn-HCl. These values can be used to compare the conformational stabilities of the D-SpG$_{C2}$-1 mutant constructs.

Comparative studies by Walker [1994] gave values of -7.26 kJmol^{-1}M^{-1} and 3.09 M under similar conditions for the proportionality factor across the transition region (*m*) and the concentration of Gdn-HCl when half the protein had denatured ([Gdn-HCl]½), respectively.

Other related studies by Alexander [Alexander *et al*., 1992] provides ΔG values calculated by differential scanning calorimetry of approximately 28 and 23 kJmol^{-1} at 25 °C for the C1 and C3 binding domains of SpG, respectively. In addition, previous work by Walker on the C2 binding domain [Walker, 1994] gives a ΔG value of 22.46 kJmol^{-1} obtained at 25 °C, pH 6.0 by equilibrium Gdn-HCl denaturation. Both of these values are higher than the ΔG value 14.09 kJmol^{-1} obtained here for the C2 binding domain of SpG by equilibrium Gdn-HCl denaturation, at 15 °C, pH 8.0. This is most likely to be due to the different

experimental conditions. Furthermore, work conducted here and by Walker [1994] is based on the C2 domain, which is a hybrid of the C-terminus of the C1 domain and the N-terminus of the C3 domain, and it is unlikely that the different amino acid sequence of the hybrid will not permit comparison with the C1 and C3 domains. In addition, these results could indicate that the introduction of a 6x His-tag at the N-terminal end of the protein has slightly altered the effect of Gdn-HCl on the denaturation of D-SpG$_{C2}$-1.

Further studies by Kuszewski and co-workers [Kuszewski *et al*., 1994] have employed equilibrium unfolding studies to establish the equilibrium unfolding curve of the B1 domain of SpG as a function of Gdn-HCl at pH 4.0, 5 °C. Although experimental conditions vary in these studies, results obtained show that the unfolding transition is a two-state mechanism, with the transition mid-point occurring at 3 M Gdn-HCl. Data also gives a ΔG_U^0 intercept value of 4.8 kcal mol^{-1}, which is equal to 20.08 kJmol^{-1}. This also compares well with data obtained by Alexander [Alexander *et al*., 1992].

It was also found that both the pre- and post-transition regions of the denaturation curve did not always give a flat baseline. This was also found in previous denaturation studies on D-SpG$_{C2}$-1 by Walker [Walker, 1994]. Alternatively, other studies have shown an independence of y_F and y_U on Gdn-HCl concentration [Jackson and Fersht, 1991; Pace *et al*., 1988], and it is now accepted that the slope of the baseline should be taken into consideration during the calculation of f_U and f_F for each concentration of Gdn-HCl used [Jackson *et al*., 1993; Pace *et al*., Lim *et al*., 1992].

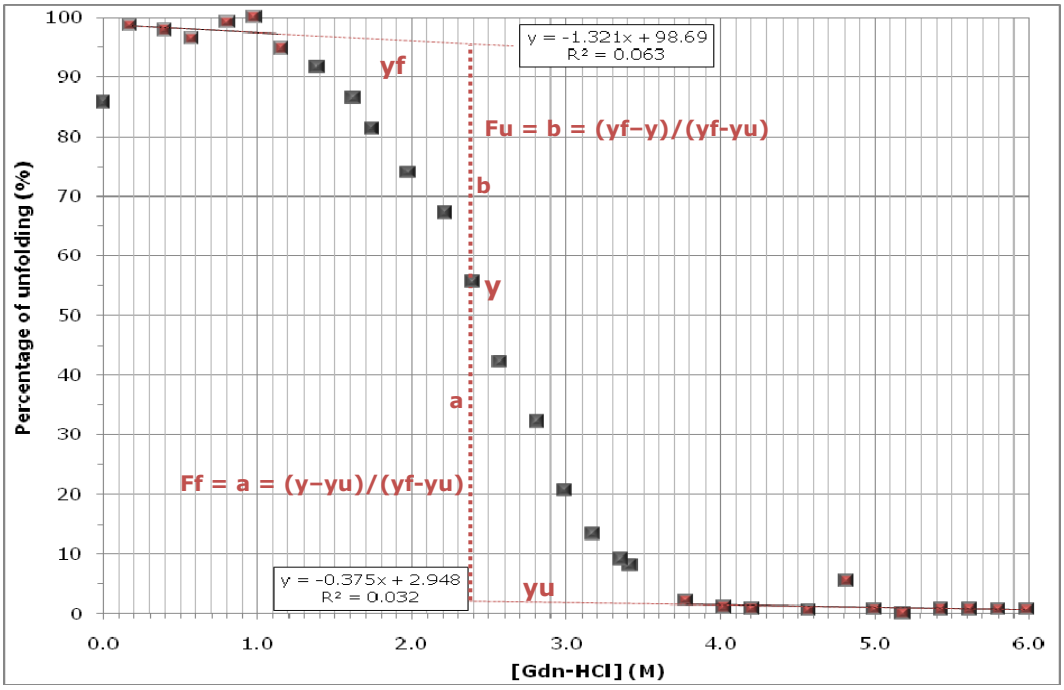


Figure 4.20 – A typical Gdn-HCl denaturation curve for Wt D-SpG$_{C2}$-1

The denaturation curve shows the result of plotting the percentage of unfolding of Wt D-SpG$_{C2}$-1 at 350 nm against concentrations of Gdn-HCl, pH 8.0 at 15 °C. From the sigmoidal curve obtained, pre- (yf), post- (yu) and transitional (y) regions are defined and are used to calculate the free energy change (ΔG_U) between folded and unfolded conformations.

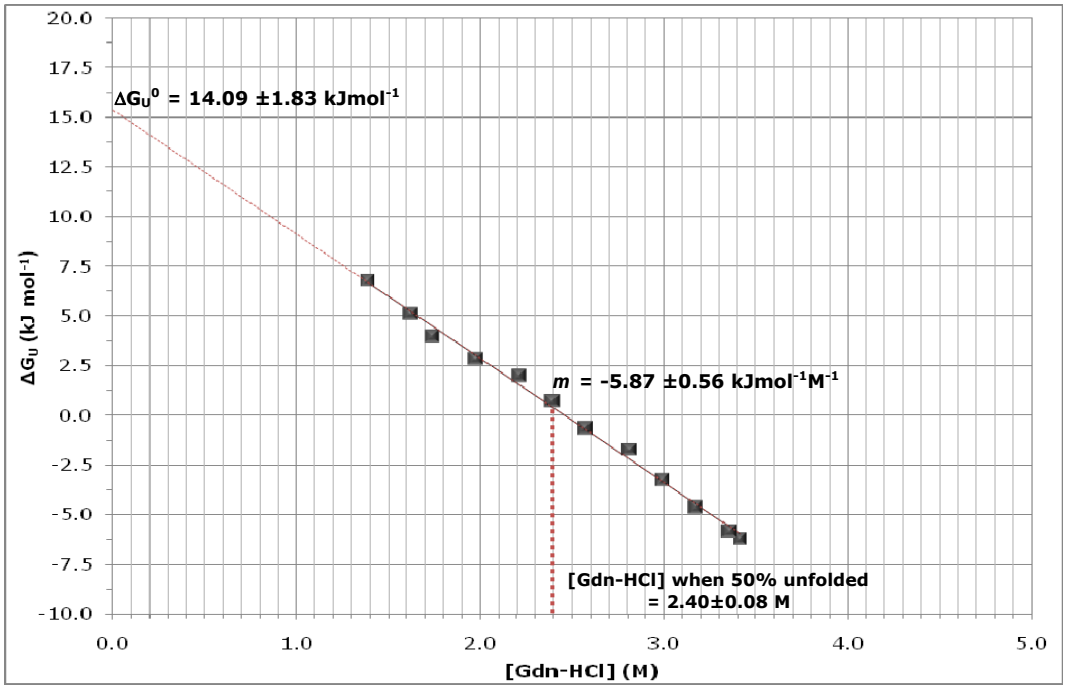


Figure 4.21 – A typical linear relationship between ΔG_U and [Gdn-HCl] for Wt D-SpG$_{C2}$-1

The plot shows the linear relationship betweenΔG_U and the concentration of Gdn-HCl for the Wt D-SpG$_{C2}$-1. From the denaturation curve presented in figure 4.21, the transitional region is used to calculate the free energy change ΔG_U between folded and unfolded conformations using the equation $\Delta G_U = \Delta G_U^0 - m$[Gdn-HCl]. ΔG_U^0 is given as ΔG_U in the absence of denaturant.

4.6.3 - Determination of conformational stability of D-SpG$_{C2}$-1 mutant constructs designed to perturb Fc binding interactions

Equilibrium denaturation studies were also carried out on mutant constructs of D-SpG$_{C2}$-1 to determine the effect of these mutations on the conformational stability of the domain. This was important since any changes in binding affinity might be due to differently folded proteins. Mutational studies offer a versatile means to probe protein structure, but point mutations may result in small, but measureable changes in structure or stability. In contrast, some proteins may be very tolerant to amino acid substitutions and it is uncommon for a single site change to completely alter the basic fold of a protein. However, mutations may affect one or more of the parameters associated with the native state and function of a protein. For example, some mutations may inhibit protein binding without dramatically affecting the stability of the protein. In contrast, other mutations may directly affect the ability of a protein to fold correctly and retain its native structure [Dinner *et al*., 2000].

As before, controlled unfolding measurements were made with 10 μM Fc D-SpG$_{C2}$-1 mutant constructs in Gdn-HCl at increasing concentrations from 0-6 M, and the final Gdn-HCl concentration was determined by refractometry using equation 4.8. The change in protein fluorescence was followed using an excitation wavelength of 280 nm; emission was measured at 350 nm. Each sample was pre-incubated at 15 °C for two hours to reach thermal equilibrium before being measured.

Six of the eight residues within the α-helix and β-strand 3 of a domain of SpG that have been identified to be involved in the interfacial interactions with the loop regions of Fc (Glu26, Lys27, Lys30, Gln31, Asn34, Asp39, Glu41 and Trp42) [Sauer-Eriksson *et al.*, 1995] were mutated to different side-chains to characterise the interaction. However, Lys27 did not express protein so could not be characterised any further.

Data were again analysed using equations 4.1 to 4.8. Typical plots demonstrating the linear dependence of ΔG_F against Gdn-HCl concentration of engineered D-SpG$_{C2}$-1 Fc mutant constructs are shown in figures 4.22 to 4.23, and compared to that obtained for the Wt D-SpG$_{C2}$-1 protein. The parameters characterising the equilibrium denaturation of the D-SpG$_{C2}$-1 Fc mutant constructs by Gdn-HCl are summarised in table 4.1.

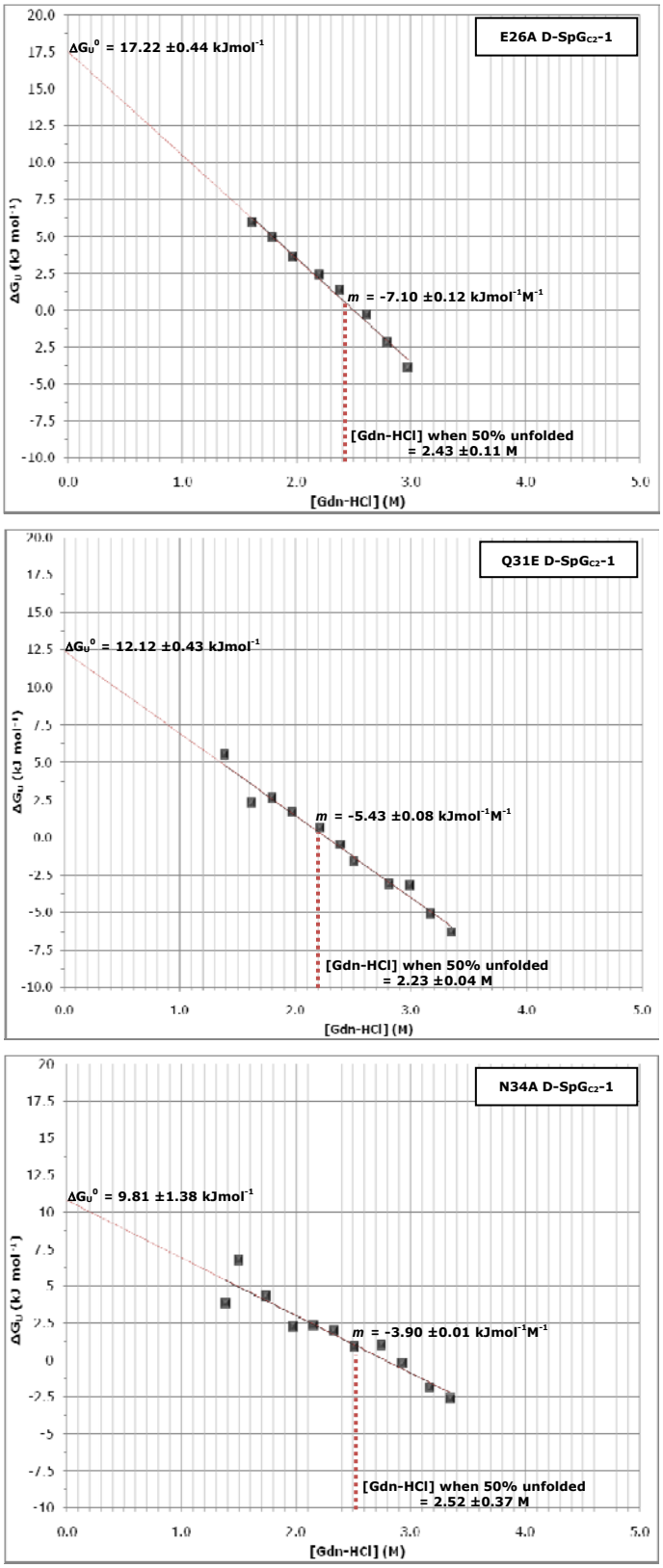


Figure 4.22 – A typical linear relationship between ΔG_U and [Gdn-HCl] for Fc D-SpG$_{C2}$-1 mutants

The plot shows the linear relationship between ΔG_U and the concentration of Gdn-HCl for the protein samples indicated. From the denaturation curve, the transitional region is used to calculate the free energy change ΔG_U between folded and unfolded conformations using the equation $\Delta G_U = \Delta G_U^0 - m[\text{Gdn-HCl}]$. ΔG_U^0 is given as ΔG_U in the absence of denaturant.

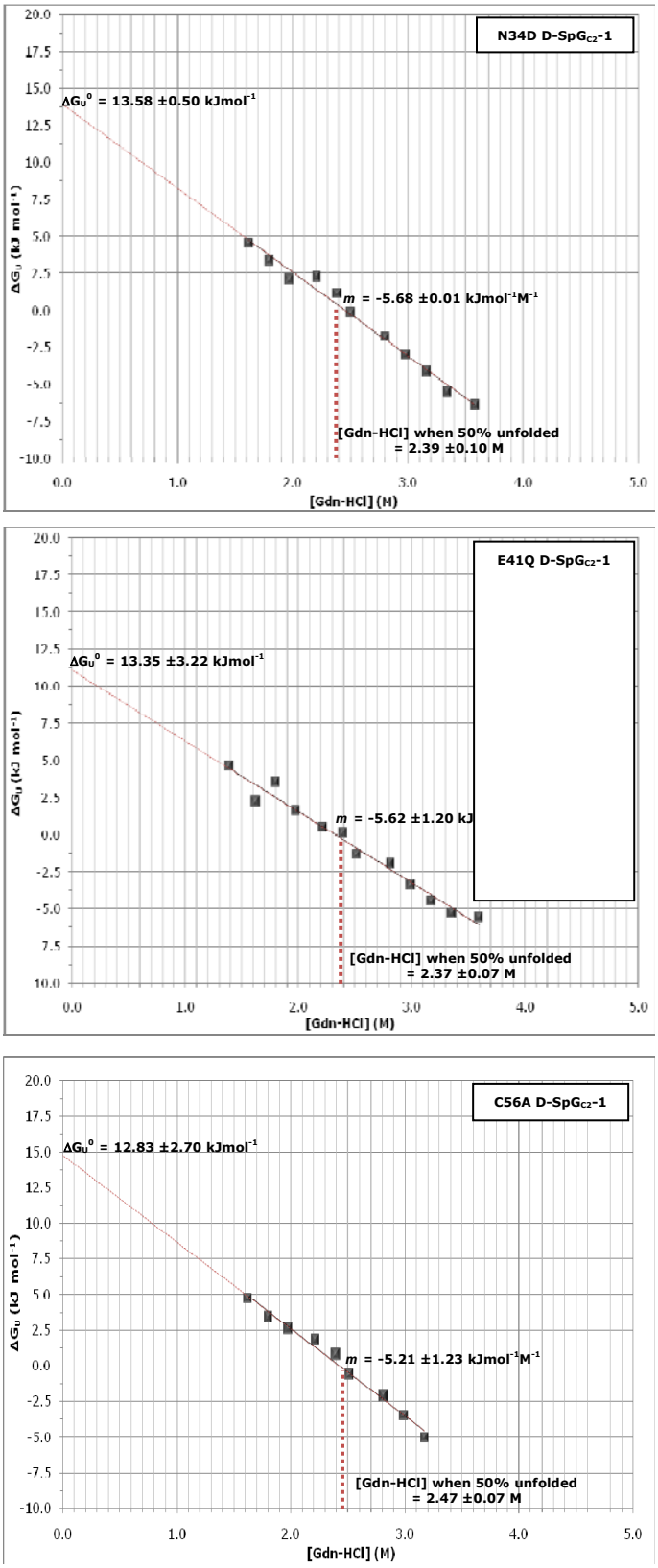


Figure 4.23 – A typical linear relationship between ΔG_U and [Gdn-HCl] for Fc D-SpG$_{C2}$-1 mutants

The plot shows the linear relationship between ΔG_U and the concentration of Gdn-HCl for the protein samples indicated. From the denaturation curve, the transitional region is used to calculate the free energy change ΔG_U between folded and unfolded conformations using the equation $\Delta G_U = \Delta G_U^0 - m[\text{Gdn-HCl}]$. ΔG_U^0 is given as ΔG_U in the absence of denaturant.

D-SpG_{C2}-1 construct	ΔG_U^0 intercept ($kJ.mol^{-1}$) ± std dev. n≥2	ΔG_U^0 calculated ($kJ.mol^{-1}$) ± std dev. n≥2	*m* ($kJ.mol^{-1}M^{-1}$) ± std dev. n≥2	[Gdn-HCl] 50% unfolded (M) ± std dev. n≥2
WT	14.09 ±1.83	14.26	-5.87 ±0.56	2.40 ±0.08
E26A	17.22 ±0.44	17.28	-7.10 ±0.12	2.43 ±0.11
Q31E	12.12 ±0.43	12.10	-5.43 ±0.08	2.23 ±0.04
N34A	9.81 ±1.38	9.82	-3.90 ±0.01	2.52 ±0.37
N34D	13.58 ±0.50	13.58	-5.68 ±0.01	2.39 ±0.10
E41Q	13.35 ±3.22	13.31	-5.62 ±1.20	2.37 ±0.07
W42F	ND	ND	ND	ND
W42Y	ND	ND	ND	ND
C56A	12.83 ±2.70	12.87	-5.21 ±1.23	2.47 ±0.07

Table 4.1 – Summary of the parameter values characterising the Gdn-HCl equilibrium denaturation of D-SpG_{C2}-1 mutant constructs designed to perturb Fc binding interactions

These results demonstrate the effect of each mutation on the parameters of unfolding. Mutant constructs W42F and W42Y D-SpG_{C2}-1 were omitted from these studies as the tryptophan residue, which acts as a fluorescent probe, has been removed. This results in a non-fluorescent protein, which makes fluorimetric determination of equilibrium denaturation difficult. Replacing tryptophan with a tyrosine residue, although an aromatic amino acid, is not fluorescent enough to act as a probe in these studies. Comparison of the parameters characterising equilibrium Gdn-HCl unfolding obtained for the majority of the mutant constructs studied, compare reasonably well with values obtained for wild-type D-SpG_{C2}-1. Mutants Q31E, N34D, E41Q and C56A have little effect on the conformational stability of D-SpG_{C2}-1. E26A D-SpG_{C2}-1, however, has an approximate 1.5-fold increased conformational stability value from 14.09 ±1.83 to 17.22 ±0.44 $kJ.mol^{-1}$ when compared to WT D-SpG_{C2}-1. This increase is due to an increased *m* value rather than a higher [Gdn-HCl]½ value, which suggests that the mutant is more co-operatively folded. N34A D-SpG_{C2}-1 shows a decreased conformational stability, with an approximate 1.7-fold decreased ΔG_U^0 intercept value from 14.09 ±1.83 to 9.81 ±1.38 $kJ.mol^{-1}$ when compared to wild-type D-SpG_{C2}-1. Thus, this mutation may well have led to a disturbed secondary structure with possible consequences to the binding reaction.

4.6.4 - Determination of the conformational stability of D-SpG$_{C2}$-1 mutant constructs designed to perturb Fab binding interactions

Equilibrium denaturation studies were also carried out as before on purified Fab D-SpG$_{C2}$-1 mutant constructs designed to probe the interactions with Fab. Key amino acid residues, identified in both the major and minor regions of contact with the Fab C_H1 domain, were chosen for mutation. The site of interaction spans residues Lys9, Thr10, Leu11, Lys12, Gly13, Glu14, Thr15 and Thr16 from β-strand 2 on SpG, with residues Tyr32, Ala33, Asn34, Asp35, Asn36 and Gly37 from the C-terminal end of the α-helix of SpG constituting the minor contact region [Derrick and Wigley, 1992; Lian *et al.*, 1994; Derrick and Wigley, 1994; Sauer-Eriksson *et al.*, 1995]. Six of these fourteen residues, from both the major and minor contact regions were chosen for mutation, namely residues Thr10, Glu14, Thr15, Thr16, Tyr32 and Asn36. The effect of these mutations on the stability of the domain was investigated.

These mutants all require a tryptophan reporter group other than W42, which needs to be removed to eliminate binding to Fc. Thus, these mutations were made on the E14W-W42F D-SpG$_{C2}$-1 template. Stability studies were therefore, carried out on this "control" protein.

4.6.5 – Determination of the conformational stability of E14W-W42F D-SpG$_{C2}$-1 by equilibrium denaturation with Gdn-HCl

Figure 4.24 shows a typical Gdn-HCl denaturation curve for E14W-W42F D-SpG$_{C2}$-1, and figure 4.25 presents a typical plot demonstrating the linear dependence of ΔG_F against Gdn-HCl concentration.

The value of ΔG_U^0 intercept was calculated as 14.31 ±1.76 kJmol^{-1}, for the conformational stability of E14W-W42F D-SpG$_{C2}$-1, and this compares well with the ΔG_U^0 intercept value calculated as 14.09 ±1.83 kJmol^{-1} for Wt D-SpG$_{C2}$-1. In these studies, the proportionality factor across the transition region (m) was found to be -5.29 ±0.51 kJmol^{-1}M^{-1}, and the concentration of Gdn-HCl at which half the protein had denatured (ΔG_U=0) was 2.70 ±0.07 M. Typical plots demonstrating the linear dependence of ΔG_F against Gdn-HCl concentration of these D-SpG$_{C2}$-1 constructs are shown in figures 4.26 to 4.29, and compared to that obtained for the E14W-W42F D-SpG$_{C2}$-1 protein. The parameters characterising the equilibrium denaturation of these D-SpG$_{C2}$-1 mutant constructs by Gdn-HCl are summarised in table 4.2.

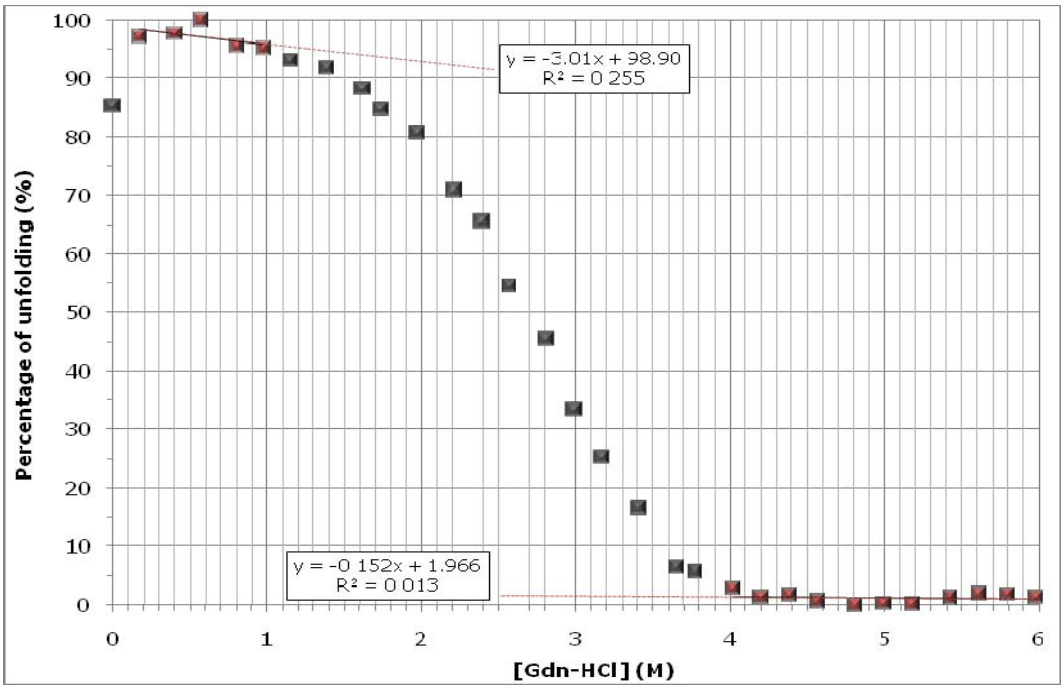


Figure 4.24 – A typical Gdn-HCl denaturation curve for E14W-W42F D-SpG$_{C2}$-1

The denaturation curve shows the result of plotting the percentage of unfolding of E14W-W42F D-SpG$_{C2}$-1 at 350 nm against increasing concentrations of Gdn-HCl, pH 8.0 at 15°C. From the sigmoidal curve obtained, pre-, post- and transitional regions are defined and are used to calculate the free energy change (ΔG_U) between folded and unfolded conformations.

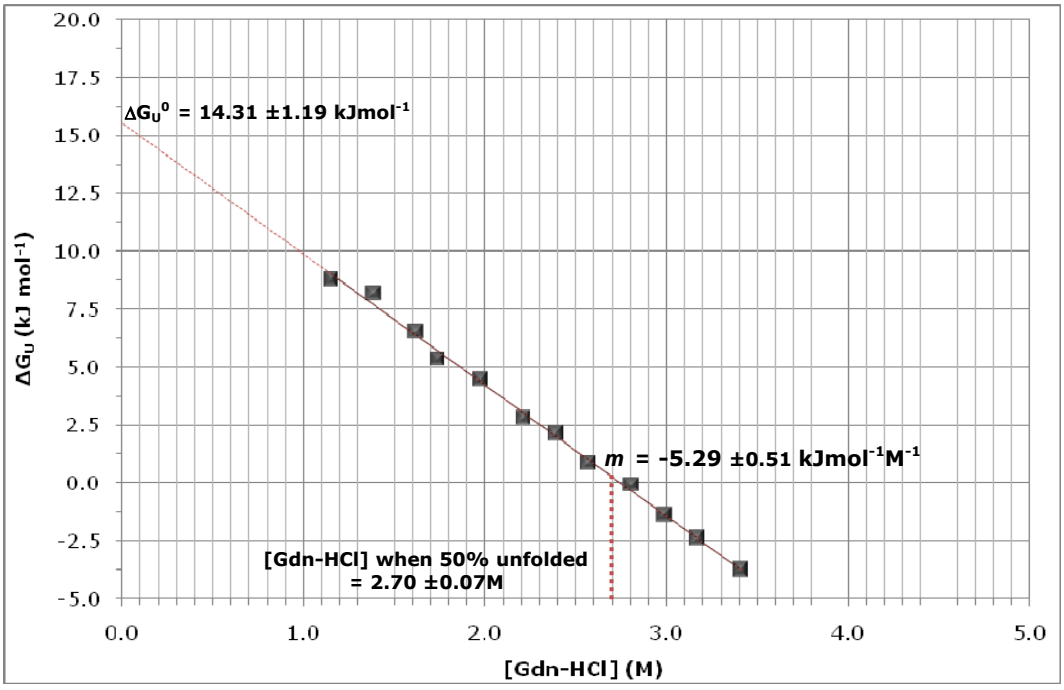


Figure 4.25 – A typical linear relationship between ΔG_U and [Gdn-HCl] for E14W-W42F D-SpG$_{C2}$-1

The plot shows the linear relationship betweenΔG_U and the concentration of Gdn-HCl for the E14W-W42F D-SpG$_{C2}$-1. From the denaturation curve presented in figure 4.52, the transitional region is used to calculate the free energy change ΔG_U between folded and unfolded conformations using the equation $\Delta G_U = \Delta G_U^0 - m$[Gdn-HCl]. ΔG_U^0 is given as ΔG_U in the absence of denaturant.

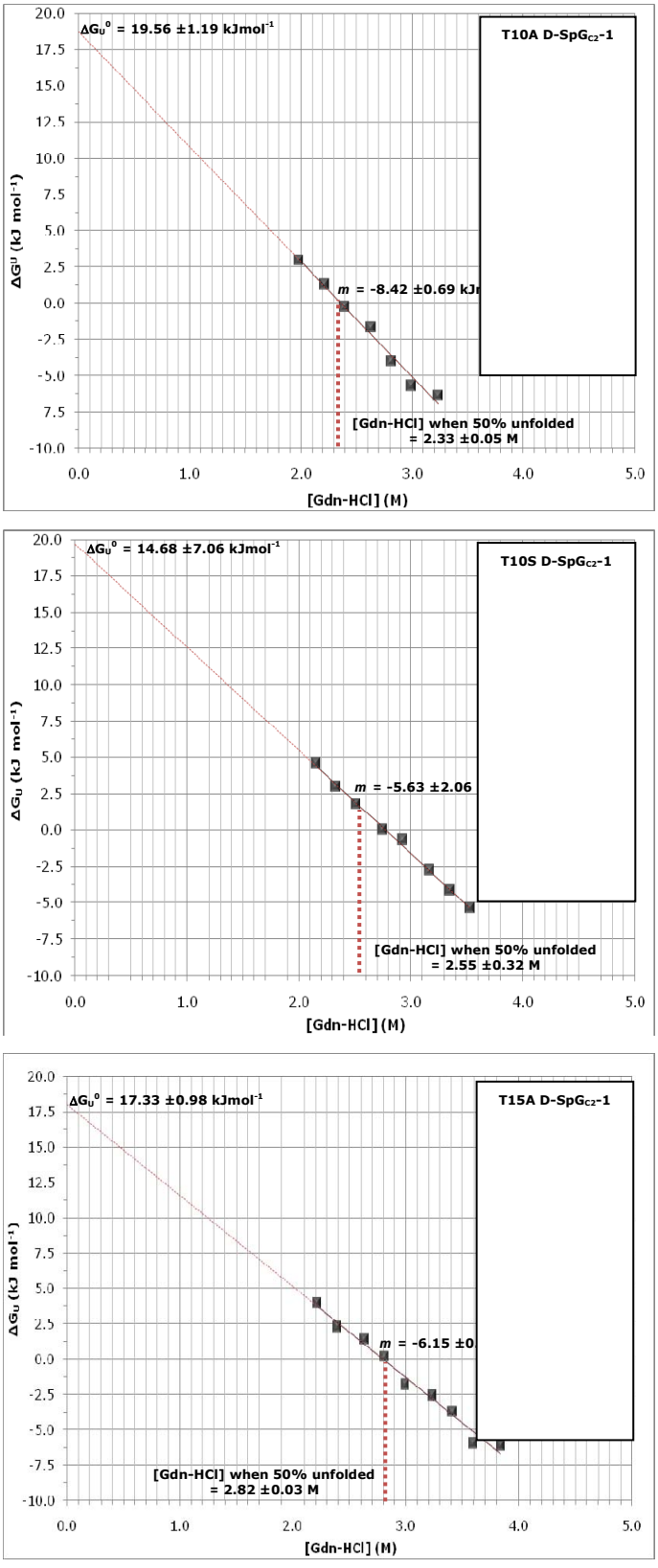


Figure 4.26 – A typical linear relationship between ΔG_U and [Gdn-HCl] for Fab D-SpG$_{C2}$-1 mutants

The plot shows the linear relationship between ΔG_U and the concentration of Gdn-HCl for the protein samples indicated. From the denaturation curve, the transitional region is used to calculate the free energy change ΔG_U between folded and unfolded conformations using the equation $\Delta G_U = \Delta G_U^0 - m[\text{Gdn-HCl}]$. ΔG_U^0 is given as ΔG_U in the absence of denaturant.

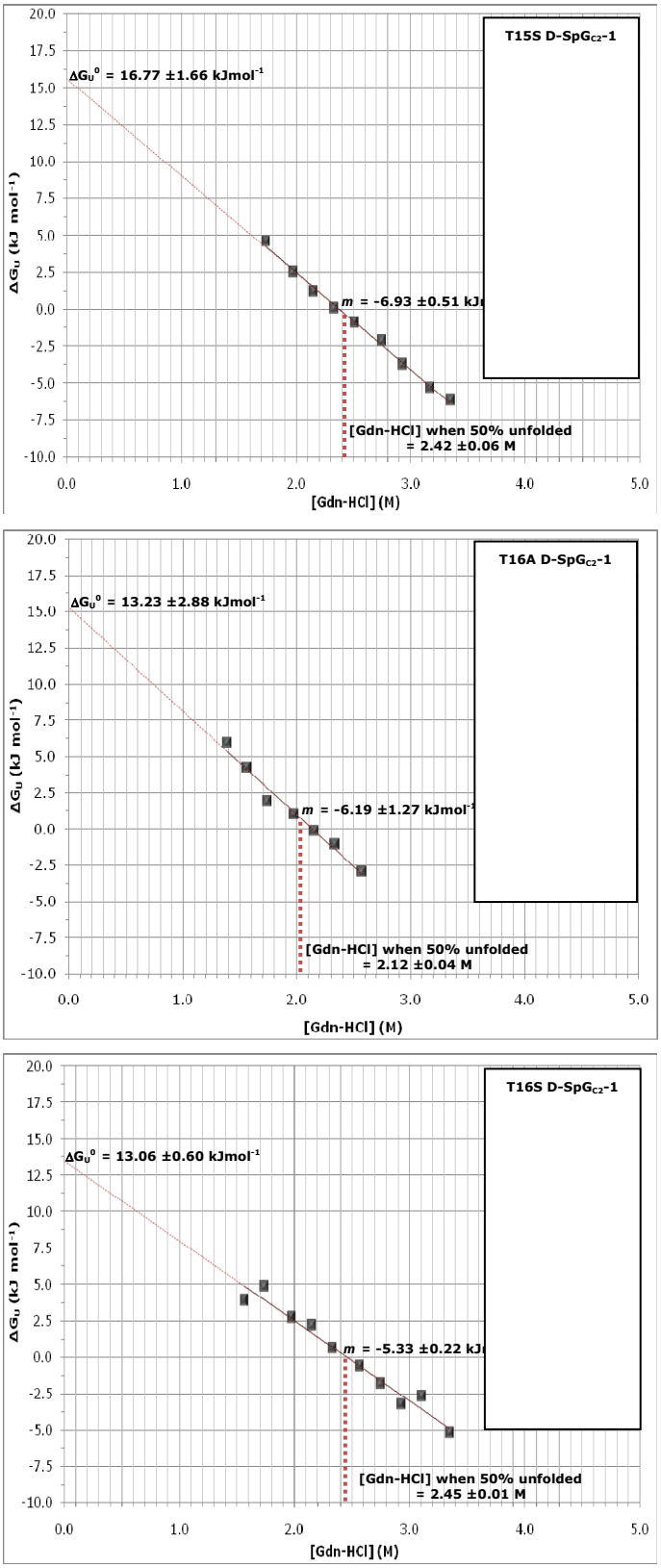


Figure 4.27 – <u>A typical linear relationship between ΔG_U and [Gdn-HCl] for Fab D-SpG$_{C2}$-1 mutants</u>

The plot shows the linear relationship between ΔG_U and the concentration of Gdn-HCl for the protein samples indicated. From the denaturation curve, the transitional region is used to calculate the free energy change ΔG_U between folded and unfolded conformations using the equation $\Delta G_U = \Delta G_U^0 - m[\text{Gdn-HCl}]$. ΔG_U^0 is given as ΔG_U in the absence of denaturant.

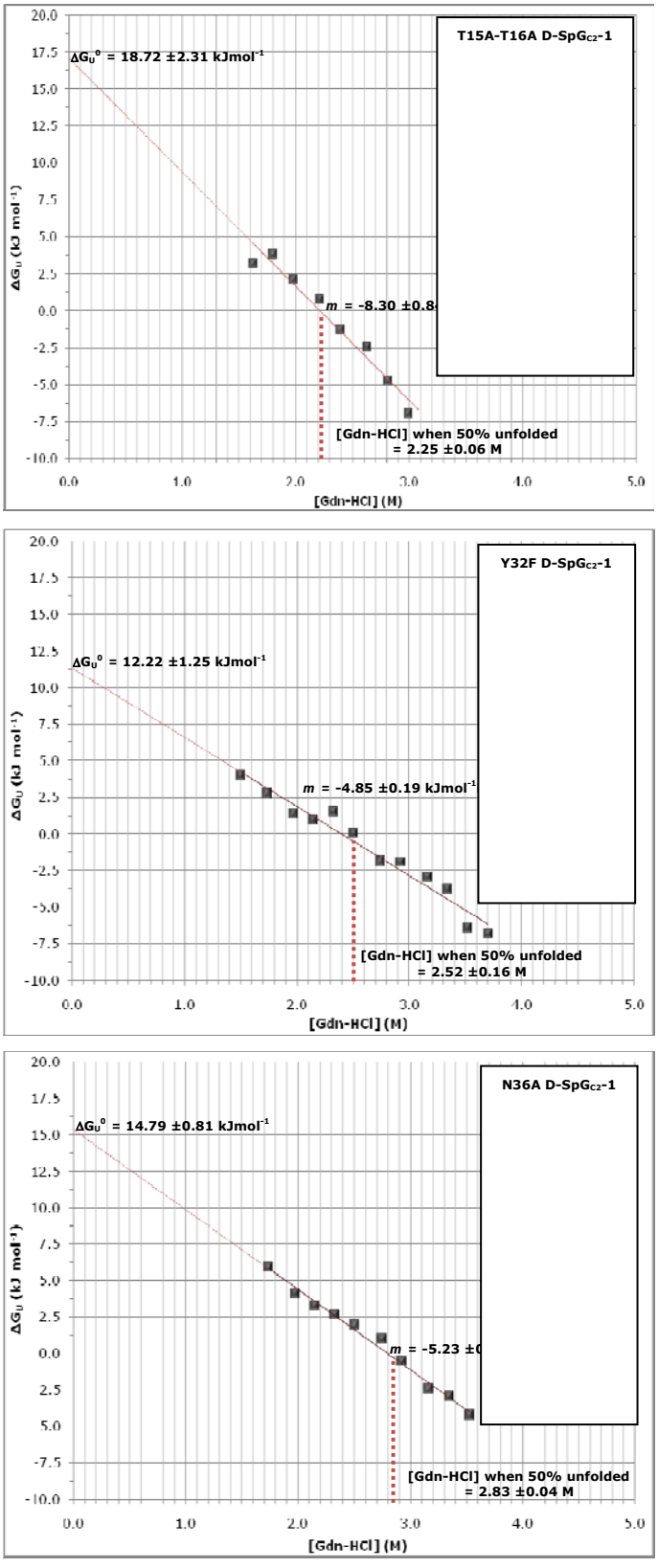
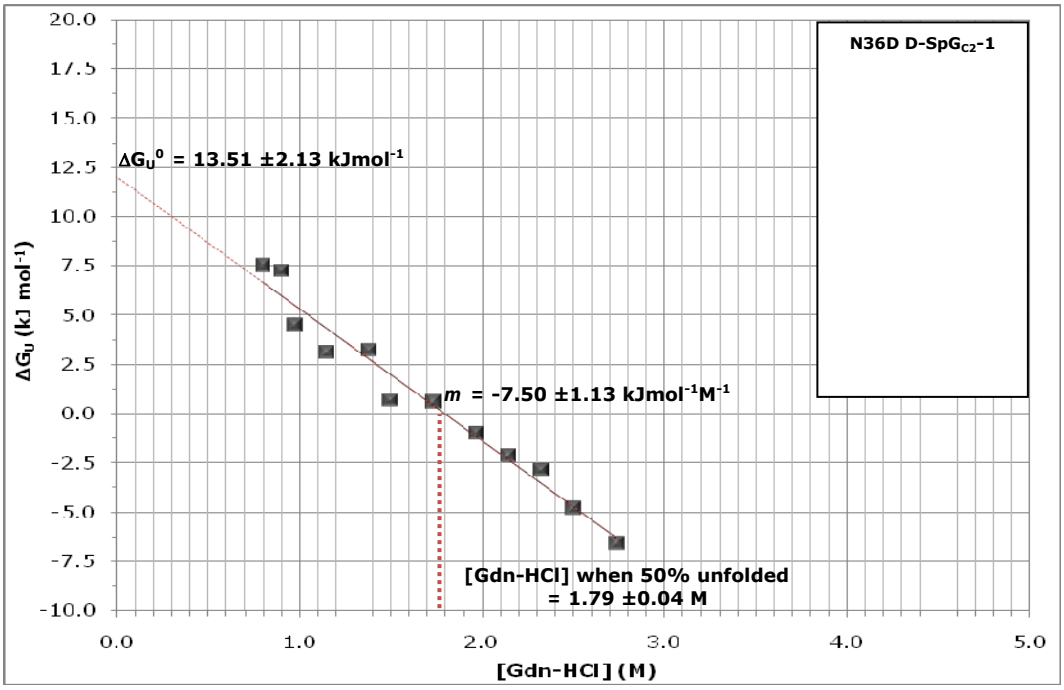


Figure 4.28 – A typical linear relationship between ΔG_U and [Gdn-HCl] for Fab D-SpG$_{C2}$-1 mutants

The plot shows the linear relationship between ΔG_U and the concentration of Gdn-HCl for the protein samples indicated. From the denaturation curve, the transitional region is used to calculate the free energy change ΔG_U between folded and unfolded conformations using the equation $\Delta G_U = \Delta G_U^0 - m[\text{Gdn-HCl}]$. ΔG_U^0 is given as ΔG_U in the absence of denaturant.



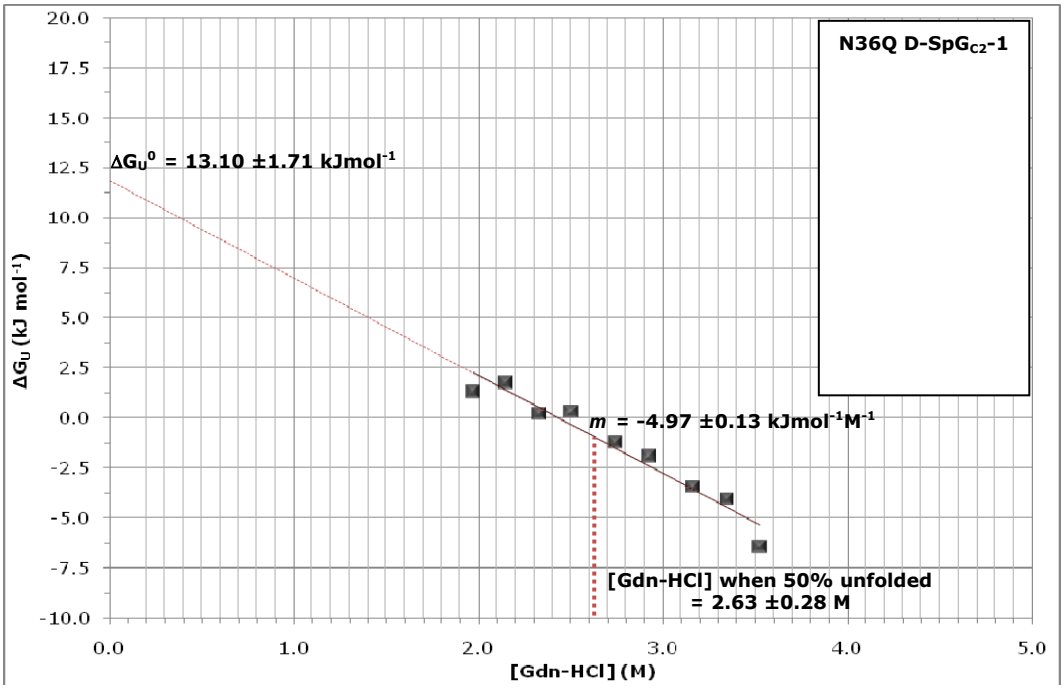


Figure 4.29 – A typical linear relationship between ΔG_U and [Gdn-HCl] for Fab D-SpG$_{C2}$-1 mutants

The plot shows the linear relationship between ΔG_U and the concentration of Gdn-HCl for the protein samples indicated. From the denaturation curve, the transitional region is used to calculate the free energy change ΔG_U between folded and unfolded conformations using the equation $\Delta G_U = \Delta G_U^0 - m[\text{Gdn-HCl}]$. ΔG_U^0 is given as ΔG_U in the absence of denaturant.

These results demonstrate the effect of each mutation on the parameters of unfolding with tryptophan 14 in the E14W-W42F construct acting as a fluorescent probe for these studies. Comparison of the parameters characterising equilibrium Gdn-HCl unfolding obtained for the majority of the mutant constructs studied, compare very well with values obtained for E14W-W42F D-SpG_{C2}-1.

As can be seen from table 4.2 the mutations T10S, T16A, T16S, Y32F, N36A, N36D and N36Q have little effect on the conformational stability of D-SpG_{C2}-1.

D-SpG_{C2}-1 construct	ΔG_U^0 intercept (kJ.mol^{-1}) ± std dev. n≥2	ΔG_U^0 calculated (kJ.mol^{-1}) ± std dev. n≥2	*m* (kJ.mol^{-1}M^{-1}) ± std dev. n≥2	[Gdn-HCl] at 50% unfolded (M) ± std dev. n≥2
E14W-W42F	14.31 ±1.76	14.28	-5.29 ±0.51	2.70 ±0.07
T10A	19.56 ±1.19	19.62	-8.42 ±0.69	2.33 ±0.05
T10S	14.68 ±7.06	14.37	-5.63 ±2.06	2.55 ±0.32
T15A	17.33 ±0.98	17.34	-6.15 ±0.41	2.82 ±0.03
T15S	16.77 ±1.66	16.77	-6.93±0.51	2.42 ±0.06
T16A	13.23 ±2.88	13.23	-6.24 ±1.27	2.12 ±0.04
T16S	13.06 ±0.60	13.06	-5.33 ±0.22	2.45 ±0.01
T15A-T16A	18.72 ±2.31	18.68	-8.30 ±0.84	2.25 ±0.06
Y32F	12.22 ±1.25	12.23	-4.85 ±0.19	2.52 ±0.16
N36A	14.79 ±0.81	14.80	-5.23 ±0.36	2.83 ±0.04
N36D	13.51 ±2.13	13.42	-7.50 ±1.13	1.79 ±0.04
N36Q	13.10 ±1.71	13.08	-4.97 ±0.13	2.63 ±0.28

Table 4.2 – A summary table of the free energy values characterising the denaturation of D-SpG_{C2}-1 Fab mutant constructs

T10A D-SpG_{C2}-1, however, has significantly increased the ΔG_U^0 intercept value from 14.31 ±1.76 to 19.56 ±1.19 kJ.mol^{-1} when compared to E14W-W42F D-SpG_{C2}-1. This increase is largely due to an increased *m* value.

T15A D-SpG_{C2}-1 has increased the ΔG_U^0 intercept value from 14.31 ±1.76 to 17.33 ±0.98 kJ.mol^{-1} when compared to E14W-W42F D-SpG_{C2}-1, but not as significantly as its neighbouring T10A mutant. T15S D-SpG_{C2}-1 has also increased

the ΔG_U^0 intercept value from 14.31 ±1.76 to 16.77 ±1.66 kJ.mol^{-1} when compared to E14W-W42F D-SpG$_{C2}$-1, but not as significantly as its neighbouring T10A and T15A mutants. Again, this increase is due to an increased *m* value.

Double mutant T15A-T16A D-SpG$_{C2}$-1 also shows an effect on the conformational stability of SpG, with an increased ΔG_U^0 intercept value from 14.31 ±1.76 to 18.72 ±1.38 kJ.mol^{-1} when compared to E14W-W42F D-SpG$_{C2}$-1. This increase is due to an increased *m* value.

These data by themselves cannot be used to deduce detailed facts about the gross protein folding. However, together with binding data presented in chapter 5.0, it may be possible to draw more conclusions. However, in almost all cases the ΔG_U^0 value suggests the domains are well folded.

4.7 – Discussion

In this chapter fluorescence was employed to study the conformational stability and the unfolding of D-SpG$_{C2}$-1 and its mutant constructs designed to probe the interaction with Fc or Fab, respectively. Tryptophan 42 and the introduction of a tryptophan residue at position 14 in D-SpG$_{C2}$-1 has allowed these residues to be exploited as reporter groups for exclusively monitoring stability changes in the mutant constructs.

For the mutations made to explore Fc binding, namely E26A, Q31E, N34A, N34D, E41Q, W42F and W42Y, only the N34 substitution led to major consequences on protein stability. It may be significant that these mutations also led to changes in binding characteristics, see table 5.5 and 6.1.

For the mutations designed to explore Fab binding interactions none led to significantly reduced stabilities, indeed mutants with decreased affinities for Fab had, if anything, slightly higher conformational stability than wild-type.

CHAPTER 5.0

EQUILIBRIUM AND PRE-EQUILIBRIUM BINDING STUDIES OF SPG-FC BINDING INTERACTION

5.1 – Introduction

Specific interactions between proteins play a crucial role in a variety of biological processes and many activities depend on the ability of proteins to form complexes with other proteins to function. Proteins must interact with very high specificity and the recognition process is usually intricate and involves formation of many specific intermolecular contacts. Knowledge of the stability of complexes and the arrangement of proteins in a complex, together with the identification of specific amino acid residues involved in these interactions, is fundamental in understanding the regulatory, kinetic and recognition properties of proteins. Protein–protein interactions are a result of optimisation of the binding affinity with the generation of specificity. Thermodynamic, kinetic and structural studies of protein–protein interactions have shown how proteins interact rapidly, tightly and in a specific manner [Perozzo *et al*., 2004].

There are many factors that control the rate of chemical reactions. These include the size and charge of ligands, the concentration of the reactants and environmental conditions, such as pH and temperature. There are a number of approaches that can be used in the study of protein-protein binding interactions for measuring equilibrium and pre-equilibrium binding constants. These include fluorescence titrations, stopped-flow fluorescence and Isothermal Calorimetry (ITC), and have all been proven to be powerful techniques in measuring equilibrium constants [Holdgate and Ward, 2005]. One method of monitoring a reaction is through spectroscopy. If the reactant or product exhibits a change in spectral properties as a function of reaction time, this method can usually be employed. A conventional cell is adequate for reactions with half-lives greater than a couple of minutes, however, reactions that are over within a second or milliseconds need specialised equipment. Of all the spectroscopic techniques that are available to study protein-protein interactions, fluorescence probably offers the greatest versatility of applications, most likely because of its inherent sensitivity and simplicity of use.

5.2 - Mechanism of equilibrium

For any reaction to be at equilibrium, the rates of the forward and reverse reactions have to be equal. Reactions that are non-reversible are said to go to completion, and after the reaction occurs the reaction mixture contains only products and any left over excess reactants. The products do not break down to reform reactants.

At equilibrium, the equilibrium association constant (K_a) equals the product of the concentrations of product divided by the concentrations of reactants.

In the reaction such as that described below, a protein (P) reversibly binds a ligand (L) to form a complex (P•L).

$$\mathbf{P + L} \underset{k_{diss}\ (s^{-1})}{\overset{k_{ass}\ (M^{-1}s^{-1})}{\rightleftharpoons}} \mathbf{P{\bullet}L} \qquad \textit{(equation 5.1)}$$

The forward velocity of the reaction $\mathbf{vf = k_{ass}\ [P][L]}$ and the reverse velocity $\mathbf{Vr = k_{diss}\ [P.L]}$

At equilibrium **vf = vr** and therefore, $\mathbf{k_{ass}\ [P][L] = k_{diss}\ [P{\bullet}L]}$. Hence,

$$\mathbf{k_{diss}/k_{ass} = [P][L]/[P{\bullet}L]} \qquad \textit{(equation 5.2)}$$

Similarly,

$$\mathbf{k_{ass}\ /k_{diss} = [P{\bullet}L]/[P][L] = Ka} \qquad \textit{(equation 5.3)}$$

$$\mathbf{Kd = [P][L]/[P{\bullet}L]} \qquad \textit{(equation 5.4)}$$

Thus, Kd may be determined by measurements of the association and dissociation rates of reaction or from the equilibrium concentrations of the reactants and products.

Equilibrium constants are determined in order to quantify chemical equilibria and describe the strength of binding or affinity between proteins and their ligands. If the Ka is large, products of the reaction are favoured and the reaction lies towards complex formation, whereas if the Ka is small, dissociation of the complex is favoured.

5.2.1 – Association and dissociation rate constants

Kinetic binding experiments can be used to determine the association and dissociation rate constants of molecular interactions and thus allow a separate estimate of equilibrium constants and maybe shed light on the reaction pathways of protein-protein interactions [Holdgate and Ward, 2005]. The velocity of complex formation is dependent upon the concentrations of each of the reactants (P and L) and is therefore a second order process with a rate constant (k_{ass}) with units of $M^{-1}\ s^{-1}$. The velocity of the dissociation process (k_{diss}) is proportional to

the concentration of the complex (P•L) and has a first order rate with units of s^{-1}, as shown in equation 5.1.

5.3 - Binding interactions of D-SpG$_{C2}$-1

Association and dissociation of complexes are distinct events and are influenced by non-covalent intermolecular interactions with different forces contributing to either. The rate of association is determined by long-range electrostatic forces, whereas the rate of dissociation is determined by short range interactions and the co-operativity between them [Schreiber and Selzer, 1999]. Therefore, characterisation of reversible interactions in terms of binding, stoichiometry, kinetics and thermodynamics enables the examination of these parameters.

Close examination of the X-ray crystallographic data of [Achari *et al.*, 1992; Derrick and Wigley, 1992; Gallagher *et al.*, 1994; Derrick *et al.*, 1994; Sauer-Eriksson *et al.*, 1995] has been undertaken in order to identify specific amino acid residues that may be critical to this binding interaction. Those residues so identified have then been substituted in a program of site directed mutagenesis and the mutated proteins produced. The effect of such mutations on the kinetic, equilibrium and thermodynamic parameters of the binding reaction have been investigated using fluorescence spectroscopy employing both equilibrium studies, which will provide information on the binding affinity and number of binding sites, and pre-equilibrium studies, which will determine rates at which equilibrium is reached and may identify structural changes. Studies were also carried out using isothermal titration Calorimetry (ITC) to study the thermodynamics of complex formation and further will allow information on the contribution of a specific amino acid to the biological function of SpG to be evaluated.

5.4 – Principles of fluorescence spectroscopy

Fluorescence is the result of a three-stage process that occurs in certain molecules, such as polyaromatic hydrocarbons, called fluorophores. When these molecules absorb photons of light and are excited to a higher electronic state, the absorbed energy can be released by the emission of a photon of light. This is luminescence, and when this release of energy is virtually immediate this is called fluorescence, thus it is this light energy, which is studied during fluorescence spectroscopy.

Stage one of the fluorescence process involves the absorption of a photon of the incident radiation (excitation) by a fluorophore, which gives rise to transitions

from the ground state (S_0) to an excited electronic state as governed by the laws of quantum mechanics, illustrated in figure 5.1. Stage two is the excited-state lifetime, which exists for a finite time typically 1 to 10 nanoseconds. Within this time, the fluorophore may be affected by conformational changes and is subject to a multitude of possible interactions within its molecular environment. As the electron returns to ground state, it may arrive at one of the vibrational levels of S_0, S_1 or S_2, consequently losing some energy in the form of heat. Within a large population of molecules, the probability of dropping from the excited state to any vibrational level in the ground state is what determines the shape of the fluorescence spectrum, from where fluorescence emission occurs, stage three. This means that the emission spectrum of the fluorophore is dependent upon these chemical and physical properties of the environment in which it has been placed. It is this phenomenon that establishes fluorescence as an important biochemical technique.

In the majority of molecules, the vibrational levels overlap with those of the first excited state, resulting in all the absorbed energy being lost in a non-radiative manner, which is heat. However, not all the molecules initially excited by absorption, in stage one, return to the ground state S_0 by fluorescence emission. Other processes such as collisional quenching and Fluorescence Resonance Energy Transfer (FRET) may also remove molecules from the S_1 state. The fluorescence quantum yield, which is the ratio of the number of fluorescence photons emitted to the number of photons absorbed, is a measure of the relative extent to which these processes occur.

Excess energy may be absorbed which allows an electron within the molecule to attain a higher electronic vibrational energy. This extra vibrational energy is quickly dispersed in the form of heat, and as a result the electron drops to the lowest vibrational level of the excited state (S_1). The excited electron then returns to the ground state either by emitting light or by a non-radiative transition. Due to energy being lost during this process, emitted light will have a lower energy and thus, will have a longer wavelength.

$$E = h\nu \qquad \textit{(equation 5.5)}$$

Where, E is the energy, h is Planck's constant and ν is frequency, which is equal to $1/\lambda$ wavelength.

Hence,

$E = h/\lambda$ *(equation 5.6)*

As h is a constant, the larger the λ, the lower the energy of the photon.

This difference between the excitation and emission wavelength maxima is termed the "Stokes shift". The Stokes shift represents the energy lost whilst the molecule was in the excited state. The Stokes shift is fundamental to the sensitivity of fluorescence techniques because it allows the emitted fluorescence photons to be easily distinguished from the excitation photons, and can be detected against a low background in fluorescent studies.

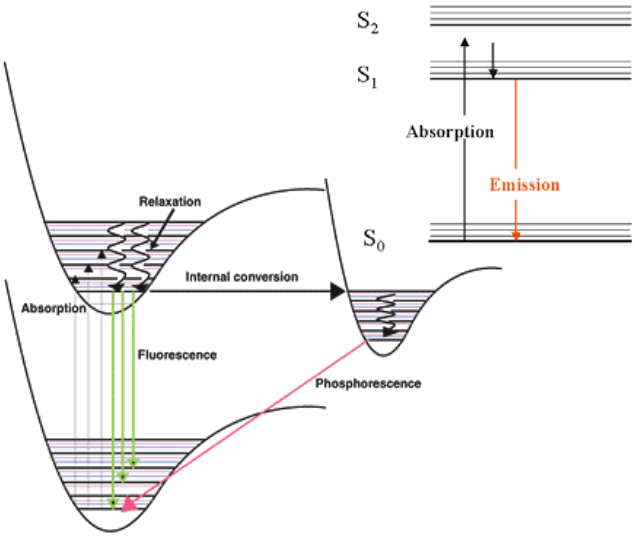


Figure 5.1 – <u>An schematic illustration of the fluorescence process</u>

Fluorescence is light emitted by molecules returning to their ground state after being initially excited by light of a specific wavelength [Lakowicz, 1999]. This diagram illustrates the electronic basis of the phenomenon of fluorescence, as first suggested by Russian scientist Professor Alexander Jabloński. Fluorescence is the phenomenon in which absorption of light of a given wavelength by a fluorescent molecule is followed by the emission of light at longer wavelengths. Excitation of a molecule leads to electrons attaining a higher electronic and vibrational energy status (S_1). On returning to ground state (S_0), an electron may arrive at one of the vibrational levels of the S_0 state, consequently losing its excess energy both radiatively and non-radiatively [Lakowicz, 1999].

Fluorescence detection has three major advantages over other light-based techniques; high sensitivity, speed and safety. Sensitivity is an important issue because the fluorescence signal is proportional to the concentration of the substance being investigated. Since this is a non-invasive technique, fluorescence does not interfere with the sample, in that they are usually not affected or destroyed in the process and no hazardous by-products are produced. The excitation light levels required to generate a fluorescence signal are low, reducing the effects of photo-bleaching of the sample.

5.4.1 – Fluorescence properties of proteins

The fluorescence of protein molecules originates from three aromatic amino acids, tryptophan, tyrosine and phenylalanine (see figure 5.2), and the fluorescence spectrum of a protein is composed of contributions from these residues. Although other factors, such as cofactors, may also contribute.

Generally, aromatic ring systems are planar, and electrons are shared over the whole ring structure [Weber and Teale, 1957]. Molecules require a rigid aromatic ring or ring system to keep the vibrational energy (and its loss) to a minimum. The fluorescence of a protein is normally dominated by the contribution from the tryptophan residues as its Molar Extinction Coefficient of excitation light is considerably higher than that of tyrosine and phenylalanine, although its quantum yield of emission is approximately the same as that of tyrosine, as shown in table 5.1.

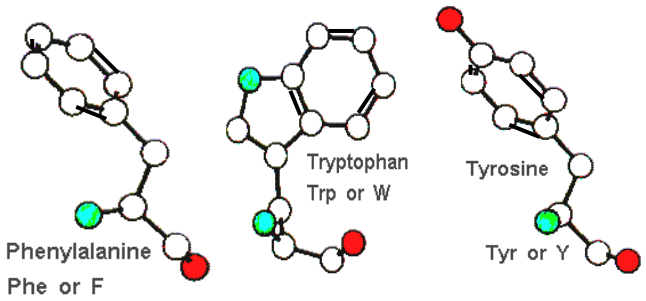


Figure 5.2 – <u>The aromatic amino acids</u>

The aromatic amino acids, shown above, exhibit ultraviolet radiation absorption properties with large extinction coefficients.

Amino Acid	φ_F	ε ($M^{-1}cm^{-1}$)	λ (nm)
Tryptophan	0.13	5600	280
Tyrosine	0.14	1450	275
Phenylalanine	0.024	300	260

Table 5.1 – <u>Summary of the fluorescent properties, quantum yield, molar extinction coefficient and the wavelength, of the aromatic amino acids</u>

[Gore, 2000]

When the tryptophan residue becomes buried within a hydrophobic environment, such as within a protein molecule, spectral changes occur. The quantum yield increases and a shift to a shorter wavelength in the emission maximum also occur. A value of 320 nm to 330 nm for an emission maximum is indicative of a very hydrophobic environment, and that the tryptophans are deeply buried within

the protein. Conversely, if the emission maximum of a protein is nearer 348 nm, one can presume that the tryptophan residue(s) are on the surface of the protein exposed to solvent [Kuipers and Gruppen, 2007].

5.4.2 – Emission spectra

Fluorescence spectroscopy can be applied as a non-quantitative method at an early stage in protein equilibrium studies. Emission wavelength scans are performed with the purified protein in the absence and presence of its ligand, using a fixed excitation wavelength and variable emission wavelength. The binding of the ligand may often be visualised as a change in the spectrum of the protein and shifts in emission maxima are the most indicative of actual binding, as well as increases or decreases of the peak fluorescence intensity.

5.5 - Determining emission spectra of D-SpG$_{C2}$-1 and its complex with Fc

In order to carry out binding studies between a domain of SpG and Fc using fluorescence, a measureable change in fluorescence must occur upon complex formation. To establish whether or not a spectral change does occur upon binding, the emission spectra for Wt D-SpG$_{C2}$-1 and Fc, both in free solution, and in complex were determined, and are presented in figure 5.3.

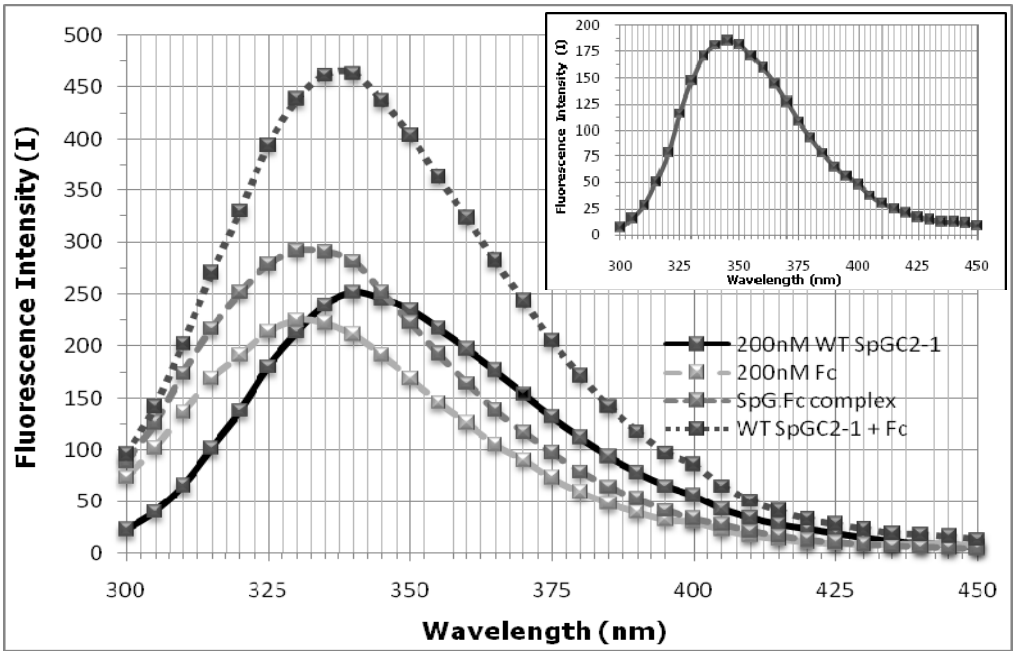


Figure 5.3 - Emission spectra for Wt D-SpG$_{C2}$-1 and Fc in complex

The figure shows the emission spectra obtained for 200 nM Wt D-SpG$_{C2}$-1, 200 nM Fc and the SpG•Fc complex. Protein samples were excited at 280 nm and fluorescence emissions were measured between 300 and 450 nm. All spectra were measured in triplicate and corrected for the buffer fluorescence. The inset shows the fluorescence change upon complex formation, calculated by subtracting the observed fluorescence intensity of the complex from the theoretical fluorescence intensity of the complex.

The spectra presented show that fluorescence intensity obtained for wild-type D-SpG_{C2}-1 displays an emission maximum at 340 nm, whereas the peaks of the Fc spectra and the complex spectra are both at 330 nm. The spectra also show that the fluorescence intensity obtained for the SpG•Fc in complex is significantly less than the theoretical fluorescence of the complex being calculated from the sum of the fluorescence spectra of the two proteins in free solution. This quench in fluorescence occurs upon complex formation and shows that the formation of this SpG•Fc complex can be monitored fluorimetrically. The quench is represented as a fluorescence difference spectrum, presented in the inset figure in figure 5.3, and shows that the difference spectrum has a maximum at 343 nm, suggesting a tryptophan residue in a solvent exposed position is being quenched on complex formation.

This fluorescence quenching effect seen on complex formation may be attributable to the tryptophan of SpG, thus providing a way of following binding and further implicates the involvement of this residue in the interaction. This has an emission wavelength maximum of 340 nm, which is approximately the same as the position of the peak of the difference spectrum in agreement with observations of others on a similar system [Sloan and Hellinga, 1998].

The involvement of this tryptophan residue of SpG in binding Fc is further supported by NMR studies [Gronenborn and Clore, 1993]. Shifts in SpG resonances were observed, which have to have arisen either directly or indirectly from contacts with the Fc fragment. No shifts at all arose from regions of SpG that were not in contact with Fc, thus it can be said that these regions are unaffected by Fc binding. But the most substantial shifts observed was for the tryptophan residue, W42, upon complexation [Gronenborn and Clore, 1993], clearly indicating the involvement of this side-chain in complex formation.

5.6 – Equilibrium binding studies using fluorescence titrations

Equilibrium binding studies can provide valuable information on the binding affinity and number of binding sites of the interaction between D-SpG$_{C2}$-1 and Fc. In the following sections, studies are described that were aimed at the determination of the stoichiometry and Kd for the binding reaction between D-SpG$_{C2}$-1 and Fc.

5.6.1 – Determining the stoichiometry and Kd of D-SpG$_{C2}$-1 and Fc complex

The binding interaction between D-SpG$_{C2}$-1 and Fc has been studied through fluorescence titrations using the change in fluorescence that occurs upon complex formation. Fluorescence emissions at 340 nm were measured, using excitation wavelengths of 280nm to study total fluorescence, and 295 nm to study tryptophan only fluorescence. 5 µl samples of 100 µM D-SpG$_{C2}$-1 were titrated into 2 ml samples of 1 µM Fc. Data obtained were corrected for dilution and the inner filter effect, and further analysed through a Klotz plot, employing the following equation:

$$\frac{Kd}{(1-\alpha)} = \frac{[L]_0}{\alpha} - [E]_0 \qquad \textit{(Equation 5.7)}$$

Where α is the fractional saturation of SpG binding sites on Fc, $[L]_0$ is the total concentration of SpG and $[E]_0$ is the total number of binding sites on Fc for SpG. A plot of $1/(1-\alpha)$ against $[L]_0/\alpha$ will have a gradient of 1/Kd and an x-axis intercept of $[E]_0$.

Figure 5.4 presents a typical saturation curve for the binding of Wt D-SpG$_{C2}$-1 to Fc and the subsequent analysis through a modified Klotz plot. All data obtained for Wt D-SpG$_{C2}$-1 characterised by fluorescence titrations are summarised in table 5.2.

The Kd value for Wt D-SpG$_{C2}$-1 binding to Fc at 15°C, pH 8.0 was able to be described by a single value of 506 $\pm$ 0.12 nM. The stoichiometry of binding from the intercept $[E]_0$ was 2.13 $\pm$ 0.36, confirming that Fc has two identical binding sites for Wt D-SpG$_{C2}$-1.

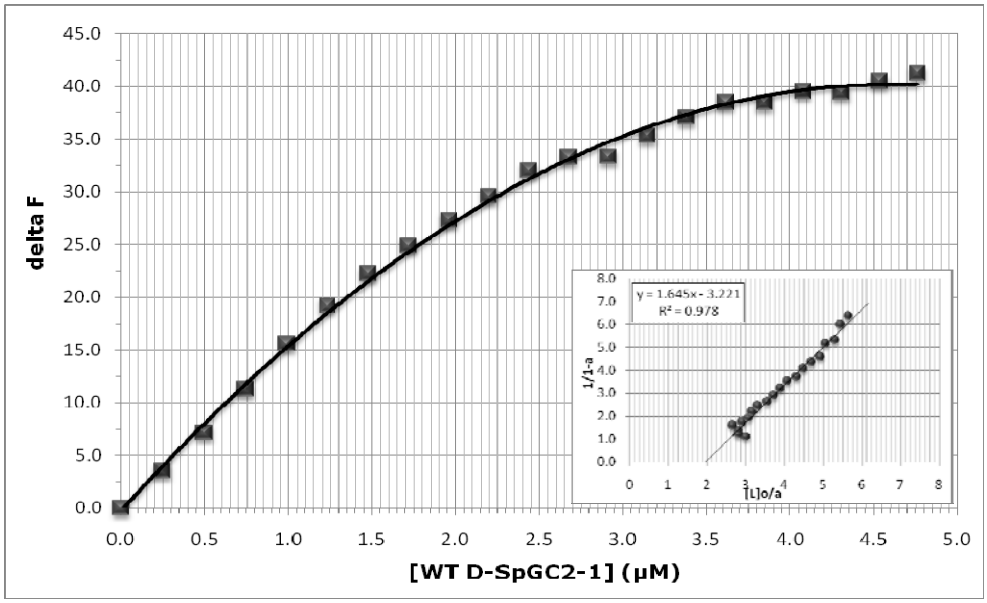


Figure 5.4 - Saturation curve for the titration of Wt D-SpG$_{C2}$-1 into Fc

The figure shows the saturation curve obtained when samples of 100 μM Wt D-SpG$_{C2}$-1 were titrated into a 2 ml solution of 1 μM Fc in 20 mM potassium phosphate buffer, pH 8.0 at 15°C. Each point on the titration curve was calculated by subtracting fluorescence values observed upon titration of Wt D-SpG$_{C2}$-1 into Fc, from the corresponding values for the titration of Wt D-SpG$_{C2}$-1 into buffer. The inset presents a Klotz plot, by which the titration curves were analysed. From the Klotz plot the Kd and stoichiometry of binding were determined. All titrations were carried out in duplicate.

TEMPERATURE (˚C)	pH 6.0 Kd (nM) (*n*) ± std dev.	pH 8.0 Kd (nM) (*n*) ± std dev.
15	173 ± 0.05 *(n = 1.7 ± 0.2)*	506 ± 0.12 *(n = 2.13 ± 0.36)*
25	250 ± 0.22 *(n = 2.01 ± 0.06)*	1270 ± 0.52 *(n = 2.52 ± 0.81)*
40	ND	4400 ± 0.13 *(n = 1.36 ± 0.49)*

Table 5.2 – Summary of dissociation constants (Kd) and the stoichiometry (n) determined by titrating Wt D-SpG$_{C2}$-1 into Fc at varying temperatures and pH

ND denotes not determined.

Previous data obtained to determine the equilibrium dissociation constant for the interaction of D-SpG$_{C2}$-1 with Fc are similar to data obtained in these studies. The Kd of the interaction between Wt D-SpG$_{C2}$-1 and Fc has been measured through fluorescence titrations carried out at 25˚C, pH 6.0 and by competitive ELISA, and was found to have values of 306 ± 65 nM and 310 nM, respectively [Walker *et al*.,

1995]. Similar values were also obtained by Gallagher [Gallagher, 1994] for the B1 (C1) domain of SpG using titration calorimetry, which gave a Kd value of 333 nM. Data obtained here under the same conditions, using the Wt D-SpG$_{C2}$-1 pQE-30 construct gives a Kd value of 250 ± 0.22 nM, as presented in table 5.1. This compares well with the previous data, but could indicate that the introduction of a 6x His-tag at the N-terminal end of the protein has slightly increased the binding affinity of D-SpG$_{C2}$-1 for Fc.

However, Walker [1994] also performed fluorescence titrations at 25°C at a higher pH (pH 8.0) and obtained a Kd value of 693 ± 100 nM. This does not compare as well with data obtained in these studies, which yielded a Kd value of 1270 ± 0.52 nM at the same pH. These results may indicate a difference in the binding affinity of D-SpG$_{C2}$-1 for Fc due to the different samples of the Fc and D-SpG$_{C2}$-1 domains.

5.7 – Pre-equilibrium binding studies

Biochemical reactions, such as those involving protein-ligand or protein-protein interactions are generally very fast with equilibria being reached within 50 to 100 milliseconds (ms) [Gore and Bottomley, 2000]. The most common method of following these reactions is by absorbance or fluorescence spectrometry, as described in section 5.4. However, conventional spectroscopic techniques cannot be used when investigating reactions that occur at a sub-second rate, especially for measuring the rates of fluorescence changes upon complex formation. If reactants are added manually, and then stirred for a few seconds to allow adequate mixing, the reaction is over and no changes in spectra are recorded. Therefore, pre-equilibrium binding studies using stopped-flow fluorescence spectroscopy is one of the most frequently used rapid kinetics techniques. Stopped-flow fluorescence provides virtually instantaneous mixing, and can be used to determine the rate at which equilibria are obtained, as well as identifying potential structural changes and observing complex formation.

5.7.1 – Stopped-Flow fluorescence

Stopped-flow fluorescence allows the observation of reaction in the millisecond to second time scale before equilibrium is reached. This provides rate constants for each individual step and allows fluorescence profiles to be measured from approximately 1 millisecond after the mixing of reagents. The key components of an Applied Photo-Physics stopped-flow fluorescence spectrophotometer platform are shown in figure 5.5.

Small volumes of protein and ligand solutions are driven from syringes under high pressure through a high efficiency mixer. The resultant mixture passes through a measurement flow cell and into a stopping syringe, and just prior to stopping, a steady state flow is achieved. The solution entering the flow cell is only approximately a millisecond old. The age of this reaction mixture is also known as the dead time of the stopped-flow system. As the solution fills the stopping syringe, the plunger hits a block, causing the flow to be stopped instantaneously. Using the manufacturer's data acquisition and analysis software the kinetics of the reaction can be measured in the cell.

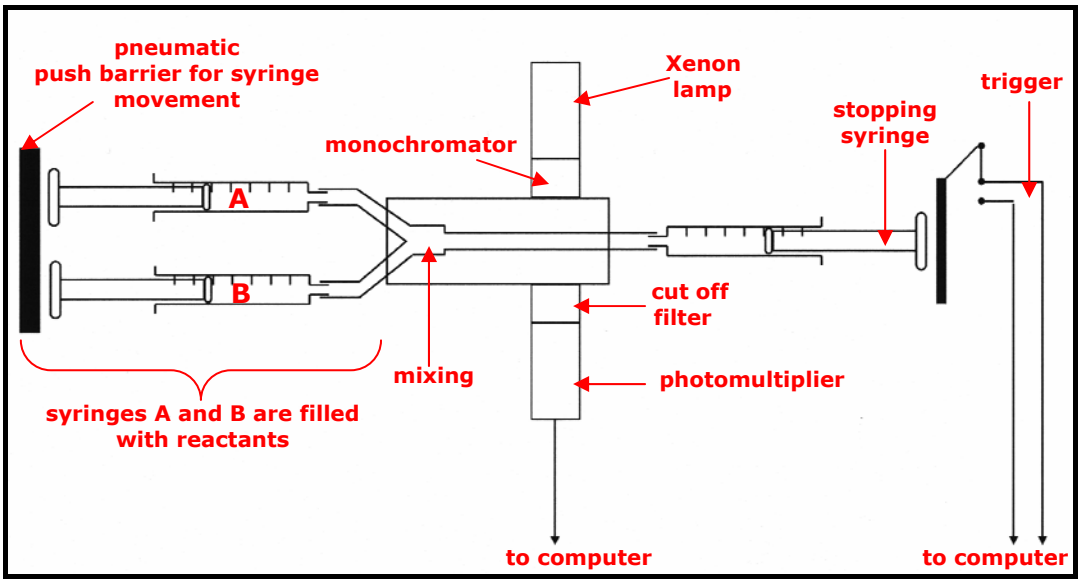


Figure 5.5 – <u>An adapted schematic diagram of the components of a stopped-flow fluorescence spectrophotometer</u>

The stopped-flow fluorescence spectrophotometer is a simple device designed to mix two reagents and to transfer the mixture to an optical cell as rapidly as possible. The reactants are placed in two syringes, and upon pressing the plunger, the solutions travel separately and only mix upon entering the cell. The reactants then enter a further syringe, the stop syringe, where the solutions stop flowing into the cell. Data collection then begins immediately, and the kinetics of the reaction is measured in the cell.

5.7.2 – Measurement of D-SpG$_{C2}$-1 and Fc complex

In the formation of the SpG•Fc complex, the Kd is described by the ratio of the rates of association, k_{ass}, and dissociation, k_{diss}, by the relationship Kd = k_{diss}/k_{ass} [Walker *et al.*, 1995]. This can be seen below in equation 5.8, followed by equation 5.9.

$$\mathbf{SpG + Fc} \underset{k_{diss}}{\overset{k_{ass}}{\rightleftharpoons}} \mathbf{SpG{\bullet}Fc} \qquad \textit{(equation 5.8)}$$

$$\frac{[SpG][Fc]}{[SpG{\bullet}Fc]} = \frac{k_{diss}\ (s^{-1})}{k_{ass}\ (M^{-1}s^{-1})} = Kd\ (M) \qquad \textit{(equation 5.9)}$$

5.7.2.1 – First-order and second-order reactions

The order of a reaction is defined as the sum of the exponents in the rate equation. Therefore, in the interaction between SpG and Fc, the forward reaction obeys a second-order and the reverse reaction obeys a first-order law of kinetics and at equilibrium, the rates of both the forward and reverse reactions are equal. At constant temperature, reaction rates vary with reactant concentration, and the rate of this process is proportional to the frequency with which the reacting molecules simultaneously come together.

In order to analyse second-order reactions more easily, conditions are used in which the reaction can be considered to be a first order process. One reactant is always present in great excess, for example SpG, over the ligand Fc concentration, which remains virtually constant throughout the complex formation. This allows an apparent first-order rate constant (k_{app}) to be determined from the fluorescence change obtained.

$$k_{app} = k_{ass}\,[Fc] + k_{diss} \qquad \textit{(Equation 5.10)}$$

The value of the rates for the association (k_{ass}) and dissociation (k_{diss}) of the interaction can be calculated by measuring k_{app} over a range of protein concentrations, all of which have SpG present in a minimum of a 10-fold excess over ligand concentration. By plotting k_{app} against the varying SpG concentrations, k_{ass} and k_{diss} can be calculated from the gradient and the intercept of the y-axis respectively, resulting in a derived second-order rate constant.

5.7.3 – Determination of rate constants for the D-SpG$_{C2}$-1 and Fc complex

All solutions of D-SpG$_{C2}$-1 and human Fc was prepared in 20 mM potassium phosphate buffer at pH 8.0. Both solutions were filtered prior to use to minimise any interference from light scattering, and were also thermo-equilibrated at the correct temperature of 15°C. Stopped-flow runs were completed using the maximum rate of data sampling, either set to 1000 data points at lower protein concentrations, and 400 data points at higher protein concentrations to reduce the amount of noise.

Data is presented as reaction progress curves, which were averaged as three sets of triplicate traces, with data being averaged and analysed using a single exponential curve fitting algorithm, using equation 5.9:

$$\mathbf{F_t = F_0 \exp(-k_{app}t) + C} \qquad \textit{(Equation 5.11)}$$

Where F_0 and F_t are the fluorescence intensities at time 0 and time t respectively, k_{app} is the apparent rate constant for the reaction and C is a constant.

5.7.3.1 – Determination of association rate constant k_{ass}

All measurements were made using an excitation wavelength of 280 nm, and the change in fluorescence intensity of the solution after mixing of the reactants was followed at 335 nm using a glass cut-off filter. Protein solutions were mixed at a 1:1 v/v mixing ratio between the ligand, Fc, which was kept at a constant initial concentration of 1 μM, with initial concentrations of D-SpG$_{C2}$-1 proteins varying between 60 μM and 100 μM.

Figure 5.6 shows a typical reaction progression curve for wild-type D-SpG$_{C2}$-1, obtained by using a final concentration of 0.5 μM Fc and a final concentration of 30 μM D-SpG$_{C2}$-1, after mixing had occurred.

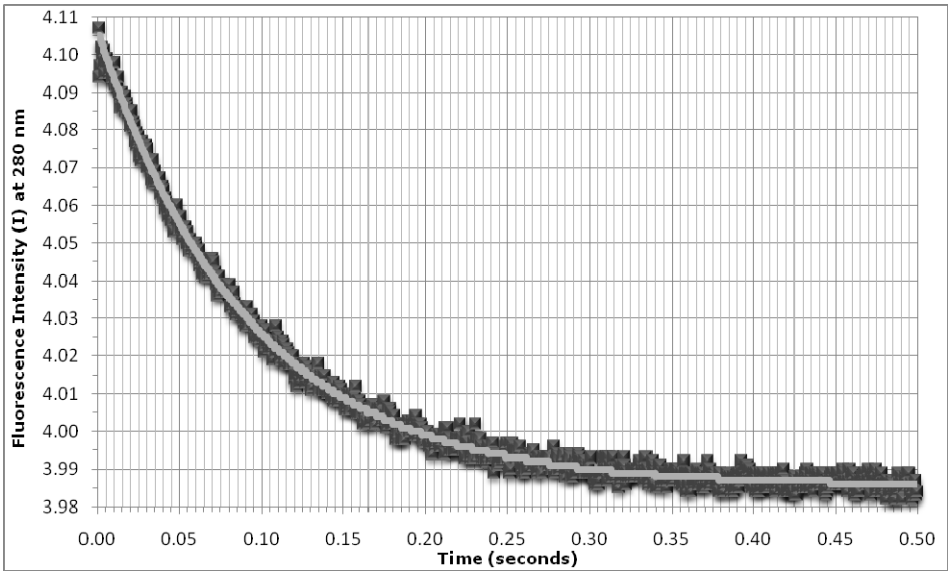


Figure 5.6 – Typical stopped-flow fluorescence profile of the association of Wt D-SpG$_{C2}$-1 and Fc

A reaction progress curve for the association of Wt D-SpG$_{C2}$-1 and Fc is presented. The association curve shows the fluorescence change upon formation of the SpG•Fc complex after mixing 0.5 μM Fc with 30 μM SpG, final concentration, at 15˚C, pH 8.0.

The association progress curve shows a decrease in the fluorescence intensity as the complex between D-SpG$_{C2}$-1 and Fc is formed. By repeating this reaction at various concentration of D-SpG$_{C2}$-1, data were collected for the determination of k_{ass}. Figure 5.7 shows typical reaction progression curves obtained by using a final concentration of 0.5 μM Fc and final concentrations of D-SpG$_{C2}$-1 ranging from 10 μM to 50 μM, after mixing had occurred.

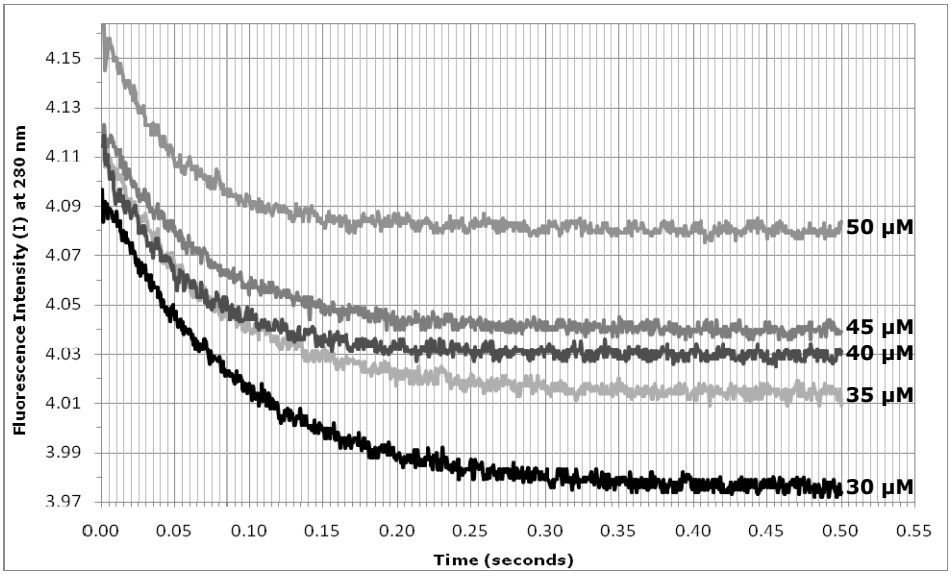


Figure 5.7 – Typical stopped-flow fluorescence profile of the association of varying concentrations of Wt D-SpG$_{C2}$-1 and Fc

Reaction progress curves for the association of Wt D-SpG$_{C2}$-1 and Fc are presented. The association curves show the fluorescence change upon formation of the SpG•Fc complex after mixing 0.5μM Fc with 30-50 μM SpG, final concentration, at 15˚C, pH 8.0.

This figure shows the change in fluorescence intensity with increasing concentration of D-SpG$_{C2}$-1. The apparent rate constants and apparent half-times for the reactions shown are given in table 5.3, and used in the graphical plot in figure 5.9.

[D-SpG$_{C2}$-1] (μM)	k_{app} (s^{-1})	Half-time of reaction (ms)
30	11.15	0.062
35	13.39	0.052
40	14.98	0.046
45	16.96	0.041
50	19.03	0.036

Table 5.3 – Summary of half-time of reaction and the apparent rate constant (k_{app}) at varying concentrations of Wt D-SpG$_{C2}$-1 with Fc

5.7.3.2 – Determination of dissociation rate constant k_{diss}

The dissociation rate constant may also be independently measured by a displacement experiment. In a displacement reaction, a solution containing protein (P) and fluorescent ligand (fL) is mixed with an excess of a non-fluorescent ligand (nfL) the following scheme applies:

$$\mathbf{P{\bullet}fL} \underset{k_{ass}}{\overset{k_{diss}}{\rightleftharpoons}} \mathbf{P} + \mathbf{fL} \overset{\mathbf{nfL}}{\underset{k_{diss}}{\overset{k_{ass}}{\rightleftharpoons}}} \mathbf{P{\bullet}nfL} \qquad \textit{(Equation 5.12)}$$

The rate of dissociation was determined by the 1:1 dilution displacement of Fc from a pre-formed 2:1 ratio (3 μM:1.5 μM) of the SpG•Fc complex with non-fluorescent 20 mM potassium phosphate buffer, pH 8.0.

Figure 5.8 presents a reaction progress curve for the fluorescence change upon dissociation.

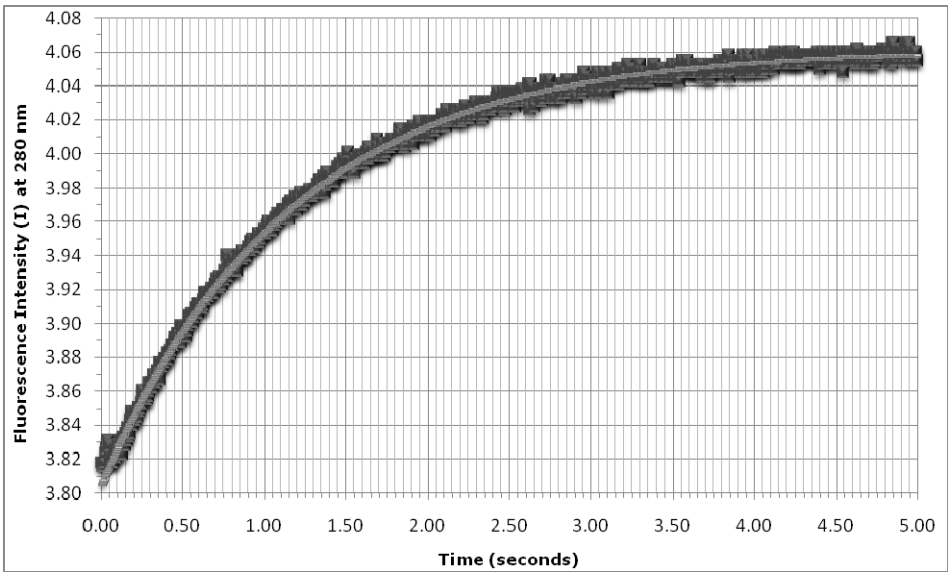


Figure 5.8 – Typical stopped-flow fluorescence profile of the dissociation of Wt D-SpG$_{C2}$-1 and Fc complex

A reaction progress curve for the dissociation of Wt D-SpG$_{C2}$-1 and Fc is presented. The dissociation curve shows the fluorescence change upon a 1:1 dilution of 3 μM:1.5 μM SpG•Fc complex, final concentration at 15˚C, pH 8.0.

The Applied Physics data analysis software allowed curves to be fitted to the data. The reaction progress curves, presented in figures 5.6 and 5.8, fit well to single-exponential equations. This suggests that both the association and dissociation of the D-SpG$_{C2}$-1 and Fc complex occur by simple processes.

5.7.3.3 – Determination of pre-equilibrium constant Kd

A plot of the observed rate constant against protein concentration will give a straight line with the slope giving the second-order association rate constant and the intercept giving the first-order dissociation rate constant. A linear plot of k_{app} (s^{-1}) against final Wt D-SpG$_{C2}$-1 concentration, as shown in figure 5.9, allows k_{ass} and k_{diss} to be determined from the gradient of the line and the intercept of the y-axis respectively. These values then enable the Kd = k_{diss}/k_{ass} to be calculated.

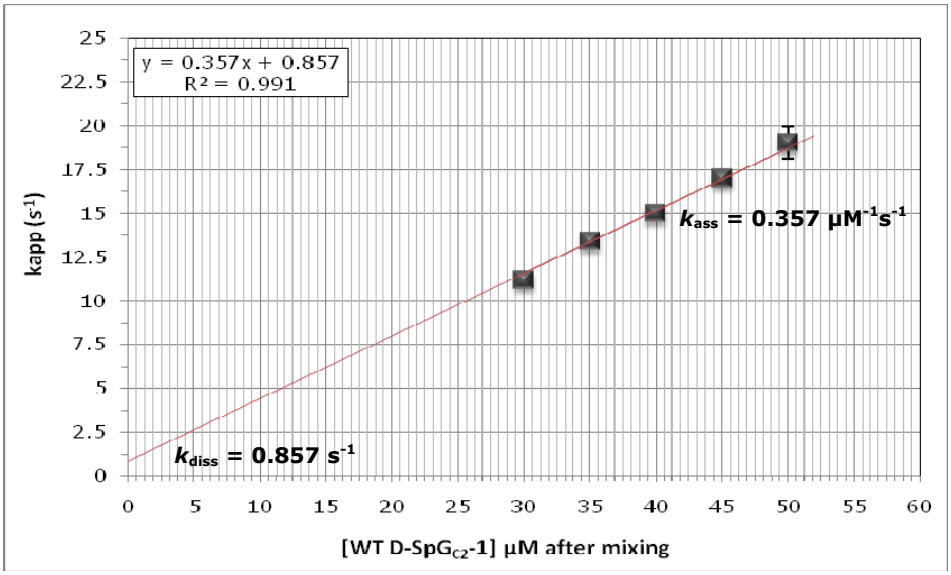


Figure 5.9 – Typical stopped-flow fluorescence analysis of the formation of Wt D-SpG$_{C2}$-1 and Fc complex

This graphical determination represents a linear plot of k_{app} (s^{-1}) against varying concentrations of Wt D-SpG$_{C2}$-1 at 15°C, pH 8.0. The rate constants k_{ass} and k_{diss} from the pseudo-first-order rates (k_{app}) can be calculated for each curve. K_{ass} and k_{diss} are determined from the gradient of the line and the intercept of the y-axis respectively.

A summary of the rate constants obtained for wild-type D-SpG$_{C2}$-1 and Fc complex formation is presented in table 5.4.

k_{ass} ($\mu M^{-1}s^{-1}$) ± std dev. n≥2	k_{diss} (s^{-1}) ± std dev. n≥2	Kd (µM) ± std dev. n≥2
0.33 ± 0.05	0.82 ± 0.08	2.52 ± 0.29

Table 5.4 – Summary of the rate constants and binding affinity obtained for the interaction of Wt D-SpG$_{C2}$-1 and Fc

Previous stopped-flow experiments studying the binding of Fc to Wt SpG$_{C2}$-1 have obtained a Kd value of 2 µM under the same conditions, at 15°C, pH 8.0 [Walker *et al.*, 1995]. This equilibrium constant is very similar to that obtained in these studies, with an average Kd value of 2.52 µM being attained here.

It has been suggested that for the SpG•Fc complex to have a 4-fold higher Kd value of 2.52 μM using pre-equilibrium techniques than that of a Kd value of 600 nM obtained in equilibrium studies, is due to a change in affinity that occurs after the initial complex formation, and is possibly due to a conformation change in either D-SpG_{C2}-1 or the Fc [Walker *et al*., 1995]. However, it is thought that this structural change is relatively small, since the X-ray structure of the backbone of SpG in free solution and in the complex is superimposable. However, small changes in Kd may occur as side chains become more immobilised in the complex. The free energy released on complex formation can be calculated from the equation below.

$$\Delta G = -RT\ln Ka \qquad \textit{(Equation 5.13)}$$

Where R is the gas constant, *T* is the absolute temperature (°Kelvin) and Ka is the equilibrium constant. Hence, if Kd is 2.52 μM, then the Ka for the same reaction is approximately $4x10^5$ M and at 25°C,

$$\Delta G = -2442 \times \ln 4x10^5$$
$$= -31.5 \text{ kJ Mol}^{-1} \qquad \textit{(Equation 5.14)}$$

Similarly, if the Kd for the reaction is 600 nM, then ΔG can also be determined from equation 5.14, to be -34.98 $kJ.mol^{-1}$. This is of the order of the contribution of a single hydrogen bond to the binding interactions. Thus, no large conformational change need take place to cause a difference in Kd of approximately 4-fold.

5.7.4– Effect of mutagenesis on D-SpG_{C2}-1 and Fc complex

As previously discussed, mutant constructs of D-SpG_{C2}-1, which have been shown to be critical for the binding interaction with Fc, have been produced to engineer favourable binding properties, as well as to allow specific amino acid residues and their contribution to the biological function of SpG to be evaluated. To achieve these objectives, mutants were designed to potentially alter the Fc binding affinity from that of the wild-type protein. This strategy is important, as it will give a complete overview of the binding interface.

5.7.5 – Determination of rate constants for Fc D-SpG_{C2}-1 mutants

SpG binds Fc at the hinge region that is situated in the cleft between C_H2 and C_H3 domains of Fc. The eight residues within the α-helix and β-strand 3 of SpG that have been identified to be involved in the interfacial interactions with the loop regions of Fc are Glu26, Lys27, Lys30, Gln31, Asn34, Asp39, Glu41 and Trp42 [Sauer-Eriksson *et al.*, 1995]. Six of these residues, Glu26, Lys27, Gln31, Asn34,

Glu41 and Trp42, formed the basis of this study and were mutated to different side-chains to characterise the interaction. However, Lys27A was not successfully expressed so could not be characterised any further. The rates of the association and dissociation of the complex between D-SpG$_{C2}$-1 constructs and Fc, as well as the pre-equilibrium Kd, are summarised below in table 5.5.

D-SpG$_{C2}$-1 construct	k_{ass} ($\mu M^{-1}s^{-1}$) ± std dev. n≥2	k_{diss} (s^{-1}) ± std dev. n≥2	Kd (μM) ± std dev. n≥2
Wt	0.33 ± 0.05	0.82 ± 0.08	2.52 ± 0.29
E26A	ND	ND	ND
Q31A	0.44 ± 0.001	1.18 ± 0.03	2.68 ± 0.08
Q31E	0.27 ± 0.001	2.56 ± 0.05	9.41 ± 0.15
Q31H	0.43 ± 0.02	1.01 ± 0.05	2.34 ± 0.01
N34A	0.37 ± 0.03	14.38 ± 2.20	38.43 ± 3.42
N34D	0.03 ± 0.003	0.99 ± 0.04	28.75 ± 1.17
E41A	0.38 ± 0.01	0.95 ± 0.02	2.50 ± 0.08
E41Q	0.43 ± 0.04	0.72 ± 0.03	1.69 ± 0.23
W42F	ND	ND	ND
W42Y	ND	ND	ND
C56A	0.56 ± 0.02	1.15 ± 0.01	2.06 ± 0.03

Table 5.5 – Summary of the rate constants and binding affinity obtained for the interaction of D-SpG$_{C2}$-1 mutants and Fc

ND denotes little or no binding determined.

The results obtained for D-SpG$_{C2}$-1 mutants Q31A, Q31H, E41A and E41Q studied by stopped-flow show very little difference in their rates of association and dissociation with Fc, and the Kd values to that of wild-type D-SpG$_{C2}$-1, confirming ITC studies (see chapter 6) that these mutants have not grossly affected the binding of D-SpG$_{C2}$-1 with Fc. Results also show that neither alanine nor glutamine side-chain replacements of glutamate 41 affected Fc binding, indicating that this glutamate residue is not significantly involved in the interaction with Fc. However, six mutant constructs were found to affect the binding affinity of D-SpG$_{C2}$-1 with Fc, namely E26A, Q31E, N34A, N34D, W42F and W42Y. These results again agree with those obtained by equilibrium methods, such as ITC (see chapter 6).

The second glutamate residue implicated in the interaction between SpG and Fc, namely Glu26, is implicated in the major interfacial binding site, and is situated on the C-terminal end of the α-helix. This residue forms hydrogen bonds with backbone nitrogen atoms of Fc residues Ile253 and Ser254, and is part of the hydrogen network that forms in the complex. Results obtained by stopped-flow studies showed that no binding was observed with Fc. This total loss of affinity suggests that this residue plays a significant role in the interaction and is essential to Fc binding.

Characterising amino-acid Gln31 by stopped-flow demonstrated that when this residue is mutated to either an alanine or a histidine residue, Fc binding is not affected. Gln31 is implicated in the major interfacial interaction of SpG with Fc and is located around the middle of the α-helix, forming a hydrogen bond with the oxygen atom of the side-chain of Fc residue Gln438. The mutant Q31A results in the loss of the hydrogen bond, but values obtained for the association and dissociation of the complex, and the binding affinity are very similar to those obtained for wild-type, suggesting that this residue does not play a significant role in the binding of D-SpG$_{C2}$-1 with Fc. This was also seen with the conservative mutant Q31H.

In contrast, replacement by glutamate dramatically affects Fc binding, which is reduced approximately 4-fold, and a typical stopped-flow fluorescence analysis of Q31E is presented in figure 5.10. Either an increase in k_{ass} or a decrease in k_{diss}, or a combination of both may bring about an increase in Kd of the complex. The calculated Kd value of the complex formed between Q31E and Fc is 9.41 ± 0.15 μM, compared to the Kd value of 2.52 ± 0.29 μM for the equivalent complex formed by wild-type. This increase in Kd is due to a 3-fold increase in the dissociation rate, whilst the association rate shows no significant change. This, therefore, implicates the significance of this ionisable group in the binding interface between D-SpG$_{C2}$-1 and Fc, and will be characterised further in pH studies.

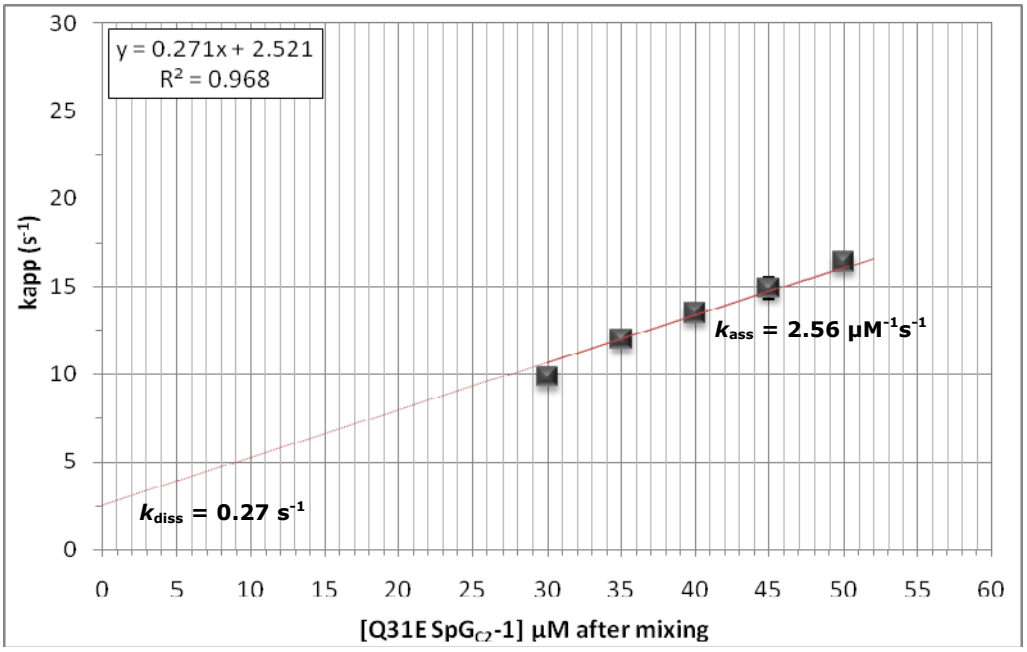


Figure 5.10 – Typical stopped-flow fluorescence analysis of the formation of Q31E D-SpG$_{C2}$-1 and Fc complex

This graphical determination represents a linear plot of k_{app} (s^{-1}) against varying concentrations of Q31E D-SpG$_{C2}$-1 at 15°C, pH 8.0. The rate constants k_{ass} and k_{diss} are determined from the gradient of the line and the intercept of the y-axis respectively.

The exclusive tryptophan residue in the D-SpG$_{C2}$-1 gene, Trp42, has also been implicated in the interfacial interaction with Fc, and is situated on the N-terminal end of β-strand 3. Although this tryptophan residue is not directly involved in binding to Fc through a hydrogen bond network, it was decided to make W42F mutant in order to collaborate the work performed by Walker [Walker, 1994]. Results obtained show that although this mutation is in a slightly different construct, the effect was still the same and a loss of binding was found. This further supports the previous studies, indicating that tryptophan 42 is very important for the binding interaction of SpG with Fc.

This was also apparent for mutant W42Y. Tyrosine is a polar residue by nature and can normally be found at the surface and the interior of proteins. Tyrosine was chosen to substitute the tryptophan residue for its ability to form hydrogen bonds, and a study by Stone *et al* [Stone *et al*., 1989] has shown that tyrosine residues are likely to be involved in SpG's interaction with Fc fragment at the C_H2-C_H3 domain interface. Therefore, it was thought that by introducing a tyrosine residue, a hydrogen bond is formed with the Fc residue Asn434, which may result in the possibility that this interaction may restore Fc binding. However, results obtained by stopped-flow studies showed that no binding was observed with Fc. This result lends itself to two possible interpretations. Either that the loss

of binding is a direct consequence of the loss in affinity, as with W42F, or that this mutant construct can still bind Fc, but no significant fluorescence change is observed as the tryptophan residue that reports changes in fluorescence upon Fc binding has been removed. Walker did not study this mutant, so there is no supporting data available as in the case with W42F to explain these results. However, further characterisation can be achieved by a non-fluorescent technique, ITC, which was used as a thermodynamic approach to study the association of the complex between D-SpG_{C2}-1 and Fc.

Trp42 is also in close proximity with a hydrogen network centralised around asparagine residue 34. Figures 5.11 and 5.12 present stopped-flow results obtained for mutant constructs N34A and N34D, respectively. As mentioned previously, either an increase in k_{diss} or a decrease in k_{ass}, or a combination of both may bring about an increase in Kd of the complex. The calculated Kd value of the complex formed between N34A and Fc is 38.43 ± 3.42 μm, compared to the Kd value of 2.52 ± 0.29 μM for the equivalent complex formed by wild-type. This 15-fold increase in Kd is due to a 17-fold increase in the dissociation rate, whilst the association rate shows no significant change. The calculated Kd value of the complex formed between N34D and Fc is 28.75 ± 1.17 μm, and in contrast, this 11-fold increase in Kd is due to a 11-fold decrease in the association rate, whilst the dissociation rate shows no significant change. Asparagine 34 is implicated in the interfacial interaction of SpG with Fc, and is located at the C-terminal end of the α-helix. This residue's side-chain forms a hydrogen bond with the side-chain of Fc residue Asn434, and continues this vast hydrogen bond network by also forming hydrogen bonds with two other Fc residues, Tyr436 and His433. These results strongly suggest that this residue plays a significant role in SpG's interaction with Fc, and the close proximity of the bulky tryptophan residue may also play a part.

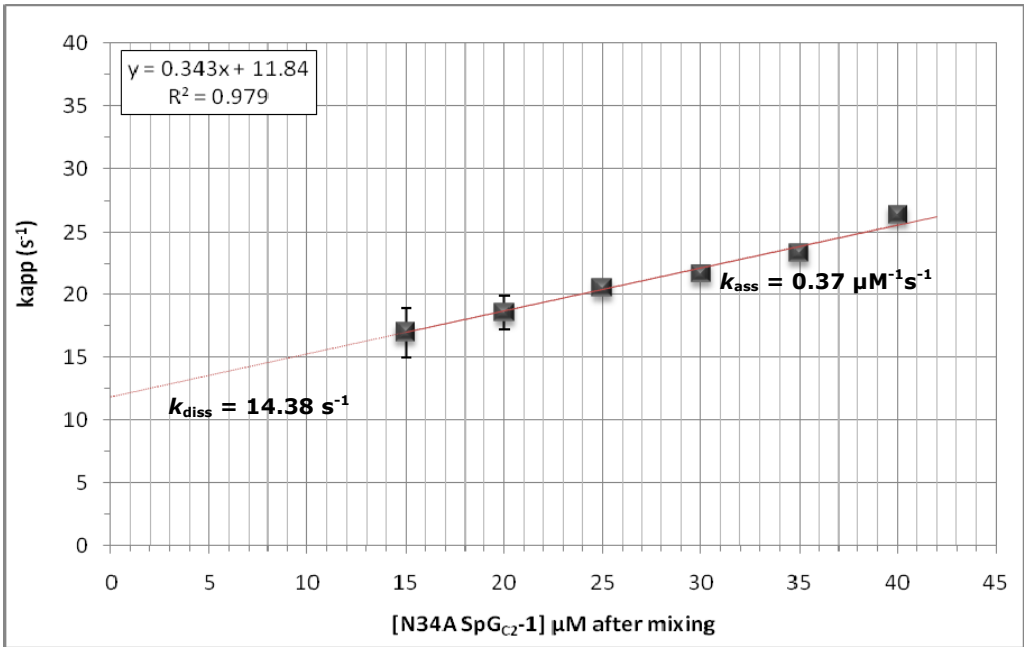


Figure 5.11 – Typical stopped-flow fluorescence analysis of the formation of N34A D-SpG$_{C2}$-1 and Fc complex

This graphical determination represents a linear plot of k_{app} (s^{-1}) against varying concentrations of N34A D-SpG$_{C2}$-1 at 15°C, pH 8.0. The rate constants k_{ass} and k_{diss} are determined from the gradient of the line and the intercept of the y-axis respectively.

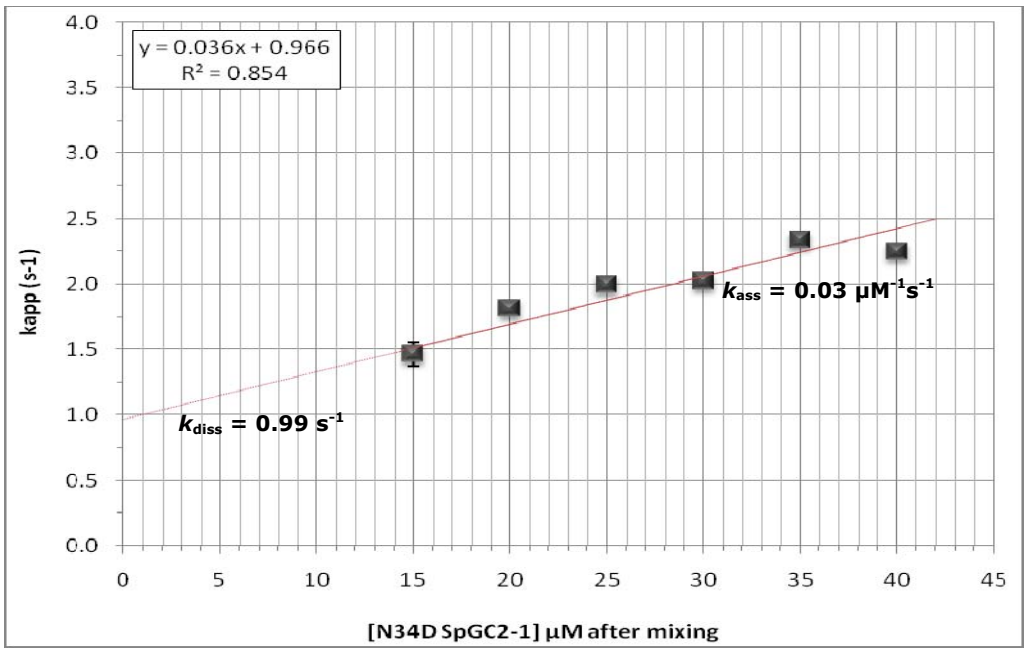


Figure 5.12 – Typical stopped-flow fluorescence analysis of the formation of N34D D-SpG$_{C2}$-1 and Fc complex

This graphical determination represents a linear plot of k_{app} (s^{-1}) against varying concentrations of N34D D-SpG$_{C2}$-1 at 15°C, pH 8.0. The rate constants k_{ass} and k_{diss} are determined from the gradient of the line and the intercept of the y-axis respectively.

5.7.6 – Observing the effect of temperature on D-SpG_{C2}-1 and Fc complex

Measurements of binding thermodynamics can extend from the determination of affinity through to the estimation of changes in free energy, enthalpy and entropy by employing Arrhenius and van't Hoff plots.

The kinetic energy of the reactant molecules is dependent upon temperature, and an increase in temperature causes an increase in both the rate constant of association and that of dissociation. Calculations based upon reactant concentration and kinetic energy indicates that only a small portion of collisions result in a chemical reaction. In order for a collision to result in a chemical reaction, the colliding molecules must be in the correct orientation and possess a critical amount of energy required for the reaction. Thereby, conducting a reaction at a higher temperature delivers more energy into the system and increases the reaction rate by causing more collisions between particles, as explained by the collision theory. This is because more of the colliding particles will have the necessary activation energy, resulting in more successful collisions, when bonds are formed between reactants. Reaction rates for many reactions double or triple for every 10 degrees Celsius increase in temperature, although the effect of temperature may be very much larger or smaller than this.

5.7.6.1 - Activation energy

The minimum energy requirement of a molecule in order to allow it to react is defined as the activation energy (Ea) for the reaction. In a reaction between two or more molecules to from a product, the energy level often passes through a maximum called the activated or transition state, as shown in figure 5.13.

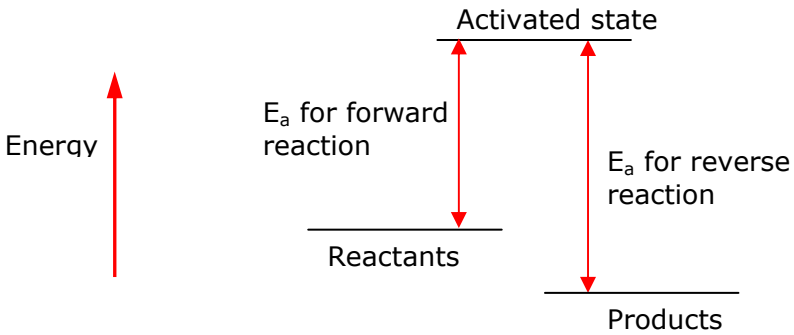


Figure 5.13 – A schematic diagram of the activation energy of a reaction

This energy barrier can be explained by the requirement of many products to be produced as a result of breaking of bonds to produce new ones. Zero activation energy only seems to occur when there is no breaking of chemical bonds [Laidler, 1978].

If the activation energy for a particular reaction is high only a small proportion of the molecules present will possess sufficient energy to react and therefore, the rate of the reaction will be slow. Conversely, if the activation energy for a reaction is low a large proportion of the molecules will have sufficient energy to react and the rate of reaction will be high.

5.7.6.2 - The Arrhenius law

The relationship between the rate constant, influence of temperature and activation energy is described by the Arrhenius equation, as shown in equation 5.15. The connection between the rate constant (k) of a reaction with absolute temperature (T) was expressed by van't Hoff and Arrhenius in 1887, and later extended by Arrhenius, to become the Arrhenius law. In addition, E_a is the activation energy (enthalpy), A is a constant independent of temperature and R is the universal gas constant.

$$k = A \exp(^{-Ea/RT)} \qquad \textit{(Equation 5.15)}$$

The approach of van't Hoff to this equation focussed on thermodynamics and energy levels. At equilibrium, the constant K is equal to the ratio of the rate constants k_{ass} and k_{diss}, and that these vary as shown in equation 5.16.

$$\frac{d\ln k_{ass}}{dT} - \frac{d\ln k_{diss}}{dT} = \frac{\Delta H^{o}}{RT^{2}} \qquad (\textit{Equation 5.16})$$

This can be split into two separate equations and integrated to produce equations 5.17 and 5.18. In these equations, E is the activation energy, and A is the frequency factor of the reaction. This can be done because the difference between the activation energies (enthalpies) E_1 and E_{-1} is equal to ΔH^{o}.

$$k_{ass} = A_1 \, e^{-E\,1/RT} \qquad (\textit{Equation 5.17})$$

$$k_{diss} = A_{-1} \, e^{-E\text{-}1/RT} \qquad (\textit{Equation 5.18})$$

Arrhenius reasoned that most collisions in chemical reactions occur with insufficient energy for products to be formed. The Arrhenius law (equation 5.15) resembles the Boltzmann distribution law (equation 5.19), which represents the fraction of molecular collisions that are equal to or in excess of particular energy, E.

$$e^{-E/RT} \qquad (\textit{Equation 5.19})$$

For the Arrhenius law, this is the activation energy, E_a. To use the law, natural logs can be taken to produce equation 5.20.

$$\ln k = \ln A - \frac{E_a}{RT} \qquad \textit{(Equation 5.20)}$$

In a moderate temperature range, E_a and A should remain approximately constant, and as such, an Arrhenius plot of ln k against 1/T gives a negative slope, which has a gradient equal to $-E_a/R$ and an intercept on the ordinate axis of ln A.

5.7.6.3 - van't Hoff's equation

ΔG can also be obtained relatively easily from the binding experiment using equation 5.21, where R is the gas constant and T is the absolute temperature in Kelvin.

$$\Delta G = -RT \ln K \qquad \textit{(Equation 5.21)}$$

van't Hoff combined this equation with equation 5.22, which relates ΔG to ΔH

$$\Delta G = \Delta H - T\Delta S \qquad \textit{(Equation 5.22)}$$

And this forms van't Hoff's equation (equation 5.23).

$$\ln K = \frac{\Delta H^o}{RT} + \frac{\Delta S^o}{R} \qquad \textit{(Equation 5.23)}$$

Similarly to an Arrhenius plot, over a moderate temperature range, ΔH and ΔS should remain approximately constant. A van't Hoff plot of ln$K$ against 1/T gives a negative slope, which has a gradient equal to $-\Delta H^o/R$ and an intercept on the ordinate axis of $\Delta S^o/R$.

5.7.7 - Determining the effect of temperature on D-SpG$_{C2}$-1 and Fc complex

The high melting point of the binding domains of SpG promoted interest in the effect of temperature on the rate constants of complex formation. Fast kinetic stopped-flow studies were carried out under identical conditions at various temperatures to measure the thermodynamic constants of a binding reaction between a domain of SpG and its ligand, Fc. The temperatures of both solutions were thermostatically controlled and k_{app} was measured as before.

The net effect of temperature on the affinity of the complex formed between D-SpG_{C2}-1 and Fc is dependent upon the activation energies of each step of the binding interaction. The values for the association and dissociation rate constants, as well as the pre-equilibrium Kd, were measured over the temperature range of 10 to 35 °C, in order to determine the activation energy of each rate constant. Typical plots, presented in figure 5.14, show the effect of increasing temperature on the association, dissociation and Kd constants respectively. The activation energy of the association and dissociation of the wild-type D-SpG_{C2}-1 binding interaction with Fc can be calculated from an Arrhenius plot of the natural logarithm (ln*k*) against 1/T, shown in figures 5.15 and 5.16, respectively. The gradient of the Arrhenius plots gives a negative slope, which has a gradient equal to $-E_a/R$ and an intercept on the ordinate axis of ln *A*.

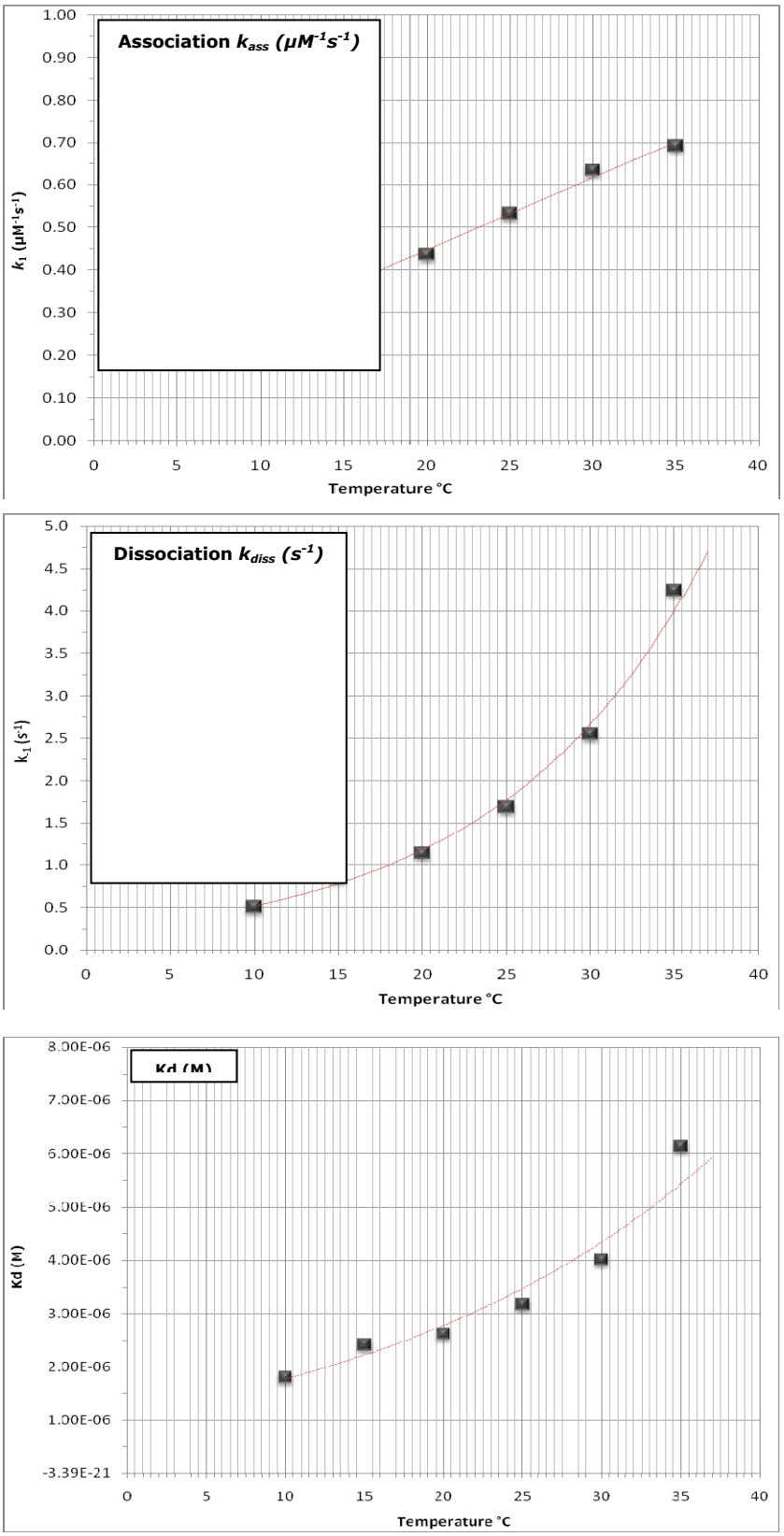


Figure 5.14 – <u>Typical plots of the effect of temperature on the Wt D-SpG$_{C2}$-1 and Fc complex</u>

These plots show the effect of temperature on the association constant k_{ass} ($\mu M^{-1}s^{-1}$), dissociation k_{diss} (s^{-1}) and Kd (M) plotted against temperature, at pH 8.0 between Wt D-SpG$_{C2}$-1 and Fc.

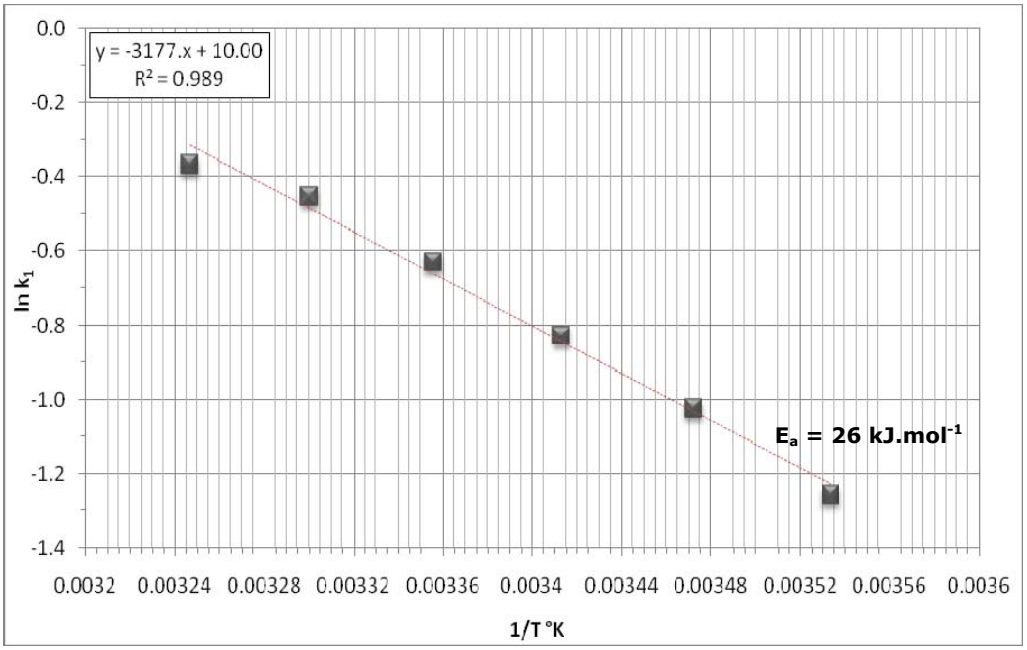


Figure 5.15 – <u>Typical Arrhenius plot of the effect of temperature on association of the Wt D-SpG$_{C2}$-1 and Fc complex</u>

This graphical determination represents an Arrhenius plot of the lnk against 1/varying temperatures (°K), at pH 8.0, for the association of the Wt D-SpG$_{C2}$-1 and Fc complex.

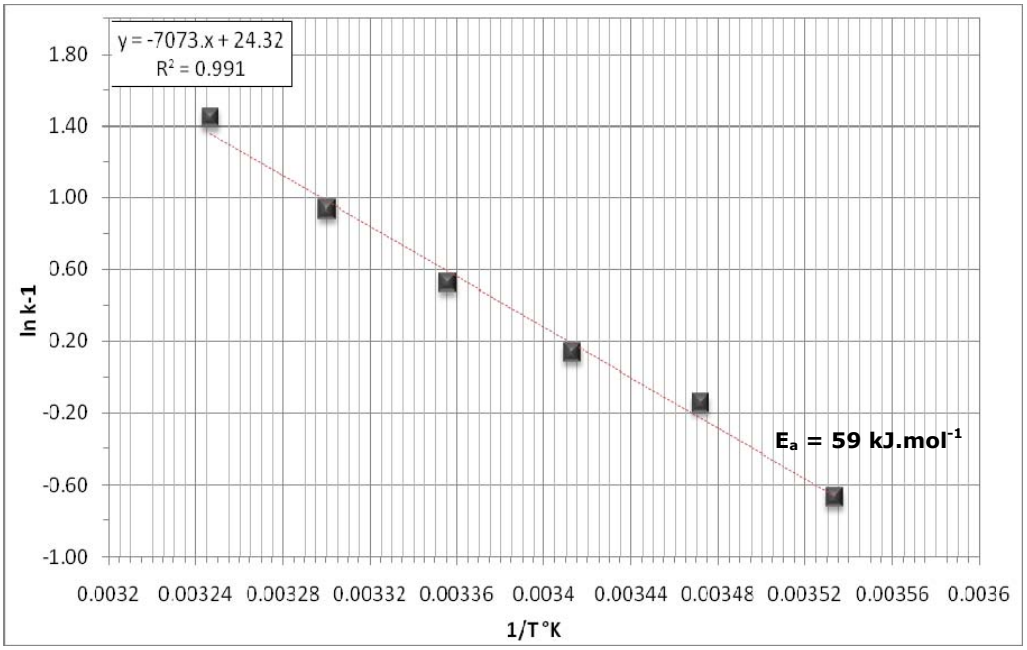


Figure 5.16 – <u>Typical Arrhenius plot of the effect of temperature on dissociation constant of the Wt D-SpG$_{C2}$-1 and Fc complex</u>

This graphical determination represents an Arrhenius plot of the lnk against 1/varying temperatures (°K), at pH 8.0, for the dissociation of the Wt D-SpG$_{C2}$-1 and Fc complex.

A summary of the effect of temperature on the association, dissociation and the binding affinity of the complex between wild-type D-SpG$_{C2}$-1 and Fc is presented in figure 5.17 and table 5.6.

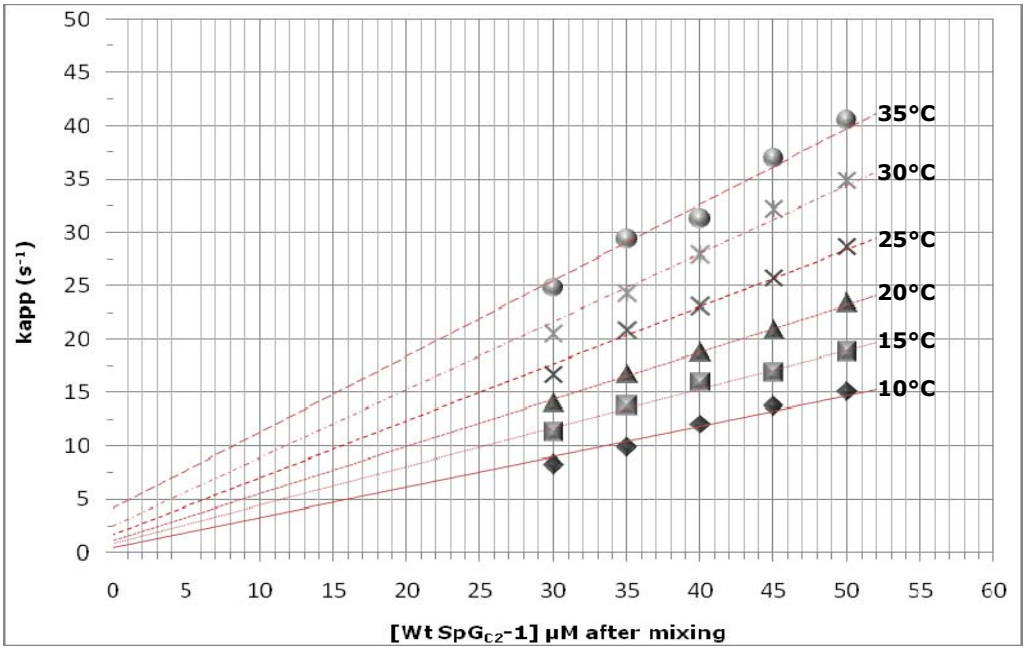


Figure 5.17 – Typical stopped-flow analysis of the Wt D-SpG$_{C2}$-1 and Fc complex at varying temperatures

This graphical determination represents a linear plot of k_{app} (s^{-1}) against varying concentrations of Wt D-SpG$_{C2}$-1 at temperatures 10 °C to 35 °C, pH 8.0. The rate constants k_{ass} and k_{diss} can be determined from the gradient of the line and the intercept of the y-axis respectively.

Temp (°C)	°Kelvin (+ 273)	k_{ass} ($\mu M^{-1}s^{-1}$) ± std dev. n≥2	k_{diss} (s^{-1}) ± std dev. n≥2	Kd (µM) ± std dev. n≥2
10	283	0.2842	0.5112	1.80
15	288	0.3587	0.8651	2.41
20	293	0.437	1.143	2.62
25	298	0.5322	1.689	3.17
30	303	0.6349	2.547	4.01
35	308	0.6925	4.246	6.13

Table 5.6 – Summary of the effect of temperature on the rate constants and binding affinity obtained for the interaction of Wt D-SpG$_{C2}$-1 and Fc

Results obtained by stopped-flow to determine the effect of increasing temperature on the association and dissociation of the complex between wild-type D-SpG$_{C2}$-1 and Fc, give positive activation energy values and show that increasing the temperature leads to an increase in the rate constants. Figure 5.18 presents a typical van't Hoff plot for the interaction of wild-type D-SpG$_{C2}$-1 with Fc. From the plot of the lnKa against 1/ temperature °Kelvin, thermodynamic parameters can be determined for the reaction. ΔH = slope = $-\Delta H/R$; ΔS = intercept = $\Delta S/R$ and $\Delta G = \Delta H - T\Delta S$. A summary of Kd and ΔG values at 15 °C (288 °K) are presented in table 5.7.

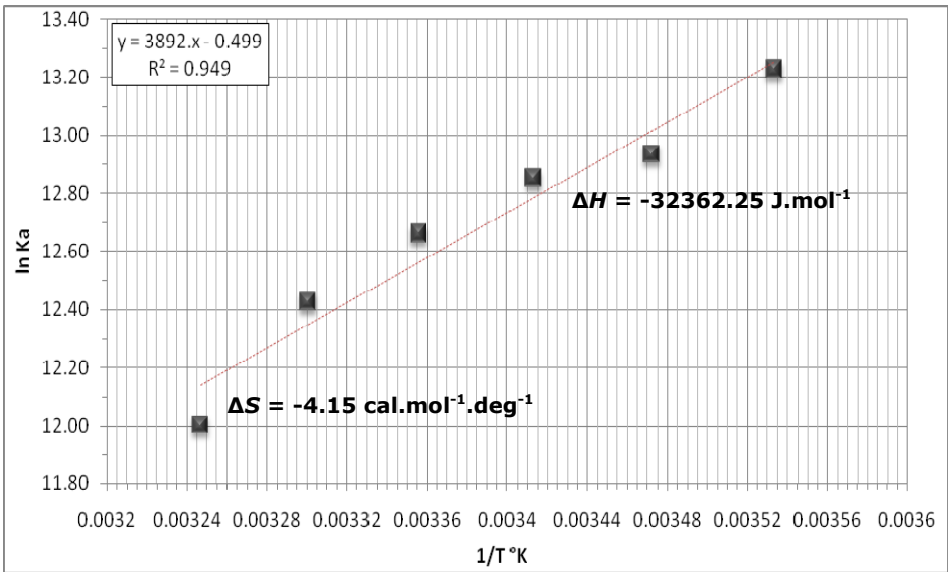


Figure 5.18 – Typical van't Hoff plot of the effect of temperature on the Wt D-SpG$_{C2}$-1 and Fc complex

This graphical determination represents a van't Hoff plot of the effect of varying temperatures on the formation of Wt D-SpG$_{C2}$-1 and Fc complex, pH 8.0.

Temp (°C)	°Kelvin (+ 273)	ΔG (cal.mol^{-1})	Kd (µM)
10	283	-31187.93	1.75
15	288	-31167.18	2.22
20	293	-31146.44	2.80
25	298	-31125.69	3.50
30	303	-31104.94	4.34
35	308	-31084.19	5.35

Table 5.7 – Summary of the effect of temperature on ΔG and the pre-equilibrium Kd obtained for the interaction of Wt D-SpG$_{C2}$-1 and Fc

5.7.8 – Determining the effect of temperature on the reaction between D-SpG$_{C2}$-1 mutants and Fc

The effect of increasing temperature was also studied on a selection of D-SpG$_{C2}$-1 mutants and their respective complex with Fc. Figures 5.19 to 5.22 present van't Hoff plots, for mutant constructs Q31E, E41Q N34A and N34D, respectively, from which thermodynamic parameters were calculated and compared with values obtained for wild-type D-SpG$_{C2}$-1. All data are summarised in table 5.8.

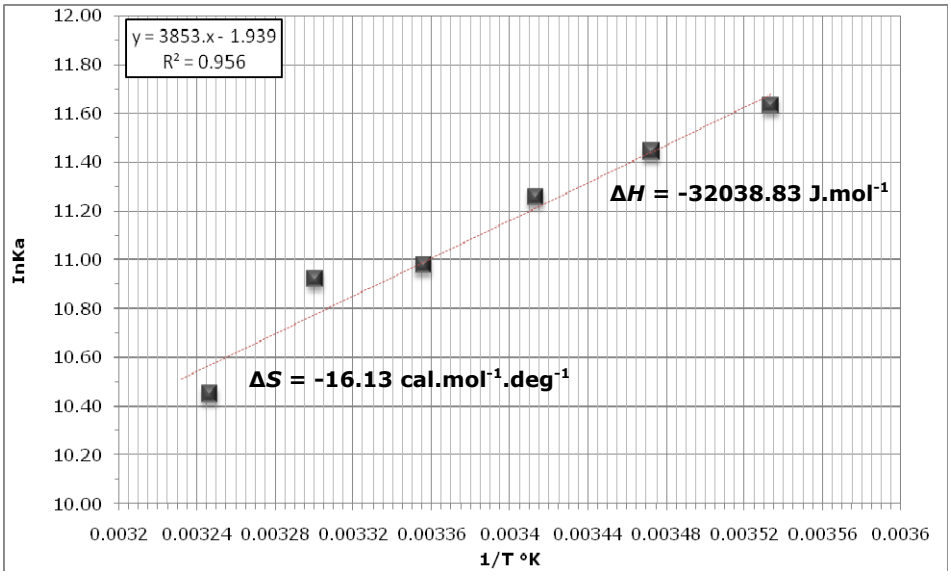


Figure 5.19 – Typical van't Hoff plot of the effect of temperature on the formation of Q31E D-SpG$_{C2}$-1 and Fc complex

This graphical determination represents a van't Hoff plot of the effect of varying temperatures on the formation of mutant Q31E D-SpG$_{C2}$-1 and Fc complex, pH 8.0.

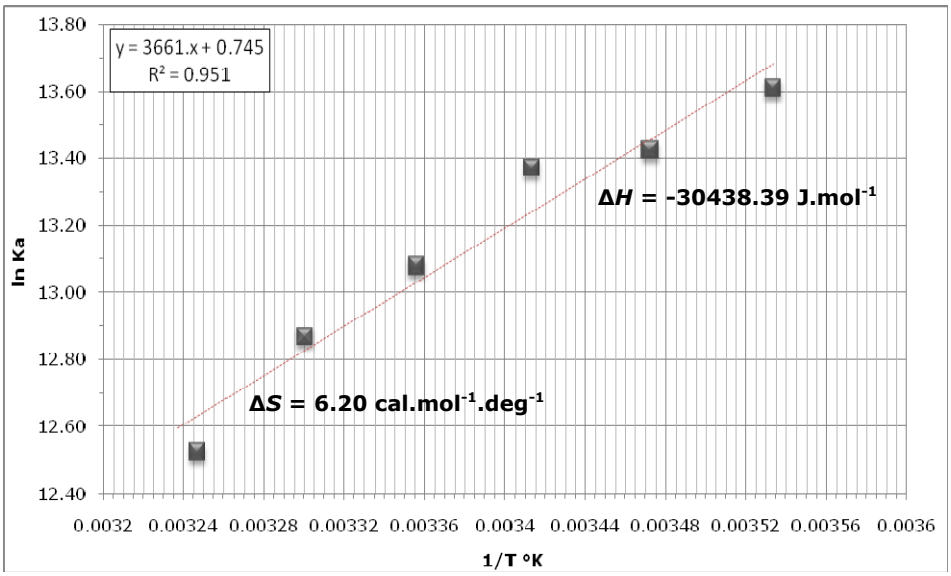


Figure 5.20 – Typical van't Hoff plot of the effect of temperature on the formation of E41Q D-SpG$_{C2}$-1 and Fc complex

This graphical determination represents a van't Hoff plot of the effect of varying temperatures on the formation of mutant E41Q D-SpG$_{C2}$-1 and Fc complex, pH 8.0.

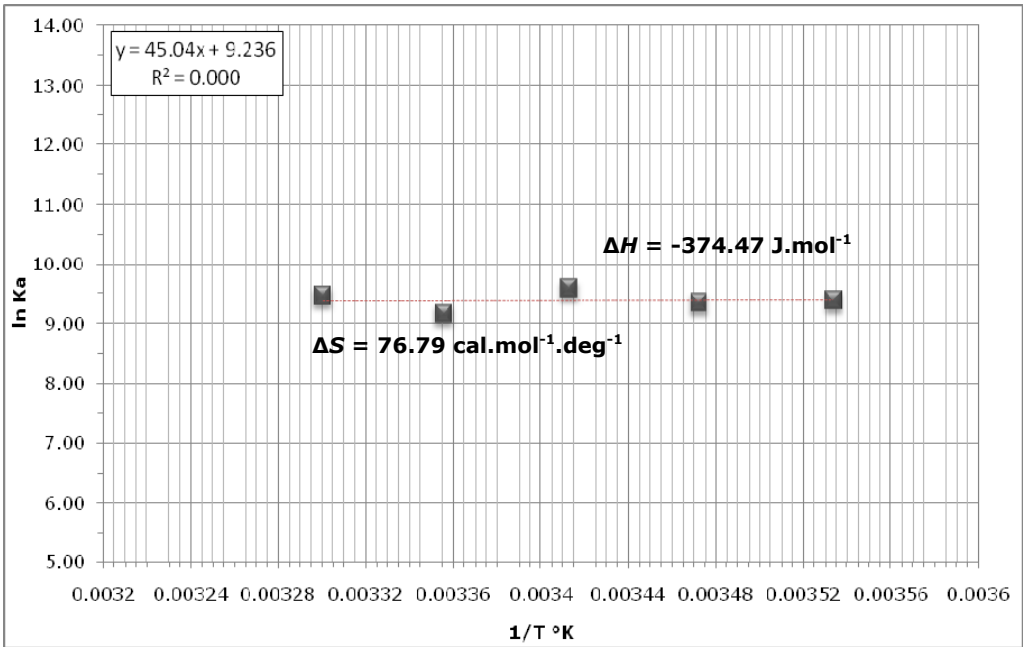


Figure 5.21 – Typical van't Hoff plot of the effect of temperature on the formation of N34A D-SpG$_{C2}$-1 and Fc complex

This graphical determination represents a van't Hoff plot of the effect of varying temperatures on the formation of mutant N34A D-SpG$_{C2}$-1 and Fc complex, pH 8.0.

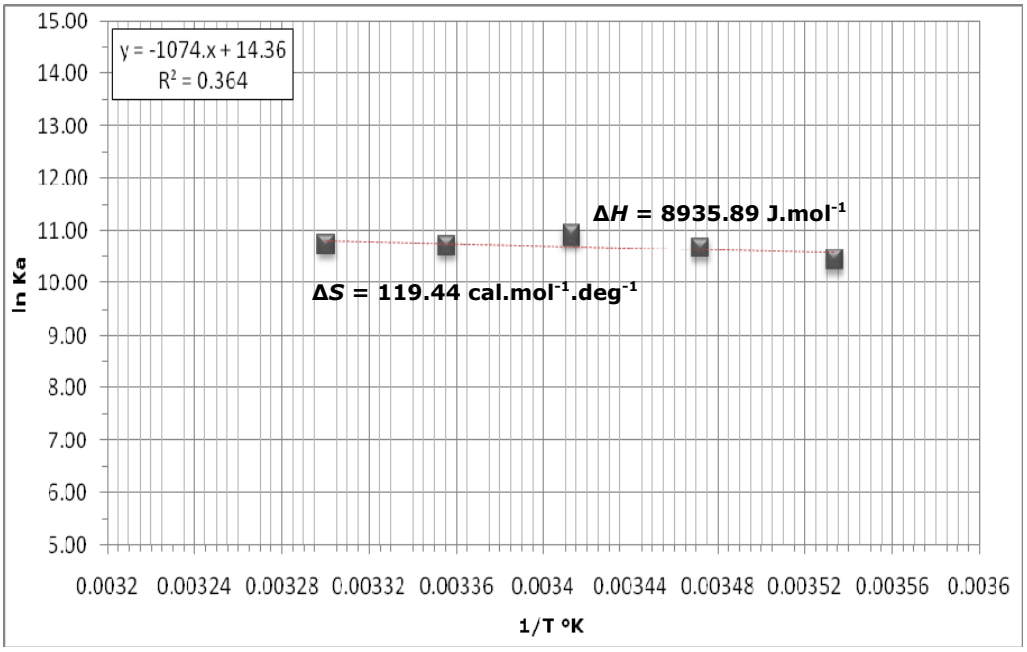


Figure 5.22 – Typical van't Hoff plot of the effect of temperature on the formation of N34D D-SpG$_{C2}$-1 and Fc complex

This graphical determination represents a van't Hoff plot of the effect of varying temperatures on the formation of mutant N34D D-SpG$_{C2}$-1 and Fc complex, pH 8.0.

D-SpG$_{C2}$-1 Construct	Kd @ 15 °C (μM) n≥2	E$_a$ k_{ass} (kJ.mol^{-1}) n≥2	E$_a$ k_{diss} (kJ.mol^{-1}) n≥2	Δ*H* (cal.mol^{-1}) n≥2	Δ*S* (cal.mol^{-1}.deg^{-1}) n≥2	Δ*G* @ 15 °C (cal.mol^{-1}) n≥2
WT	2.22	26.42	58.81	-32362.25	-4.15	-31167.18
Q31E	10.75	35.92	68.44	-32038.83	-16.13	-27394.11
N34A	83.58	57.83	58.20	-374.47	76.79	-22489.44
N34D	24.08	48.87	39.88	8935.89	119.44	-25462.52
E41Q	1.40	33.26	63.79	-30438.39	6.20	-32223.20

Table 5.8 – Summary of the effect of temperature on the thermodynamic parameters and binding affinity rate constants obtained for the interaction of D-SpG$_{C2}$-1 mutants and Fc

Studies to determine the effect of increasing temperature on the interaction between D-SpG$_{C2}$-1 mutants and Fc are presented in table 5.7. Thermodynamic constants for each binding interaction with Fc were determined from Arrhenius and van't Hoff plots. Δ*G* values were calculated at 15 °C so that a direct comparison with data obtained from ITC studies, chapter 6.0, can be made.

5.7.9 - Calculating the effect of pH on D-SpG$_{C2}$-1 and Fc complex

The native three-dimensional structure of a protein is determined by the physicochemical nature of its constituent amino acids. The twenty different types of amino acids, depending on their physicochemical properties, can be grouped into three major classes, hydrophobic, hydrophilic and charged. Charged residues are also known as ionisable groups, and a protein's structure, stability and function, as well as the net charge on a protein depends on the balance of the individual positive, negative and neutral charges present on the amino terminus, the carboxyl terminus and any ionisable side-chains that are present [Shaw *et al*., 2001].

Amino groups of proteins are basic groups and are positively charged except at high pH. The overall charge on a protein influences the ionisation of individual residues, and the net charge varies with pH, as well as being determined by the p*K* values of the ionisable groups [Thurlkill *et al*., 2006]. There are two groups of ionisable amino-acid side-chains depending on the charge found on the acidic form. Type 1 groups are neutral in their acidic form, dissociate to form a proton and a conjugate base, and carry a negative charge when the pH is greater than the p*K*a. This group includes the carboxyl group of the C-terminus, as well as on

the side-chains of aspartate (aspartic acid) and glutamate (glutamic acid), and the phenolic side-chain of tyrosine. In contrast, type 2 ionisable groups are positively charged, dissociate to form a proton and an uncharged conjugate base, and carry a positive charge when the pH is less than the p*K*a. Type 2 groups found in proteins include the alpha-amino group of the N-terminus, and the amino groups found on the side-chain of lysine, the imidazole side-chain of histidine, and the guanidinium group of arginine [Schutz and Warshel, 2001].

To evaluate the ionisation states in proteins requires the p*K* to be determined. The term p*K* is related to the state of ionisation or dissociation of a molecule or functional group on a molecule and the p*K*a is a measure of proton binding affinity [Pace *et al*., 2000]. These values provide important electrostatic information of the ionised groups in proteins, as well as the electrostatic energies in protein binding sites. The ionisation of charged residues in a binding interaction depends on their appropriate ionisation states at physiological pH [Pace *et al*., 2000], and the distribution of surface-charged residues is also critical for protein-protein association. Furthermore, changes in the distribution of charge can occur at the site of the interaction [Sheinerman *et al*., 2000]. Ionisable groups can be influenced by the chemical structure and when proteins fold and interact, the perturbations of the p*K*s of the ionisable groups on the surface of the protein are determined mainly by charge-charge interactions with other ionisable groups. However, large positive and negative perturbations often occur if these groups are partially or fully buried in the protein interior [Schutz and Warshel, 2001].

If a functional group on a protein is said to have a low pK (<2.5), it has a low affinity for protons and gives them up readily, whereas a high pK (>3.0) binds protons strongly, and gives them up reluctantly. The p*K*a rule for predicting the dissociation and therefore, determining the ionisation state of a molecule or functional group is that when the pH of a solution is below the p*K* of a molecule then the un-dissociated molecule and the acid form of the compound predominates. Whereas, when the pH of a solution is the same as the p*K* of a molecule then the concentration of the un-dissociated molecules is equal to the concentration of the dissociated molecules, and when a solution pH is above the p*K* of a molecule then the dissociated molecule and the conjugate base [Thurlkill *et al*., 2006].

The effect of an ionisation process on the binding interaction can be determined by using stopped-flow to study the pH dependence of the apparent second-order complex-formation rate constant (k_{ass}) and of the first-order complex-dissociation

rate constant (k_{diss}) of D-SpG$_{C2}$-1 with Fc. These kinetic parameters will change as the reaction pH increases or decreases below the p*K*, depending on whether the group is protonated or deprotonated in order to function.

The pH profile of the reaction can be constructed by employing an adapted Dixon-Webb plot, whereby the log of Kd is plotted against the varying pH values of the study [Dixon and Webb, 1979]. If an ionisation process has an effect on the interaction, the adapted Dixon-Webb plot will consist of straight-line sections joined by short curved parts. The p*K* of the ionising group in one of the components is defined by the curve, concurrently with the straight-line sections that intersect at a pH and correspond to the p*K* [Dixon and Webb, 1979].

5.7.10 – Determining the effect of pH on D-SpG$_{C2}$-1 and Fc complex

To enable the binding interaction between SpG and Fc to be fully characterised, knowledge of the contribution from each amino-acid involved is required. The p*K* changes of ionisable groups that may occur upon binding can be determined by the pH dependence of the kinetics of Fc binding. From these studies it is possible to obtain information regarding the charge and p*K* values of the important residues.

A detailed pre-equilibrium binding study of the electrostatic nature of the interaction was performed using a triple buffer system to buffer a wide range of pH conditions, namely pH 4.0 to pH 9.0, at which both proteins were considered stable. The pH dependence of the second-order association rate constant k_{ass} and of the first-order dissociation rate constant k_{diss} for wild-type D-SpG$_{C2}$-1 and Fc is presented in figure 5.23. The effect on the Kd binding affinity is also presented in figure 5.23, and a summary of all the values obtained is shown in table 5.9. These results show that the interaction between wild-type D-SpG$_{C2}$-1 and Fc is pH dependent, exhibiting a maximum binding affinity at pH 6.0, which compares well with results obtained by Walker [Walker, 1994]. This also agrees with results obtained by immobilisation techniques for SpG [Pilcher *et al*., 1991].

The different affinities at pH values outside this range may be contributed to a change in the association and dissociation rates or a combination of both. The results obtained for both of these rate constants show that the effect of pH on the Kd appears to be mediated through changes in the rate of association of the complex, as the dissociation is virtually unchanged over the same pH range. This seems to indicate that charged residues are involved in the formation of the

complex between wild-type D-SpG$_{C2}$-1 and Fc, but not with the dissociation of the complex. These results also compare well with results obtained by Walker [Walker, 1994].

pH	k_{ass} ($\mu M^{-1}s^{-1}$) ± std dev. n≥2	k_{diss} (s^{-1}) ± std dev. n≥2	Kd (μM^{-}) ± std dev. n≥2
4.0	0.43 ± 0.04	1.29 ± 0.06	2.99 ± 0.13
5.0	1.04 ± 0.14	2.10 ± 0.46	2.08 ± 0.74
6.0	1.01 ± 0.10	1.54 ± 0.15	1.51 ± 0.23
7.0	0.65 ± 0.04	1.23 ± 0.25	1.90 ± 0.13
8.0	0.34 ± 0.02	1.08 ± 0.11	3.21 ± 0.54
9.0	0.24 ± 0.02	1.02 ± 0.04	4.23 ± 0.24

Table 5.9 – Summary of the effect of pH on the rate constants and binding affinity obtained for the interaction of D-SpG$_{C2}$-1 mutants and Fc

The influence of the ionisation process of one or more of the residues participating in the interaction is then determined from the adapted Dixon-Webb plot, presented in figure 5.24. This plot indicates that an ionisable group with a p*K* of around 6.3 may be important for binding.

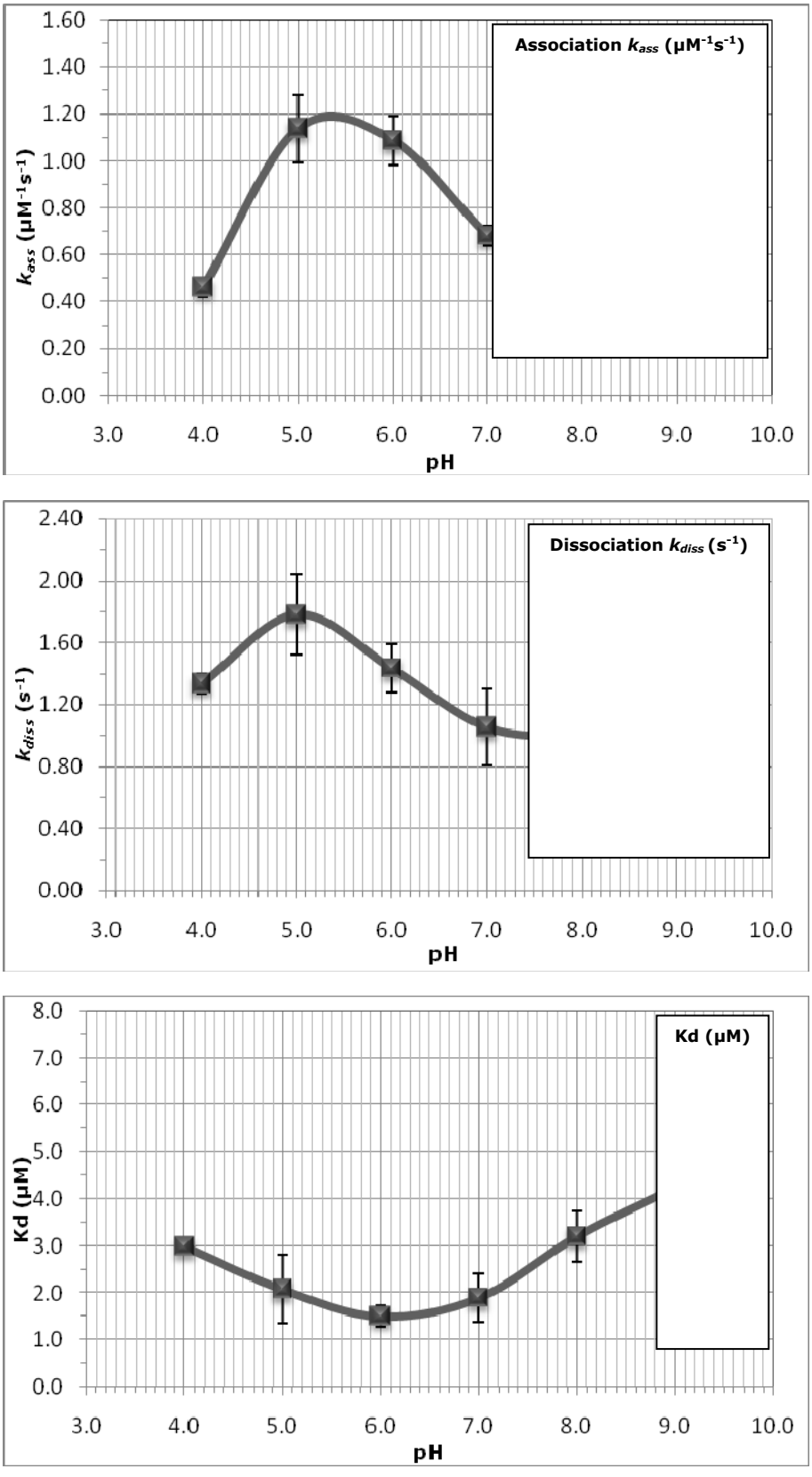


Figure 5.23 – <u>pH profiles of the effect of pH on the Wt D-SpG$_{C2}$-1 and Fc complex</u>

These plots represent pH profiles of the association, dissociation rate constants and Kd against varying pH, at 15 °C, between Wt D-SpG$_{C2}$-1 and Fc.

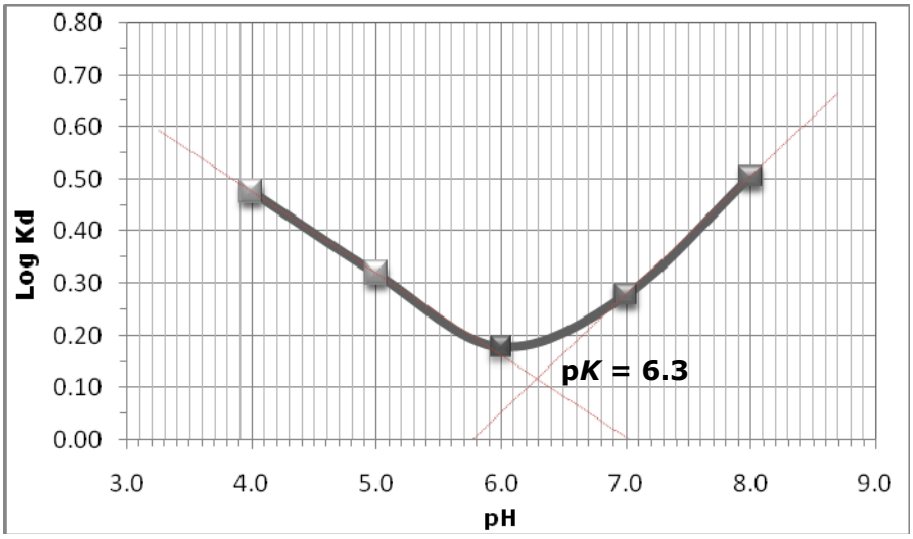


Figure 5.24 – Dixon-Webb plot of the effect of pH on the formation of Wt D-SpG$_{C2}$-1 and Fc complex

This graphical determination represents a Dixon-Webb plot of the formation of Wt D-SpG$_{C2}$-1 and Fc against varying pH at 15 °C.

There are a number of ionisable residues in SpG that are involved in the interfacial interaction with Fc, namely Glu26, Lys27, Lys30, Asp39 and Glu41. At high pH, tyrosines on the Fc, which have been implicated in binding [Stone *et al*., 1989] may also begin to ionise and disrupt the interaction. The p*K* of around 6.3 suggests that one or more of the acidic residues Glu26, Asp39 or Glu41 may be responsible for this pH effect.

5.7.11 – Determining the effect of pH on D-SpG$_{C2}$-1 mutants and Fc complex

To further determine the effect of pH on the formation of the complex between SpG and Fc, the p*K* of the ionisable groups derived from aspartate, glutamate and histidine in a range of mutant constructs of D-SpG$_{C2}$-1 was calculated and compared with p*K* values obtained for wild-type. The mutant constructs possessing these ionisable groups are Q31E, Q31H and N34D, and will allow the effect of pH on the individual charged amino acid side-chains to be determined. The effect of removing an ionisable group was also investigated with E41Q, as this mutant has been shown to affect the binding of SpG with Fc.

The influence of the ionisation process of one or more of the residues participating in the interaction is then determined from the adapted Dixon-Webb plot, presented in figure 5.25 for mutants Q31E, Q31H and N34D, respectively. A summary of the Dixon-Webb plots of these mutants compared with wild-type is presented in figure 5.26.

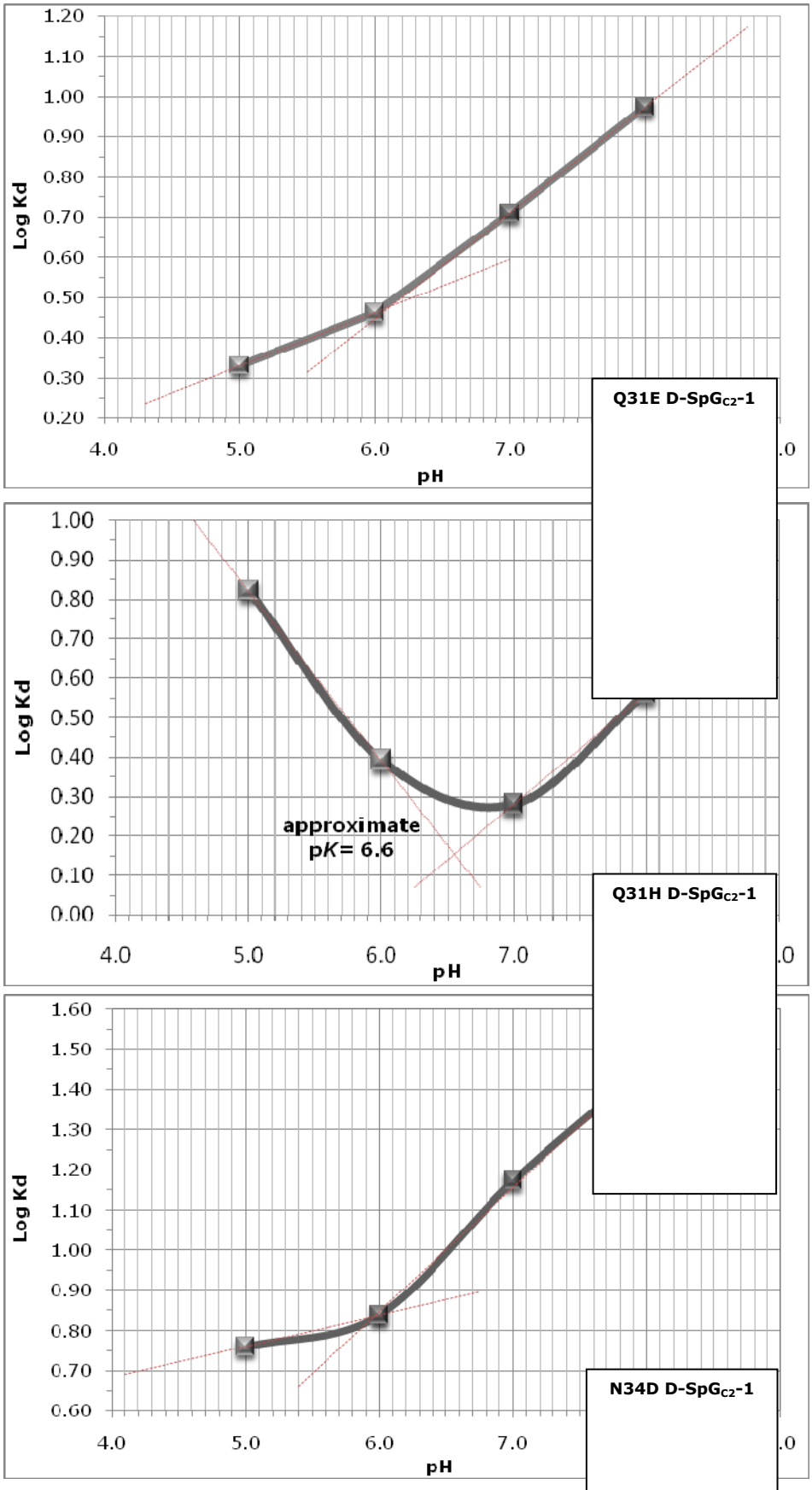


Figure 5.25 – Typical Dixon-Webb plots of the effect of p ion of D-SpG$_{C2}$-1 constructs and Fc complex

This graphical determination represents a Dixon-Webb nation of D-SpG$_{C2}$-1 constructs (as indicated) and Fc against varyi

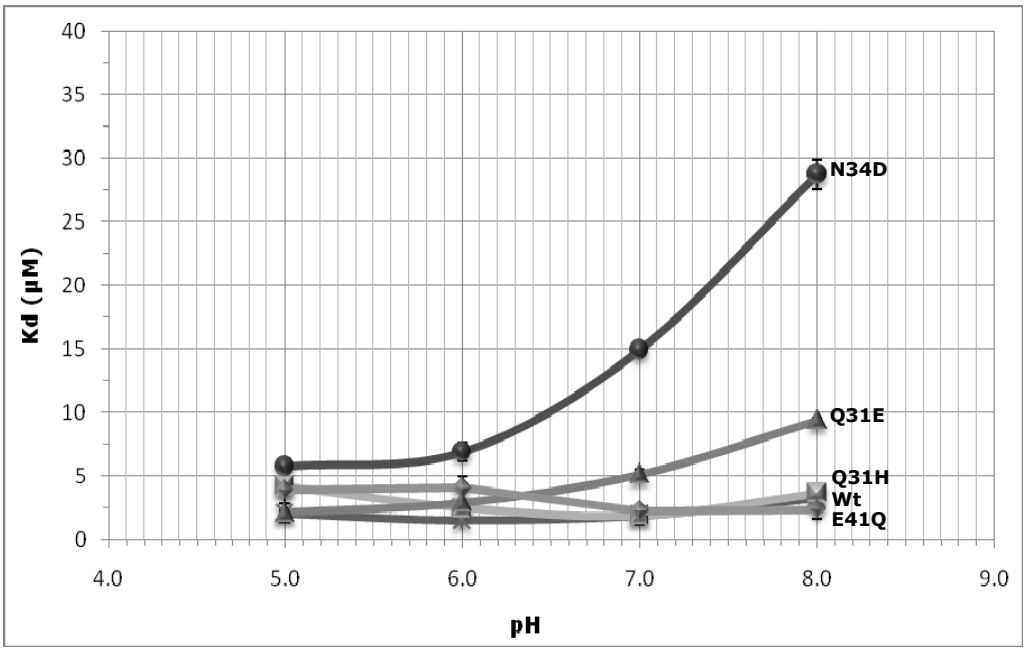


Figure 5.26 - Summary pH profile of the effect of pH on Kd of the formation of D-SpG_{C2}-1 mutants and Fc complex

This graphical determination represents a summary pH profile of the Kd (µM) against varying pH, at 15 °C, between D-SpG_{C2}-1mutant constructs and Fc.

The mutations Q31E and N34D have both altered the pH vs log Kd profile indicating that ionisation of the replacement is affecting the Kd of the complexes. The Q31H and the E41Q mutants have both got wild-type profiles, suggesting these mutations have little or no effect on pH sensitivity.

5.8 – Discussion

In this chapter fluorescence has been employed to study the binding interaction between SpG and Fc. The presence of a native tryptophan residue in SpG and its implication with Fc has allowed this residue to be exploited as a reporter group for monitoring Fc binding. In addition, this binding interaction has permitted the characterisation of the reversible interactions in terms of binding, stoichiometry, kinetics and thermodynamics, and enabled the reliable determination of these parameters.

Equilibrium studies using fluorescence titrations provided a simple and effective method for determining the true dissociation constant for the complex and allowed the number of binding sites to be calculated. Pre-equilibrium studies employed fast kinetic stopped-flow to determine the rates of association and dissociation of the complex between SpG and Fc. Temperature and pH studies also showed that the interaction was dependent on these environmental

conditions. The results suggest that amino acid residues at positions 31 and 34 should be neutral for maximum binding affinity.

CHAPTER 6.0

ISOTHERMAL TITRATION CALORIMETRIC STUDIES ON THE SPG-FC BINDING INTERACTION

6.1 – Introduction

Molecular recognition events involving macromolecules, such as proteins, are fundamental to almost every cellular process, and understanding the basis for protein-protein interactions is aided by the full characterisation of the thermodynamics of their association. Determination of reaction thermodynamics can extend the determination of affinity and stoichiometry to the estimation of changes in free energy, enthalpy, entropy and heat capacity.

These and other important parameters of a reaction may be derived from secondary graphical plots of kinetic and equilibrium data (Arrhenius and van't Hoff analyses) obtained at different temperatures, but can also be obtained directly from calorimetric studies. The direct assessment of the thermodynamics associated with a binding reaction by Isothermal Titration Calorimetry (ITC) often provides more information than that gained by equilibrium studies such as fluorescence titrations. Firstly, it is sometimes necessary to engineer a fluorescence reporter group into one reactant which may alter the structure of that molecule and hence the stability of the complex. Secondly, changes in the thermodynamical contributions of a reaction are not always accompanied by changes in affinity. The value of ΔG (from which the Kd of a complex is calculated) may remain unchanged even if the contributions of enthalpy and entropy to the binding reaction both change. Thus, Kd would not reflect the difference of the changes in bonding characteristics. The enthalpy of binding involves the formation and breaking of non-covalent bonds, particularly hydrogen bonds and van der Waals; a negative ΔH value indicates a net gain of the strength or number of these bonds, whereas a positive ΔH value indicates a loss [Holdgate and Ward, 2005].

In the following studies, ITC was used as a direct thermodynamic approach to study the association of the complex between SpG and its ligands Fc and Fab, and to provide a complete characterisation of the binding energetics of the reaction. This technique is based on the measurement of the heat absorbed or released (the enthalpy change) upon titration of a solution of one reactant with that of a second reactant. This is because every molecular interaction either generates heat or absorbs small amounts of heat in direct proportion to the amount of binding that occurs [Doyle, 1997]. When the macromolecule becomes saturated with the ligand, the heat signal diminishes until only the background heat of the dilution of the titrating reactant is observed. If the reactants are in the appropriate concentration range and if the Kd of the equilibrium is fairly small

(e.g. Kd > μM) then ITC experiments can detect these small changes in heat. ITC therefore represents a very powerful tool for the characterisation of protein-protein ligand interactions.

6.2 - ITC instrumentation

An Isothermal Titration Calorimeter instrument is composed of two, thermally insulated and identical cells - a sample cell and a reference cell. Sensitive thermocouple circuits are used to detect temperature differences between the reference cell (which is filled with water or buffer) and the sample cell which contains the non-titrating reactant. Both cells are surrounded by an inner shield, which is maintained at the same temperature as the cells to minimise any heat flow. The outer shield is also maintained at a constant temperature, to compensate for any changes in room temperature. A long-needle syringe, fitted with a paddle on its end, contains the ligand. After filling, the syringe is seated into the holder, which is coupled to an injector system under software control. The syringe paddle is located inside the sample cell and the entire syringe assemble rotates at a pre-selected speed to continuously provide mixing to the ligand contents of the sample cell. The main components of a MicroCal VP-ITC system employed for all ITC experiments are presented in figure 6.1. The binding of the two proteins causes heat to be either taken up or evolved, depending on the nature of the reaction.

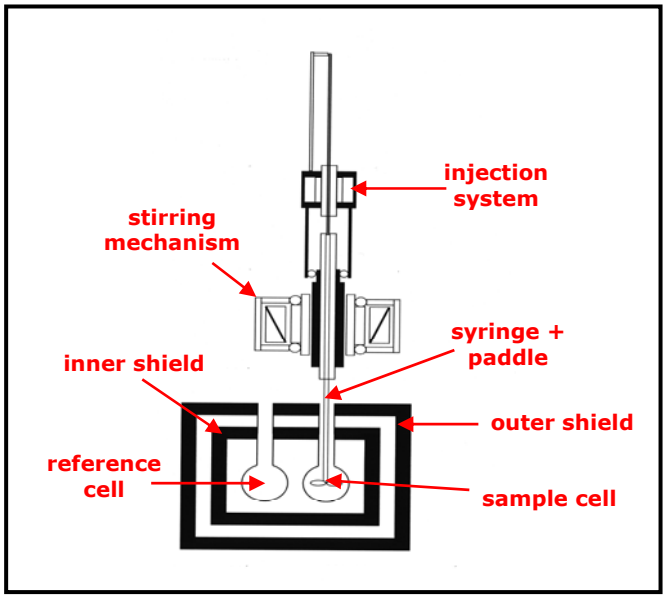


Figure 6.1 – An adapted diagram of the components of an isothermal titration calorimeter *[Chellani, 1999]*

The VP-ITC instrument employs two identical cells. All binding processes occur in the sample cell. An experiment consists of series of injections of a ligand (in this case SpG solution) from the syringe into a solution of Fc contained in the sample cell.

In an exothermic reaction, the temperature in the sample cell increases upon addition of ligand. This causes the feedback power to the sample cell to be decreased, and power is applied to the reference cell in order to maintain an equal temperature between the two cells. In an endothermic reaction, the opposite occurs, and the feedback circuit increases the power to the sample cell in order to maintain a constant temperature.

Observations are plotted as the power in µcal/sec needed to maintain the reference and the sample cell at an identical temperature. These pulses of power are integrated with respect to time, giving the total power applied per injection. The power applied (equivalent to heat (ΔH) released) upon each injection is presented in the top panel of the isotherm (see figure 6.2). As successive amounts of the ligand are titrated into the cell, the quantity of heat absorbed or released is in direct proportion to the amount of binding occurring. The majority of the added ligand becomes bound during the initial injections because of the excess macromolecules, and gives a measure of ΔH. The heat change becomes smaller throughout the titration as the available binding sites are filled. As the system reaches saturation, the signal changes diminish until only the heat of dilution of the ligand are observed. A binding curve, together with a line of best fit, is then obtained from a plot of the heat values from each injection against the ratio of ligand and binding partner in the cell, and allows the estimation of Kd. This is presented in the lower panel of the isotherm. The isotherm describing complex formation as a function of the molar ratio of ligand to macromolecule can then be analysed to give the thermodynamic parameters of the interaction under study. In addition, assuming that the concentrations of the SpG and Fc or Fab are known, the stoichiometry of the complex is easily calculated.

6.3 - Calculation of thermodynamic constants from ITC data

The main factor contributing to the entropic component of a binding reaction is the effect of the movement of water molecules from an ordered state - when hydrating one or more components - to a disordered state if they are "squeezed out" from binding interfaces when a complex forms, in which case their entropy increases. That is the burial of the water-accessible surface areas of reactants on binding, results in the release of solvent molecules and contributes a positive effect to the total entropy of the interaction, which is usually greater than the decrease in entropy caused by the reactant molecules becoming more ordered. Negative entropy values are proposed to be due to conformational restrictions that must be overcome for complex formation [Perozzo *et al*., 2004].

Gibbs free energy (ΔG) can be the same for a binding reaction, which has positive ΔH and ΔS values, denoting that binding is dominated by the hydrophobic effect, as complex formation, which has negative ΔH and ΔS values, denoting that binding is dominated by specific interactions [Holdgate and Ward, 2005]. The change in ΔG for the formation of the SpG•Fc complex, as shown in equation 6.2, is related to the standard Gibbs free energy change (ΔG°), under defined conditions, using the following relationship:

$$\Delta G = \Delta G^\circ + RT\ln\left(\frac{[SpG\bullet Fc]}{[SpG][Fc]}\right) \qquad \textit{(Equation 6.1)}$$

At equilibrium, under standard conditions, where $\Delta G=0$, equation 6.1 may become:

$$\Delta G^\circ = -RT\ln\left(\frac{[SpG\bullet Fc]}{[SpG][Fc]}\right) = -RT\ln Ka = -RT\ln(1/Kd) \qquad \textit{(Equation 6.2)}$$

Where, R is the gas constant and is equal to 8.314472 $kJ^{-1}mol^{-1}$, T is the absolute thermodynamic temperature in °Kelvin and is equal to 288 at 15 °C, Ka is the equilibrium association constant and Kd is the equilibrium dissociation constant.

This relationship demonstrates that the value of Kd is dependent upon the free energy of binding, ΔG°, and can be measured using a variety of methods. However, a full thermodynamic characterisation requires that the enthalpy change, which reflects the heat released or taken up during the association reaction, will also be measured. Free energy changes associated with protein-protein interactions are usually the result of strong enthalpic and entropic contributions [Pierce *et al*., 1999]. The relationship between the free energy change, and enthalpy and entropy changes, is given by equation 6.3.

$$\Delta G = \Delta H - T\Delta S \qquad \textit{(Equation 6.3)}$$

This is known as the Gibbs-Helmholtz equation, and is a thermodynamic equation used for calculating changes in the Gibbs energy of a system as a function of temperature.

6.4 – Observation of the formation of the Wt SpG•Fc complex by ITC

During a typical binding experiment, a solution of SpG (ligand in syringe) was titrated in precisely known volumes against a solution of Fc or Fab fragments, located in the sample cell, at constant temperature. Typically a solution of 600 µM Wt D-SpG_{C2}-1 was titrated into 2.0 ml of 10 µM Fc using injections of 1 × 2 µl and

then between 20 and 49 × 5 µl. This method produces a binding isotherm for the Fc-Wt D-SpG_{C2}-1 interaction with a minimum of -0.3 µcal/sec, which is sufficient to be analysed. Data were analysed using the manufacturer's Origin™ software. Figure 6.2 presents a binding isotherm of Wt D-SpG_{C2}-1 titrated into 2.0 ml of 20 mM phosphate buffer at pH 8.0, and figures 6.3 and 6.4 present typical binding isotherms of Wt D-SpG_{C2}-1 titrated into 2.0 ml of Fc observed by ITC at 15 °C and 25 °C respectively. These experiments demonstrate that it is the addition of Fc that produces the heat released upon the interaction. Titrations were carried out in triplicate for the wild-type protein with Fc, and typical isotherms of the data obtained are presented. Analysis of all data by the Origin™ software, is also presented as an inset to each binding isotherm. Formation of the SpG•Fc complex is an exothermic reaction, with heat being released upon complex formation, therefore injections of SpG into Fc resulted in a decrease in the electrical energy required to maintain the temperature difference between the sample and reference cells.

Thermodynamic constants, *ΔH* and *ΔS,* and the equilibrium dissociation constant, Kd, obtained by titrating 600 µM Wt D-SpG_{C2}-1 into 10 µM Fc, are presented in table 6.1 which shows that the binding reaction is driven by negative enthalpic changes (Δ*H* **=** -50.02 ± 5.15 kJ.mol^{-1}) although entropic changes (Δ*S* = -56.07 ± 17.15 J.mol^{-1}.deg^{-1}) are unfavourable for the reaction in this direction. The Gibb's free energy change (Δ*G)* of -33.93 ± 0.48 kJ.mol^{-1} shows that the reaction occurs spontaneously with a Kd of the complex under these conditions of 677 nM ± 154.1. This latter value is very similar to that determined by fluorescence titrations performed under the same experimental conditions in which the Kd value for Wt D-SpG_{C2}-1 binding to Fc was found to be 506 nM ± 0.1 (see chapter 5). This also compares reasonably well with previous data obtained by Gallagher [Gallagher, 1994] for the B1 domain of a different SpG construct using titration calorimetry, which gave a Kd value of 333 nM. Although these values are slightly different, the slight increase in Kd found for the construct described in this thesis could indicate that the introduction of a 6x His-tag at the N-terminal end of the protein has slightly decreased the binding affinity of SpG for Fc**.** Similarly, the stoichiometry of binding for the interaction between SpG and Fc compares well for the two different experiments. Data from the ITC experiments gave a stoichiometry of 1.70 ± 0.05 equivalents of SpG to 1 equivalent of Fc, whereas the stoichiometry of binding from the fluorescence titrations was 2.1 ± 0.4 mol SpG: mol Fc.

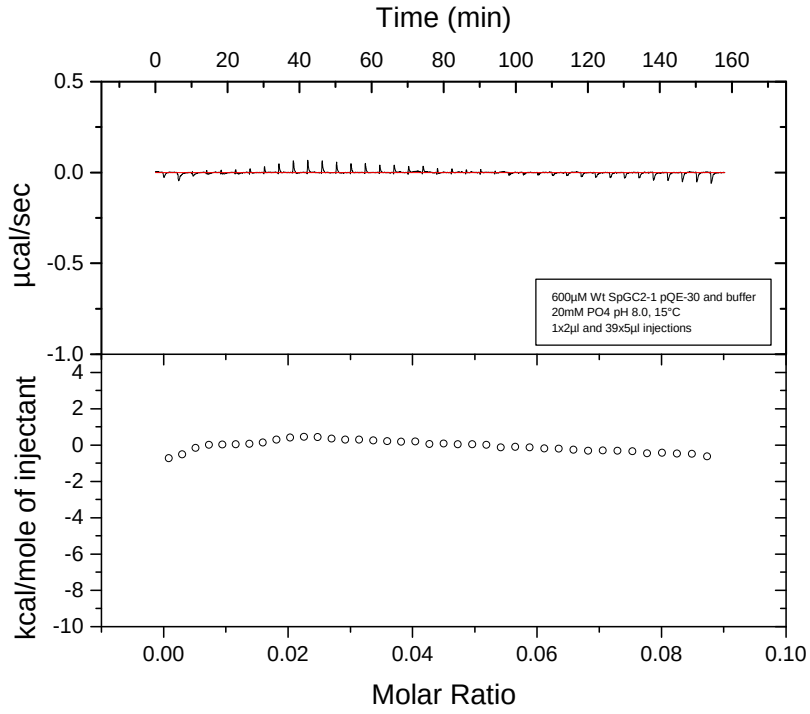


Figure 6.2 – Binding isotherm of the interaction between Wt D-SpG$_{C2}$-1 and buffer

The analysis of the binding interaction between Wt D-SpG$_{C2}$-1 and buffer measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into phosphate buffer using injections of 1 x 2 µl and 39 x 5 µl.

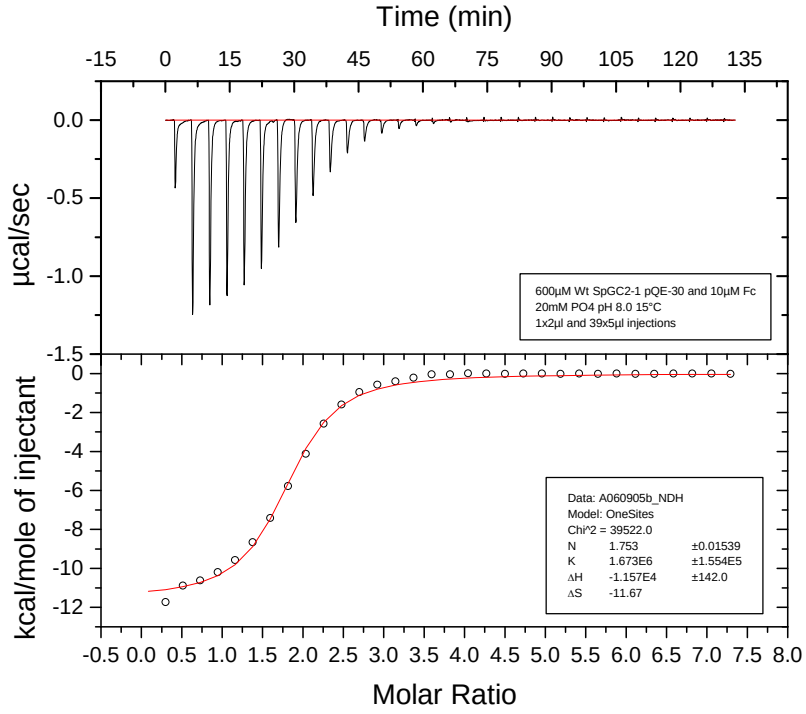


Figure 6.3 – Binding isotherm of the interaction between Wt D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between Wt D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 29 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0.

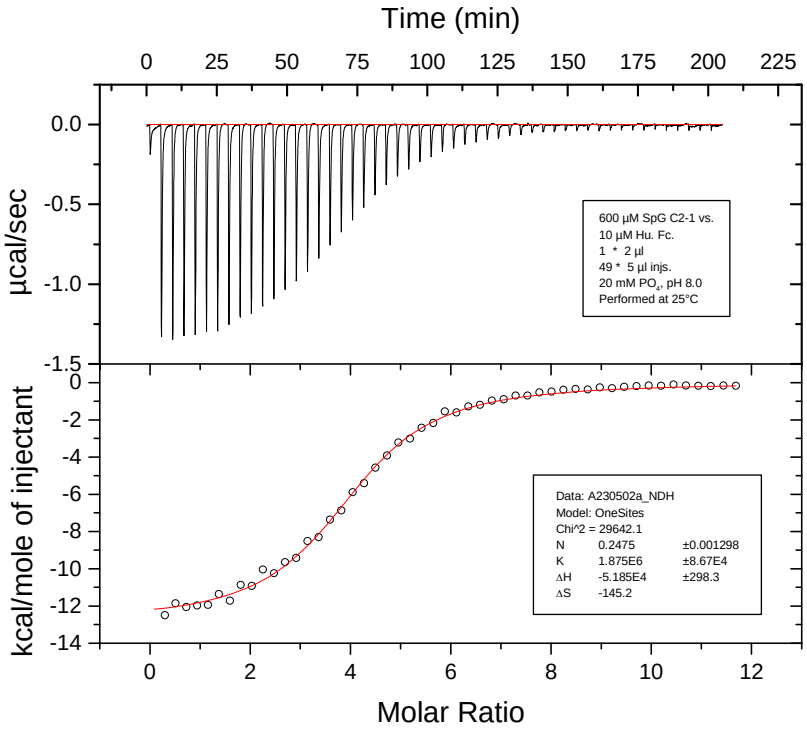


Figure 6.4 – Binding isotherm of the interaction between Wt D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between Wt D-SpG$_{C2}$-1 and Fc measured by ITC at 25 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 49 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0.

6.5 – Effect of mutagenesis on the stability of the SpG•Fc complex

Mutant constructs of SpG that have substitutions of residues that may be critical for the binding interaction with Fc have been produced to allow the contribution of these residues to the biological function of SpG to be evaluated. This is important, as it may give a complete overview of the binding interface and should greatly add to the understanding of the binding interaction. ITC is well placed to characterise the binding activities of these mutant protein constructs because it does not require the development of a new assay, and as well as being precise it does not demand catalytically active protein to generate a product to follow the reaction. Thus, titrations were performed as described above on all D-SpG$_{C2}$-1 mutant constructs.

D-SpG$_{C2}$-1	Kd ± std dev. n≥2	Δ*H* (kJ.mol^{-1}) ± std dev. n≥2	Δ*S* (J.mol^{-1}.deg^{-1}) ± std dev. n≥2	Δ*G* (kJ.mol^{-1}) ± std dev. n≥2	N ± std dev. n≥2
Wt	677 nM ± 154.1	-50.02 ± 5.15	-56.07 ± 17.15	-33.93 ± 0.48	1.7 ± 0.1
E14W	856 nM ± 76.4	-39.14 ± 0.81	-19.67 ± 2.00	-33.51 ± 0.24	2.40 ± 0.0
E14W,W42F	4.5 µM ± 1.4	-15.53 ± 3.33	48.54 ±14.23	-29.55 ±0.76	0.8 ± 0.1
E26A	ND	ND	ND	ND	ND
C56A	290 nM ± 4.2	-47.26 ± 3.64	-38.91 ± 12.55	-36.08 ± 0.02	1.7 ± 0.1
Q31A	664 nM ± 43.1	-44.65 ± 5.98	-36.82 ± 21.34	-34.10 ± 0.19	1.9 ±0.1
Q31E	4.5 µM ±0.7	-46.59 ± 0.74	-58.58 ± 1.25	-29.69 ± 0.37	1.8 ± 0.1
Q31H	651 nM ± 19.1	-51.51 ±2.01	-60.25 ± 7.11	-34.14 ± 0.07	2.0 ±0.0
N34A	7.0 µM ± 0.6	-11.77 ± 0.84	57.32 ± 3.35	-28.28 ± 0.17	1.6 ± 0.4
N34D	8.5 µM ± 3.7	-30.66 ± 3.39	-7.53 ± 15.90	-28.32 ± 3.83	3.3 ± 0.2
E41A	517 nM ± 142.8	-28.18 ± 1.77	-19.67 ±7.53	-33.76 ± 0.42	2.1 ± 0.0
E41Q	320 nM ± 24.0	-30.24 ± 0.89	-24.27 ± 2.93	-34.89 ± 0.81	2.1 ±0.1
W42F	ND	ND	ND	ND	ND
W42Y	40 µM ±12.0	-40.92 ± 5.92	-57.32 ± 17.99	-24.42 ± 0.79	0.8 ± 0.1

Table 6.1 – Summary of the thermodynamic parameters obtained from titrating Wt and mutant D-SpG$_{C2}$-1 constructs into Fc

6.6 - Determination of the thermodynamic constants for Fc mutants

The thermodynamic data obtained for all Fc mutants is given in figures 6.5 to 6.17 and the thermodynamic parameters summarised in table 6.1. In general, the Kd values for the mutants E14W, Q31A, Q31H, and E41A (figures 6.5 to 6.8) show very little difference, within error, from the Kd of 677 nM that was determined for wild-type SpG interacting with Fc. This indicates that the replacement of these residues thought to be involved in the binding of SpG and Fc does not result in a reduced binding affinity. However, although the overall Kd does not change greatly, the thermodynamic constants of the reaction contributions are different. The enthalpic contributions of these reactions are around -40 to -50 $kJ.mol^{-1}$ for the Wt, E14W, Q31A and Q31H proteins, but for the E41A protein, it is lower at -28.18 $kJ.mol^{-1}$. However, compensatory changes in the entropic contribution result in very similar Kd values. Thus, the overall conclusion is that these residues may contact the Fc, but have a minor involvement in stabilising interactions with Fc. It is of course also possible that the substitutions made have compensated for the loss in enthalpic changes. Hydrophobic changes may for example have caused some water molecules to be released thus increasing entropy in relative terms (e.g. from -55.95 $cal.mol^{-1}deg^{-1}$ to -19.70 $cal.mol^{-1}deg^{-1}$). Residue E41 has been implicated in Fc binding by virtue of the fact that in the C3 domain (which has a higher affinity for human IgG than C1), a valine is found at this position. It is therefore interesting that for the mutant E41Q (figure 6.9), the removal of the negative charge and its replacement by a more hydrophobic amide side chain causes an increase in affinity (Kd = 320 nM). This, compared with that for the E41A mutant (517 nM), suggests that the amide side chain may play an important role in the interaction and that the loss of the negative charge may not, by itself, be sufficient to increase affinity.

The binding affinities of mutants Q31A and Q31H with Fc do not change when observed by ITC, even though the replacement of the glutamine residue with an alanine residue removes the hydrogen bond, and hence part of the stabilising hydrogen network with Fc residue Gln438. In all cases the stoichiometry of binding remains close to the expected value of 2 due to the presence of two identical binding sites for SpG on Fc. However, the mutation Q31E also involves the loss of the hydrogen bond between the two proteins, and presents a negative charge. Results obtained show that this mutant, and the mutants N34A, N34D and W42Y, do show greatly reduced affinities for Fc. Figures 6.10 to 6.13 present

typical binding isotherms for these D-SpG$_{C2}$-1 mutants respectively, under the same conditions. Kd values were 4.5 μM, 7 μM, 8.5 μM and 40 μM, respectively. Titrations involving the mutant N34D gave irreproducible analyses, possibly due to some instability in this protein, therefore the standard deviation is high and the Kd value and stoichiometry should be taken as estimates only.

In general, these mutants show lower enthalpic contributions suggesting that fewer bonds are being made although in some cases (e.g. Q31E and W42Y) there is a higher loss of entropy, which is unfavourable for binding. The results emphasise the importance of Q31E, N34A, N34D and W42Y residues in the binding of SpG with Fc. In most cases the stoichiometry calculated was close to 2 mol SpG to 1 mol of Fc, however the difficulty encountered in fitting the data to a binding profile where Kd is very large (e.g. N34D and W42Y) results in a what is believed to be a poor estimate for the stoichiometry (see table 6.2).

Figures 6.14 and 6.15 present typical binding isotherms obtained for each of the mutants E26A and W42F D-SpG$_{C2}$-1 titrated into Fc. No heat changes were observed for either of these mutants, indicating very large losses or total loss of binding activity. The latter data support previous data on the mutant W42F obtained by Walker [1994], who implicated the importance of this tryptophan residue in Fc binding by ELISA studies. Clearly E26 has an important role in binding Fc.

Mutants E14W (see figure 6.5) and E14W-W42F D-SpG$_{C2}$-1 (see figure 6.16) were specifically constructed to study the interaction between SpG and Fab. By using the W42F D-SpG$_{C2}$-1 construct as a template, it was planned that the re-introduction of a tryptophan residue at the alternative position of residue E14, would allow the effect of Fab mutations to be probed by fluorescence studies without any background signal changes from position 42 (normally a tryptophan). To ensure that the E14W-W42F D-SpG$_{C2}$-1 construct had not been affected by the introduction of tryptophan at an alternative position, the ability of this mutant to bind Fc, and that of E14W D-SpG$_{C2}$-1 (which contains two tryptophan residues W42 and E14W) was also studied by ITC as controls. Figures 6.5 and 6.16 present typical binding isotherms that were obtained for mutants E14W and E14W-W42F D-SpG$_{C2}$-1, respectively, titrated into Fc.

Results obtained show that E14W-W42F D-SpG$_{C2}$-1 has a much higher Kd value (4.5 μM) and thus a reduced binding affinity for Fc compared to values obtained

for wild-type SpG (677 nM ±154 nM) and E14W (856 nM ± 76 nM), which contains two tryptophan residues, and binds with a similar Kd value to that obtained for wild-type SpG. These results confirm that residue W42 plays an important part in the binding of SpG with Fc. However, due to the lack of binding exhibited by construct W42F, it would seem that the addition of tryptophan E14W in the construct may weakly compensate for the absence of W42. Again the low affinity resulted in a low estimate of the binding stoichiometry.

Finally, figure 6.17 shows the binding isotherm for the mutant C56A (made to remove the carboxyl terminal cysteine), as its substitution removes any possibility of formation of dimers between the SpG domains and the need for the presence of a reducing agent. C56A was found to have a similar Kd (290 μM) to that of the wild-type protein, confirming that the presence of the cysteine has little or no effect on the affinity for Fc.

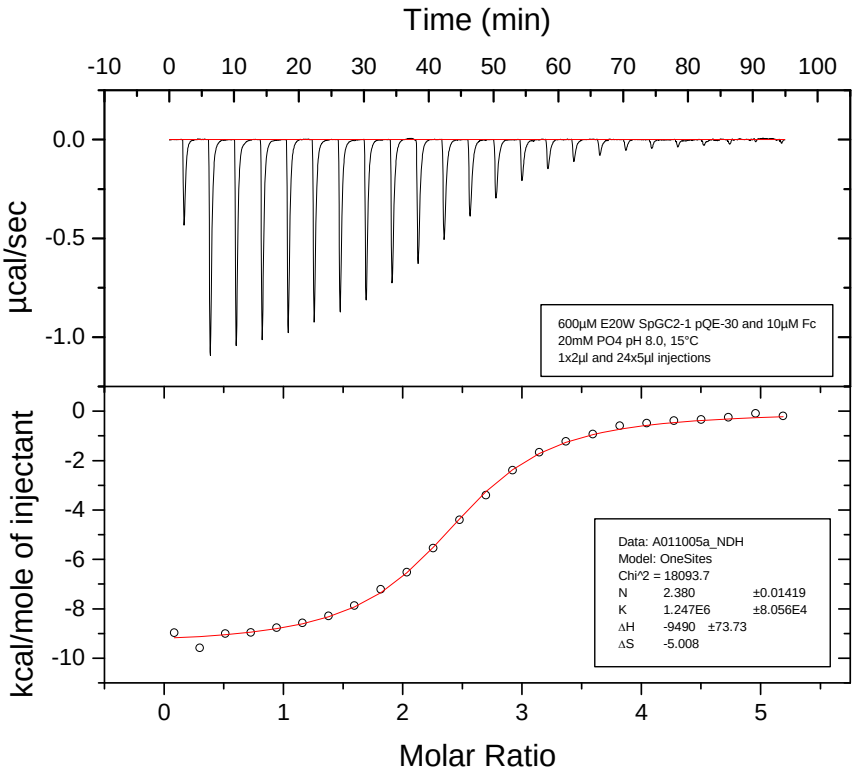


Figure 6.5 – Binding isotherm of the interaction between E14W D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between E14W D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 μM SpG was titrated into 10 μM Fc using injections of 1 x 2 μl and 24 x 5 μl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0.

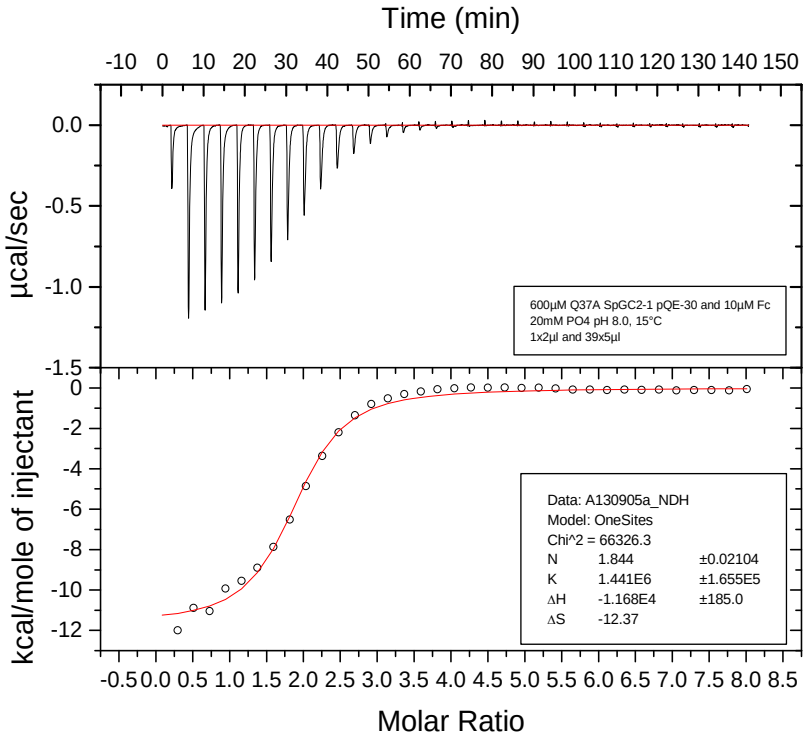


Figure 6.6 - Binding isotherm of the interaction between Q31A D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between Q31A D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 39 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0.

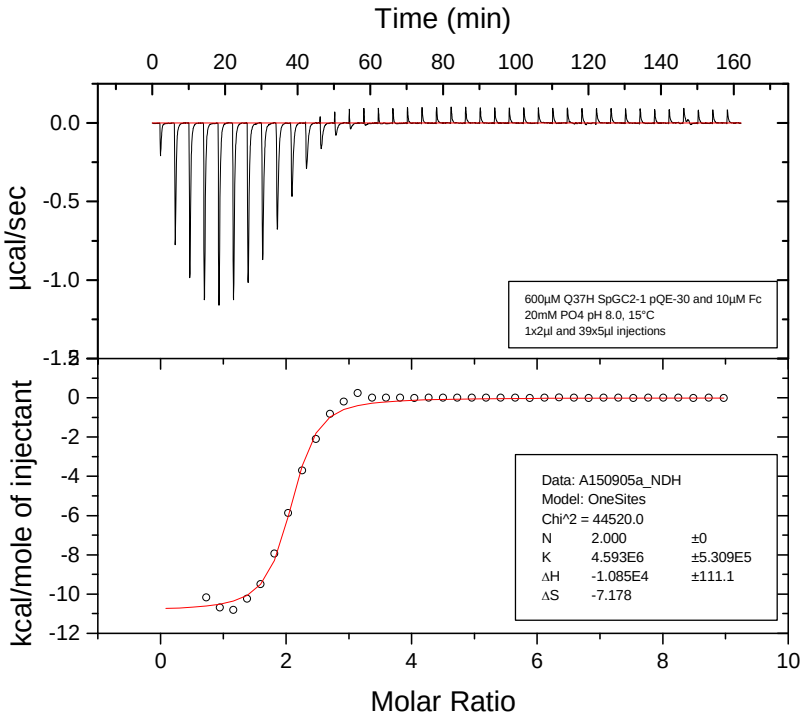


Figure 6.7 - Binding isotherm of the interaction between Q31H D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between Q31H D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 39 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0.

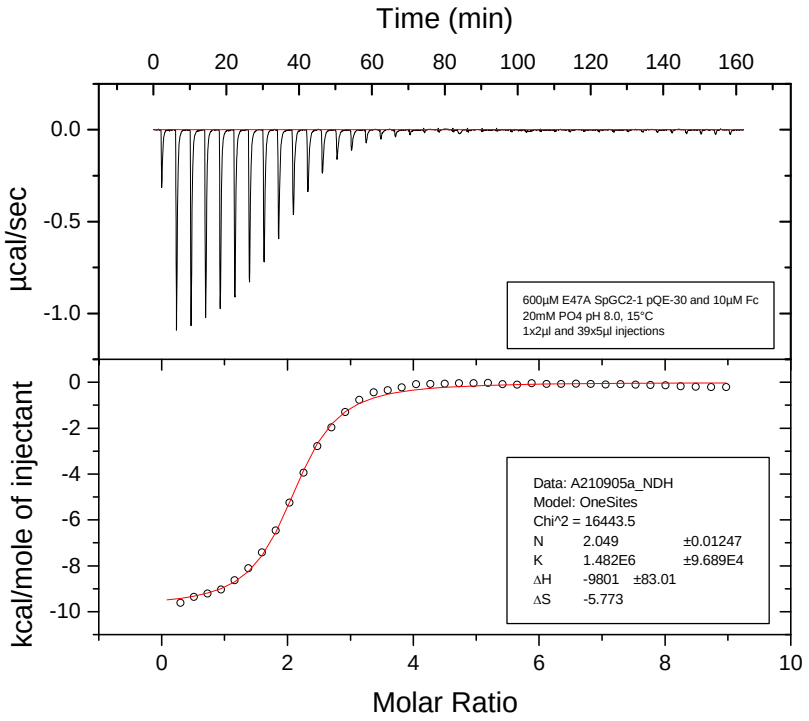


Figure 6.8 - Binding isotherm of the interaction between E41A D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between E41A D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 39 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0.

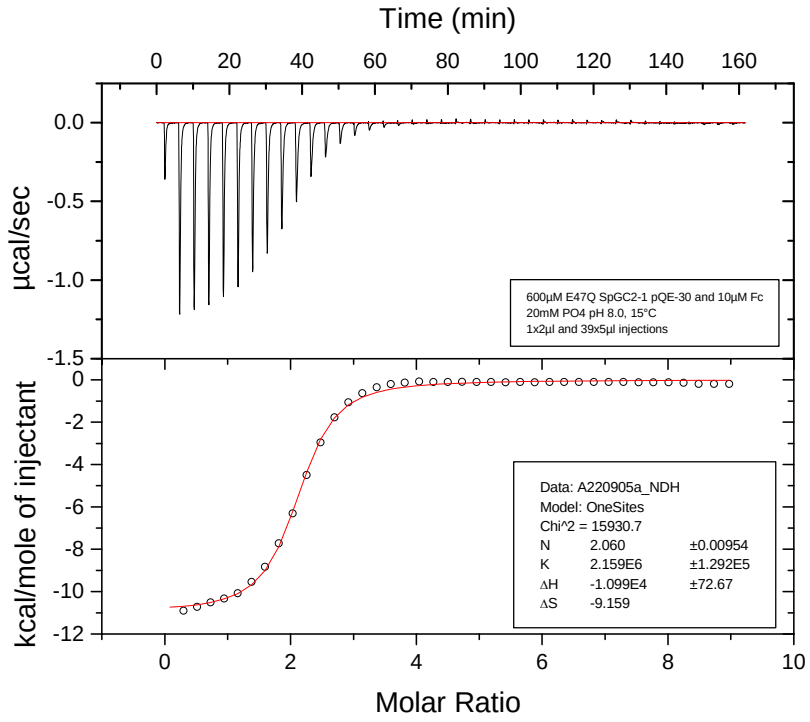


Figure 6.9 - Binding isotherm of the interaction between E41Q D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between E41Q D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 39 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0.

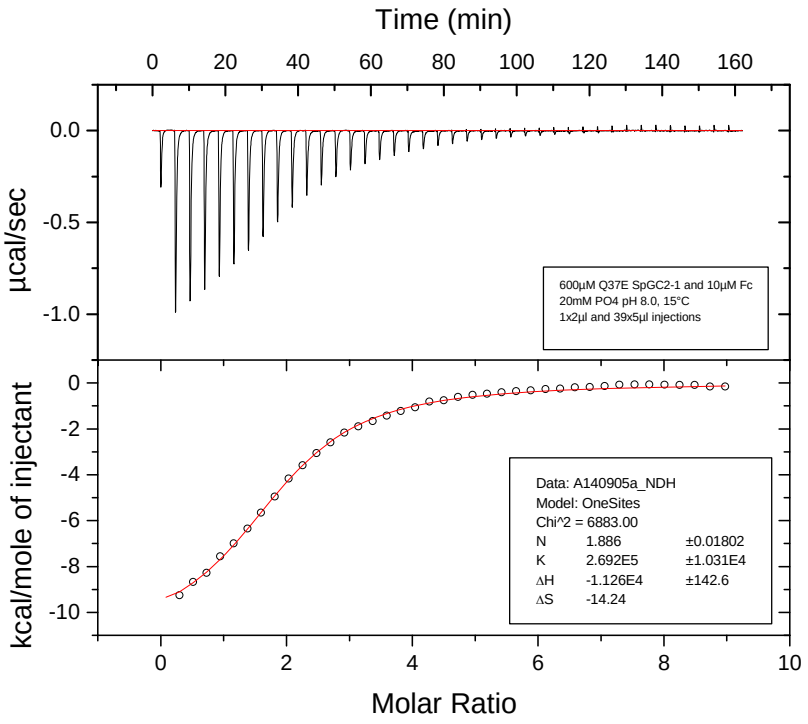


Figure 6.10 - Binding isotherm of the interaction between Q31E D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between Q31E D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 39 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0.

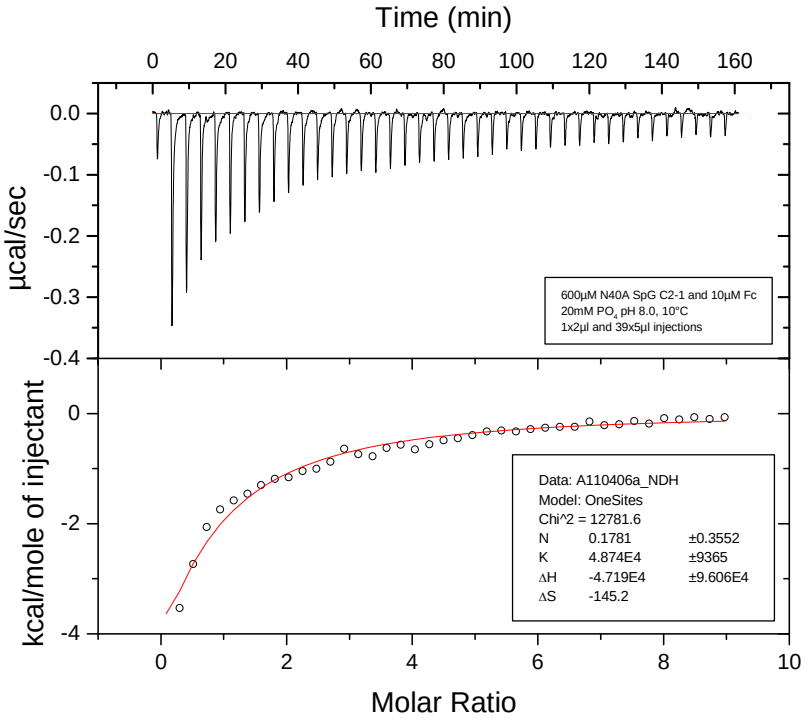


Figure 6.11 - Binding isotherm of the interaction between N34A D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between N34A D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 39 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0.

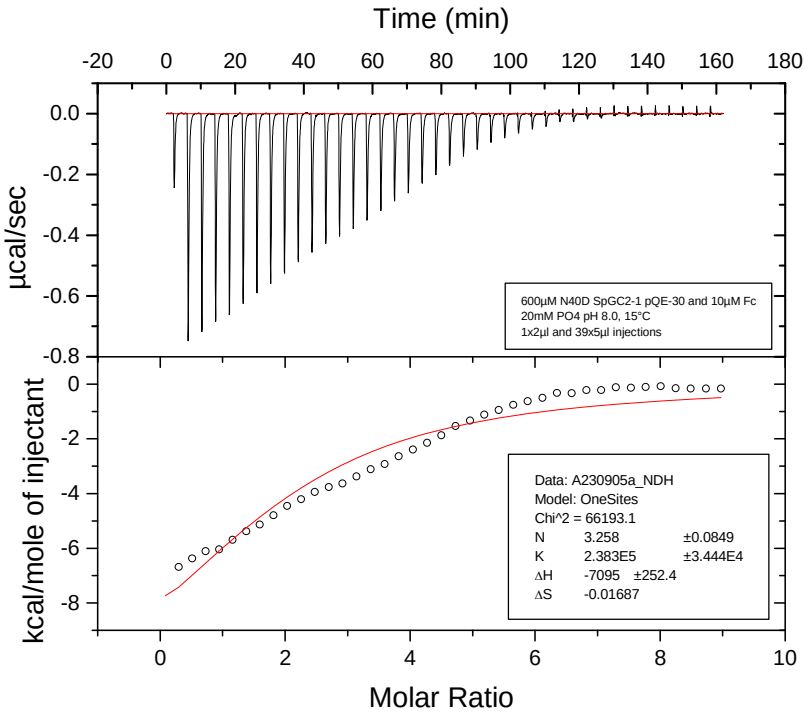


Figure 6.12 - Binding isotherm of the interaction between N34D D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between N34D D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 39 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0.

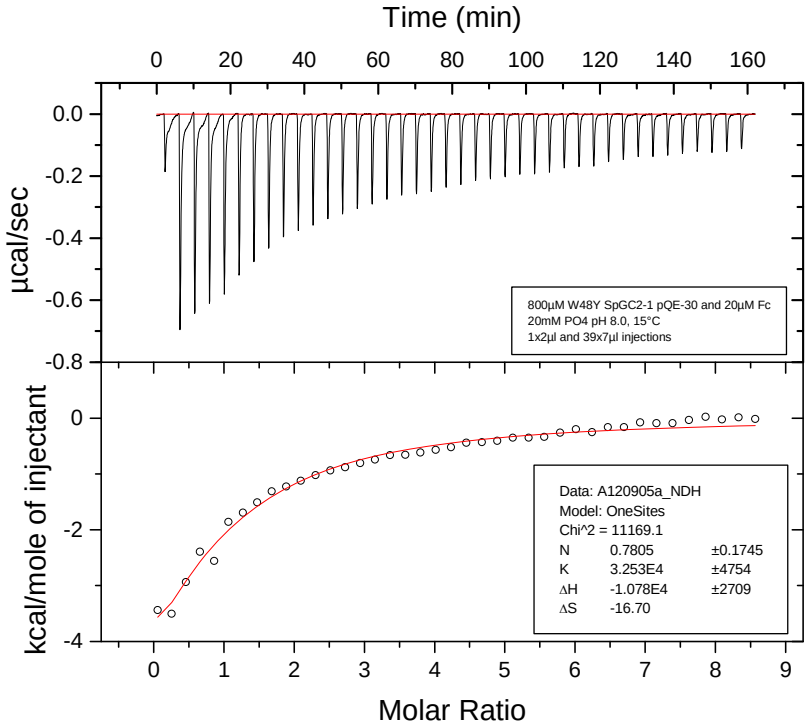


Figure 6.13 - Binding isotherm of the interaction between W42Y D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between W42Y D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 39 x 7 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0

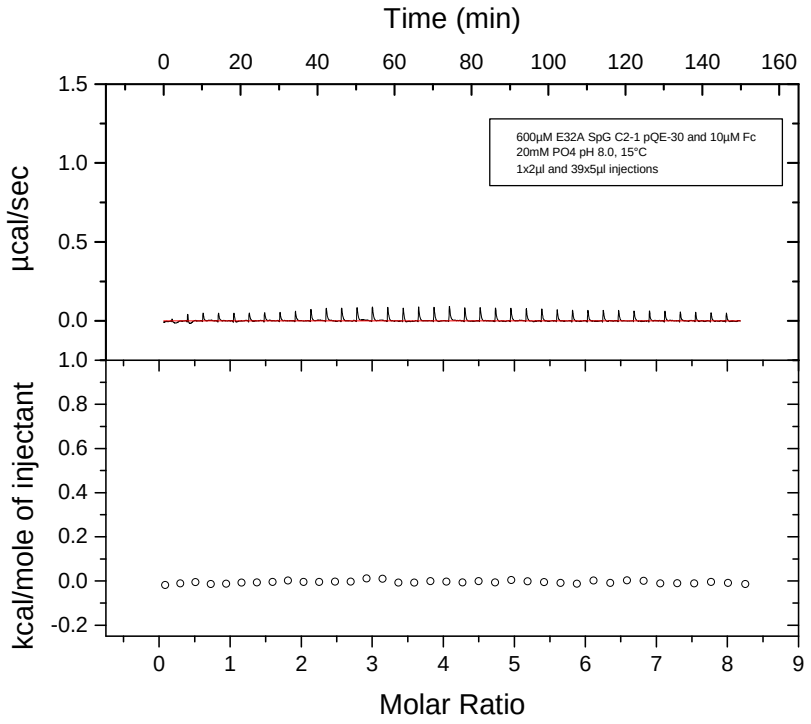


Figure 6.14 – Binding isotherm of the interaction between E26A D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between E26A D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 39 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0

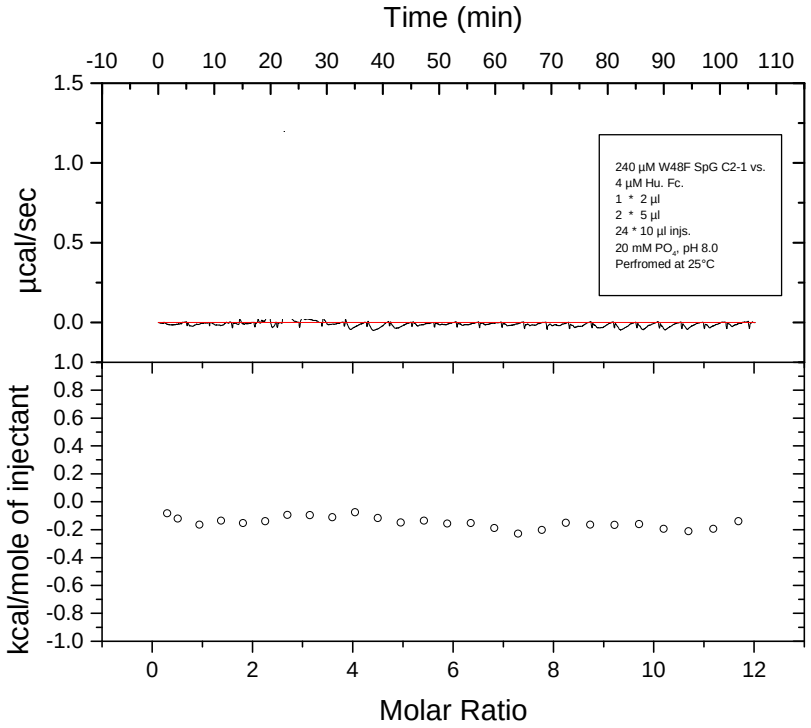


Figure 6.15 - Binding isotherm of the interaction between W42F D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between* W42F *D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 4 µM Fc using injections of 1 x 2 µl, 2 x 5µl and 24 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0

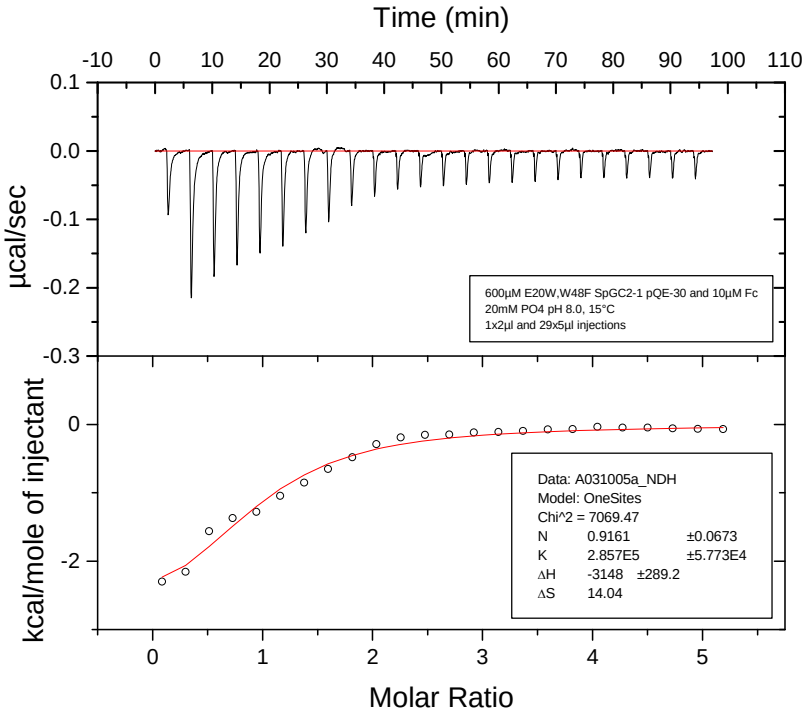


Figure 6.16 - Binding isotherm of the interaction between E14W-W42F D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between* E14W-W42F *D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 29 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0

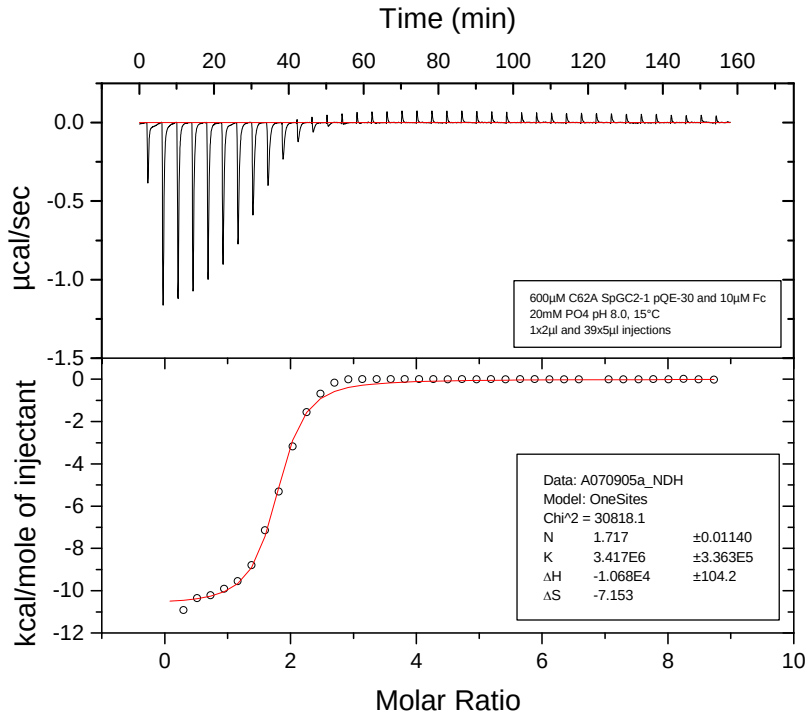


Figure 6.17 - Binding isotherm of the interaction between C56A D-SpG$_{C2}$-1 and Fc

The analysis of the binding interaction between C56A D-SpG$_{C2}$-1 and Fc measured by ITC at 15 °C is presented. A solution of 600 µM SpG was titrated into 10 µM Fc using injections of 1 x 2 µl and 39 x 5 µl. All titrations were carried out in 20 mM potassium phosphate buffer, pH 8.0

6.7 - The effect of temperature on the binding interaction between Wt SpG and Fc

By performing similar equilibrium experiments over a range of temperatures it is possible to examine the effect of temperature on the reaction and derive values of Δ*H*° and Δ*S*° from a van't Hoff plot. The equation for the plot used was derived by van't Hoff by combining equation 6.2 with equation 6.3.

$$\Delta G = -RT\ln Ka = -\Delta H - T\Delta S \qquad (Equation\ 6.4)$$

This can be re-arranged to give the van't Hoff equation below:

$$\ln Ka = -\Delta H°/RT + \Delta S°/R \qquad (Equation\ 6.5)$$

Thus a plot of lnKa against 1/*T* will yield a line of slope -Δ*H*°/R and an intercept of Δ*S*°/R. These values will therefore be derived not from a single experimental point, but from the average of all experiments carried out at different temperatures. The range of temperatures able to be used will depend upon the stabilities of the proteins used, in this case SpG and Fc. Both of these are stable up to 35 °C and therefore a temperature range of 10 °C to 35 °C was chosen with experiments being carried out at 5 °C intervals.

For these experiments wild-type SpG was titrated into Fc as described above, except that the solutions were equilibrated at different pre-determined temperatures for the titration. The data given in table 6.2 show the temperatures in °C and the values of the average Ka determined from each titration. These data are used to generate the van't Hoff plot seen in figure 6.18.

From the data it is clear that the affinity of SpG for Fc decreases as temperature increases from 10 °C to 35 °C. This is as expected since increased temperature raises the level of molecular motion and destabilises the complex. This also suggests that the complex is stabilised by hydrogen bonds and not stabilised by hydrophobic or ionic interactions alone, since an increase in temperature should increase the strength of these latter interactions. This follows because a hydrophobic interaction is normally regarded as resulting in an increase in entropy due to the exclusion of water molecules from the interface. Thus the entropic force (*T*Δ*S*) would increase with *T*. Similarly, ionic interactions are described by Coulomb's law in which the attractive force between two oppositely charged particles is reciprocally related to the Dielectric constant (D) of the

solvent. In this case the solvent is water whose dielectric constant decreases as temperature rises.

Temperature (°C)	Kd (μM) ± std dev. n≥2
10	2.233 ± 0.025
15	1.237 ± 0.039
20	0.699 ± 0.0741
25	0.611 ± 0.056
30	0.482 ± 0.025
35	0.325 ± 0.031

Table 6.2 – Summary of the average dissociation constant (Kd) at different temperatures by titrating Wt D-SpG$_{C2}$-1 into Fc

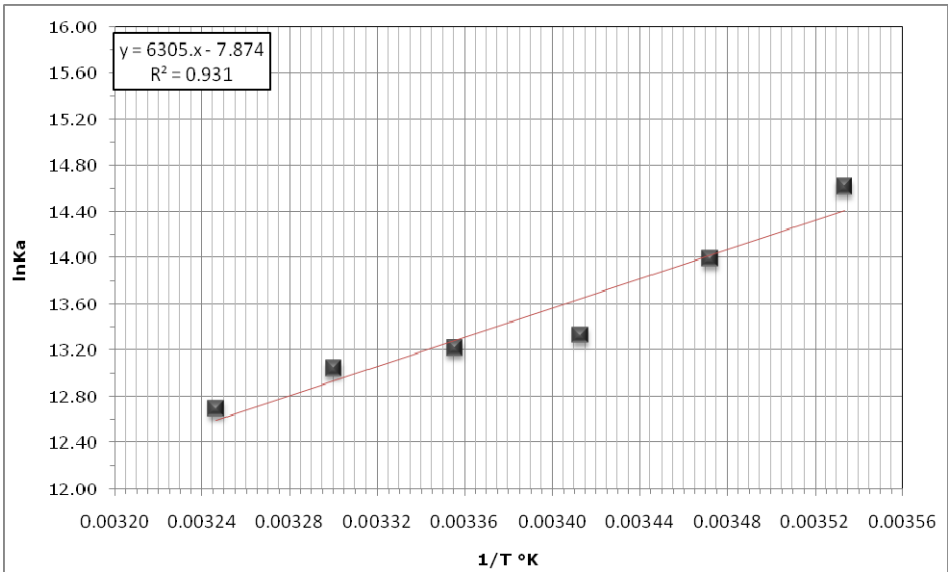


Figure 6.18 - A van't Hoff plot of data from Table 6.2

This graphical determination represents a van't Hoff plot of the effect of varying temperatures on the formation of Wt D-SpG$_{C2}$-1 and Fc complex. All experiments were observed by ITC in triplicate and carried out in 20 mM phosphate buffer, pH 8.0.

The van't Hoff plot gives ΔH = -52.19 kJ.mol^{-1}. Similarly, the intercept results in a value for ΔS of 65.18 J.mol^{-1}.deg^{-1}. From these values the calculated ΔG at 15 °C (288K) is -33.42 kJ.mol^{-1}, which yields a Kd for the reaction at this temperature of around 1150 nM in close agreement with that obtained from a single titration at 15 °C (677 ± 154 nM). Thus, the data obtained over the temperature range appear to confirm that the plot is linear (R = 0.931) and that the two methods for estimating the thermodynamic parameters for the reaction are comparable.

6.8 - Discussion

The results of the experiments above show that the affinity of SpG for Fc can be dramatically altered by substitution of some key residues. It is clear that W42 is essential for binding activity as previously found by Walker [Walker, 1994]. However, it appears that some restoration of activity is brought about by the placement of a tryptophan at residue 14 and a Kd of 4.5 μM was observed. In practice this may be a useful protein for affinity chromatographic applications as a Kd in the micromolar range is ideal for affinity ligands where easy elution of the bound target protein (in this case IgG or Fc) is required (personal communication Dr. Jim Pearson, Affinity Chromatography Ltd).

These studies also show that in addition to the importance of W42, residues E26, Q31 and N34 play an important role in the binding. In the case of E26, substitution by alanine results in no significant change in the Kd value of the complex, but substitution by a glutamate residue increases Kd by approximately 6-fold, demonstrating that a negative charge at this position is detrimental to binding. Residue N34 is clearly more important for binding since the N34A substitution increases Kd by about 10-fold and the N34D substitution effectively eliminates all apparent binding activity.

CHAPTER 7.0

EQUILIBRIUM & PRE-EQUILIBRIUM BINDING STUDIES OF SPG-FAB BINDING INTERACTION

7.1 – Introduction

The Fab region is composed of one constant and one variable domain from each heavy and light chain of the antibody, as shown in figure 7.1. The variable domains from the heavy and light chains provide a framework that supports three regions of hypervariability in sequence called the complementarity-determining regions (CDRs). These are located in the loops that connect the beta strands of the variable Ig domains. When the V_L and V_H chains pair up in the antibody structure, the CDR-containing loops are brought together to form the antigen binding site at each amino terminal end of the Fab fragment [Hurst *et al.*, 2007]. They are responsible for the specificity and affinity of the epitope on their specific antigens.

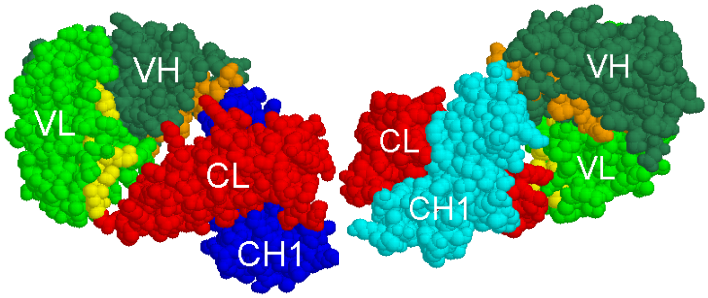


Figure 7.1 – <u>An adapted illustration of two Fab fragments</u>

This model presents two Fab fragments, which are composed of one constant (C) and variable (V) domain from each heavy (H) and light (L) chain of the antibody. At the bottom of each Fab fragment is the hinge region, where each heavy chain extends into the Fc fragment.

7.2 – Interaction of SpG with Fab

Single Ig-binding domains of SpG have been shown to bind to Fab by forming an anti-parallel interaction with the last β-strand of the C_H1 domain of Fab [Derrick and Wigley, 1992; Lian *et al.*, 1994; Derrick and Wigley, 1994; Sauer-Eriksson *et al.*, 1995]. SpG's binding site for the interaction with Fab lies predominantly within β-strand 2, but also involves a minor region at the C-terminal end of the α-helix [Derrick and Wigley, 1994]. The amino acid residues that form the major contact regions have been identified, and the site of interaction spans residues Lys9 to Thr16 from β-strand 2 on SpG and Ser209 to Lys216 from the 7th β-strand of the C_H1 domain of Fab. In addition, residues Tyr32 to Gly37 from the C-terminal end of the α-helix of SpG and Pro125 to Tyr129 on the 1st β-strand of the C_H1 domain of Fab constitute the minor contact region [Derrick and Wigley, 1994].

These interfacial interactions between Fab and SpG are stabilised through a network of hydrogen bonds and salt links, involving mainly charged and polar

residues contributed by the last β-strand of the C_H1 domain to the second β-strand in SpG. The complex is further stabilised by the association of exposed non-polar residues from Fab and SpG, providing a continuous hydrophobic core, which is shielded from solvent [Derrick and Wigley, 1992].

As previously mentioned, SpG possesses a unique tryptophan residue at position 42, located on β-strand 3, which is known for its important role in Fc binding. Its substitution by phenylalanine reduces the binding interaction of SpG with Fc approximately 300-fold [Walker *et al.*, 1995]. Therefore, by using this W42F D-SpG_{C2}-1 construct, IgG Fab binding can be studied without interference from any Fc interactions. However, the removal of the tryptophan in the W42F mutation would create a non-fluorescent protein, resulting in Fab binding studies using fluorimetry being difficult or impossible. Therefore, a tryptophan residue was re-introduced at an alternative position (E14W, see chapter 3.6.1) to allow the effect of Fab mutations to be probed in fluorescence studies.

By incorporating the E14W mutant close to the proposed Fab binding site it was anticipated that this tryptophan residue would act as a reporter group solely for SpG binding to the C_H1 domain of Fab. Therefore, the mutant W42F D-SpG_{C2}-1 construct was used to make the E14W mutant and the resulting double mutant, E14W-W42F D-SpG_{C2}-1 *pQE*-30, was then employed as a template for subsequent mutants. The single mutant, E14W, was also employed as a control for all experiments undertaken in this study.

Six residues were selected to be mutated and used to study and characterise the binding affinity of Fab to SpG. A selection of residues from both the major and minor contact regions, namely residues Thr10, Glu14, Thr15, Thr16, Tyr32 and Asn36 were chosen for their side-chain involvement in the interaction with Fab. This complex is stabilised through a network of hydrogen bonds, and the importance and effect of these bonds to the stability and nature of the complex by removing these key residues was the prime aim of the studies.

7.3 – Determining the emission spectra of D-SpG_{C2}-1 and Fab complex

In order to carry out binding studies between SpG and Fab using fluorescence, a measureable change in fluorescence must occur upon complex formation. To establish whether or not a spectral change does occur upon binding, the emission spectra for E14W-W42F D-SpG_{C2}-1 and Fab, both in free solution and in complex, were determined and is presented in figure 7.2.

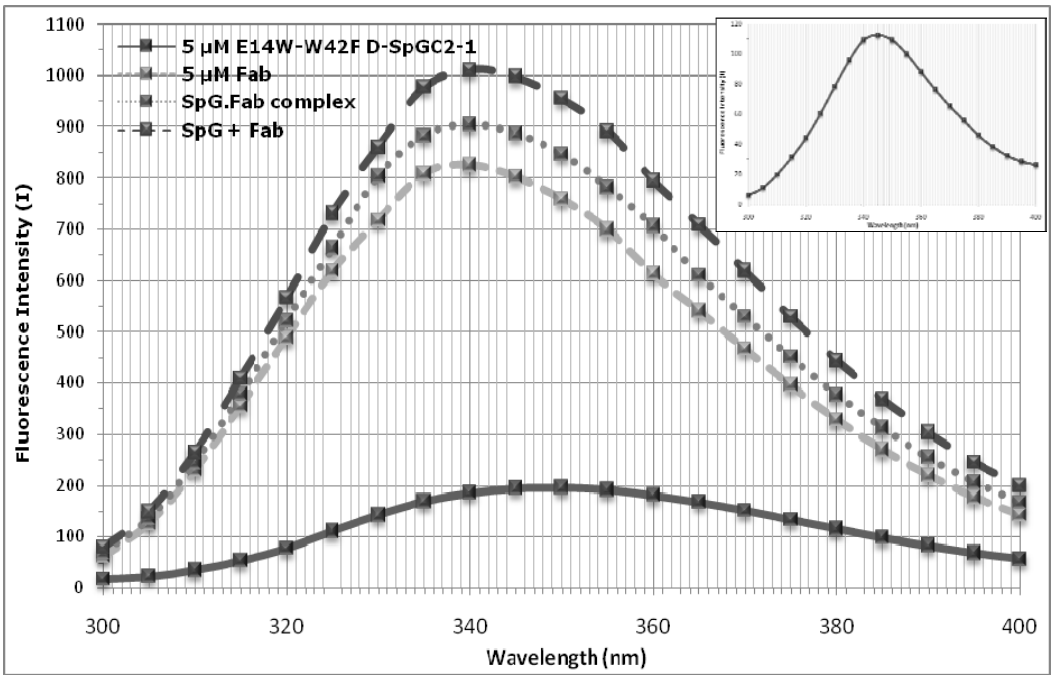


Figure 7.2 - Emission spectra for E14W-W42F D-SpG$_{C2}$-1 and Fab in complex

The figure shows the emission spectra obtained for 200 nM E14W-W42F D-SpG$_{C2}$-1, 200 nM Fab and the SpG•Fab complex. Protein samples were excited at 280 nm and fluorescence emissions were measured between 300 and 400 nm. All spectra were measured in triplicate and corrected for the buffer fluorescence. The inset shows the fluorescence change upon complex formation, calculated by subtracting the observed fluorescence intensity of the complex from the theoretical fluorescence intensity of the complex.

The spectra presented show that fluorescence intensity obtained for E14W-W42F D-SpG$_{C2}$-1 displays an emission maximum at 345 nm, whereas the peak of the Fab spectra and the complex spectra is 340 nm. The spectra also show that the fluorescence intensity obtained for the SpG•Fab in complex is slightly less than the theoretical fluorescence of the complex, being calculated from the sum of the fluorescence spectra of the two proteins in free solution. This quench in fluorescence seems to occur upon complex formation and shows that the formation of the SpG•Fab complex can be monitored fluorimetrically. The quench is represented as a fluorescence difference spectrum, and is presented in the inset figure in figure 7.2. The inset shows that the difference spectrum also has a maximum at 345 nm, which suggests that tryptophan 14, like tryptophan 42, is in a solvent exposed position, and is being quenched on complex formation.

7.4 – Equilibrium binding studies

Equilibrium binding studies were carried out to provide valuable information on the binding affinity and number of binding sites of the interaction between SpG and Fab. In the following sections, studies are described that were aimed at the determination of the Kd for the binding reaction between SpG and Fab.

7.4.1 - Fluorescence titration

The binding interaction between SpG and Fab has been studied through fluorescence titrations using the change in fluorescence that occurs upon complex formation. Fluorescence emissions at 340 nm were measured, using excitation wavelengths of 280 nm to study total fluorescence, and 295 nm to study tryptophan only fluorescence. 5 µl samples of 100 µM D-SpG$_{C2}$-1 were titrated into 2 ml samples of 1 µM Fab. Data obtained were corrected for dilution and the inner filter effect, and further analysed through a Klotz plot, employing equation 5.5 as defined in Chapter 5.

Figure 7.3 presents a typical saturation curve for the binding of E14W-W42F D-SpG$_{C2}$-1 to Fab and the subsequent analysis through a modified Klotz plot. The curve obtained for E14W-W42F D-SpG$_{C2}$-1 binding to Fab at 15 °C, pH 8.0 was able to be described by a single Kd value of 3.33 ± 1.40 µM. The stoichiometry of binding from the intercept $[E]_0$ was 1.17 ± 0.18, confirming that Fab has one binding site for E14W-W42F D-SpG$_{C2}$-1. This is expected as Fab fragments have one heavy C_H1 domain with a potential binding site for SpG.

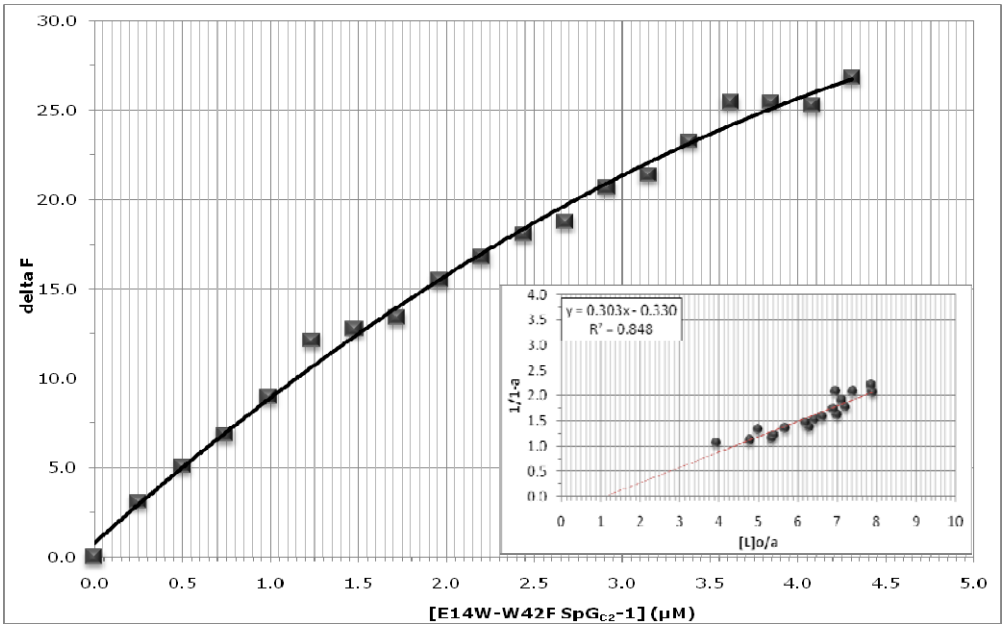


Figure 7.3 - <u>Saturation curve for the titration of E14W-W42F D-SpG$_{C2}$-1 into Fab</u>

The figure shows the saturation curve obtained when samples of 100 µM E14W-W42F D-SpG$_{C2}$-1 were titrated into a 2 ml solution of 1 µM Fab in 20 mM potassium phosphate buffer, pH 8.0 at 15 °C. Each point on the titration curve was calculated by subtracting fluorescence values observed upon titration of E14W-W42F D-SpG$_{C2}$-1 into Fab, from the corresponding values for the titration of E14W-W42F D-SpG$_{C2}$-1 into buffer. The inset shows the Klotz plot of the data (n = 3).

7.4.2 – Isothermal Titration Calorimetry (ITC)

7.4.2.1 – Determination of Wt SpG•Fab binding by ITC

Analysis of the binding of Fab to SpG by ITC is typically carried out by injecting a small volume of SpG at high concentration into a large volume of Fab fragment at a lower concentration. Therefore, a solution of 400 μM Wt D-SpG$_{C2}$-1 was titrated into 20 μM Fab using injections of 1 × 2 μl and between 20 and 30 × 7 μl. This method aims to produce a binding isotherm that is sufficient to be analysed, with a minimum -0.3 μcal/sec value.

Figure 7.4 presents a binding isotherm of Wt D-SpG$_{C2}$-1 titrated into 20 mM phosphate buffer at pH 8.0, and figures 7.5, 7.6 and 7.7 present typical binding isotherms of wild-type, E14W-W42F and E14W D-SpG$_{C2}$-1, respectively, titrated into Fab at 15 °C.

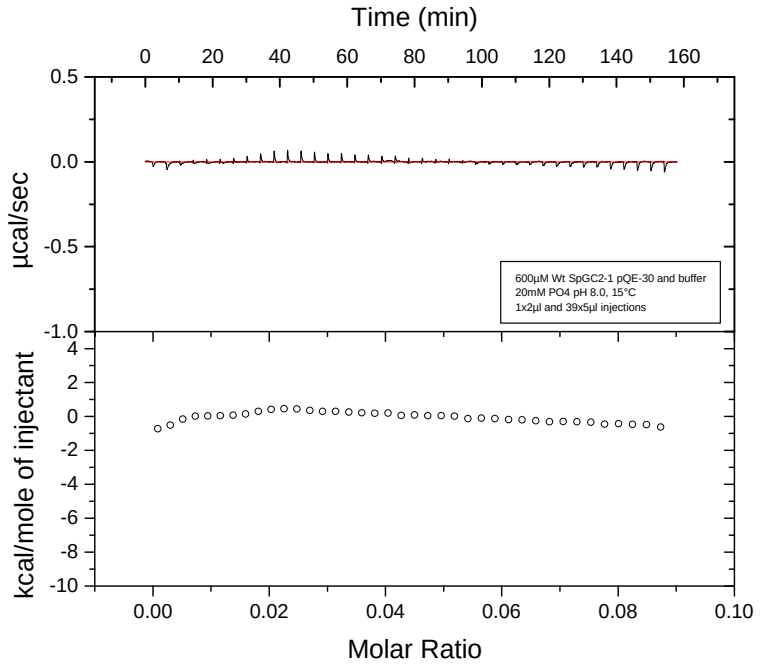


Figure 7.4 – <u>Binding isotherm of the interaction between Wt D-SpG$_{C2}$-1 and buffer observed by ITC</u>

The analysis of the binding interaction between Wt D-SpG$_{C2}$-1 and buffer measured by ITC at 15 °C is presented. A solution of 600 μM SpG was titrated into phosphate buffer using injections of 1 x 2 μl and 39 x 5 μl.

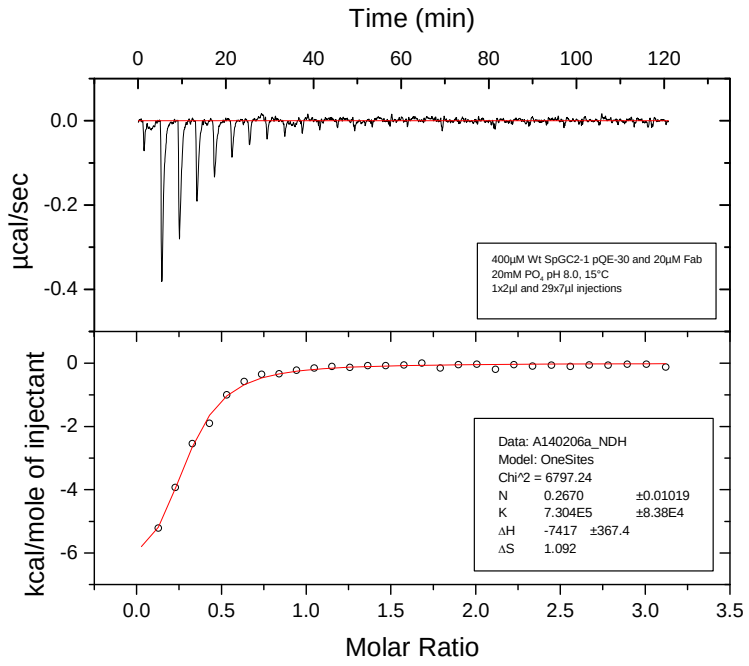


Figure 7.5 – Binding isotherm of Wt D-SpG$_{C2}$-1 and Fab observed by ITC

The analysis of the binding interaction between Wt D-SpG$_{C2}$-1 and Fab measured by ITC at 15 °C is presented. A solution of 400 µM SpG was titrated into 20 µM Fab using injections of 1 x 2 µl and 29 x 7 µl.

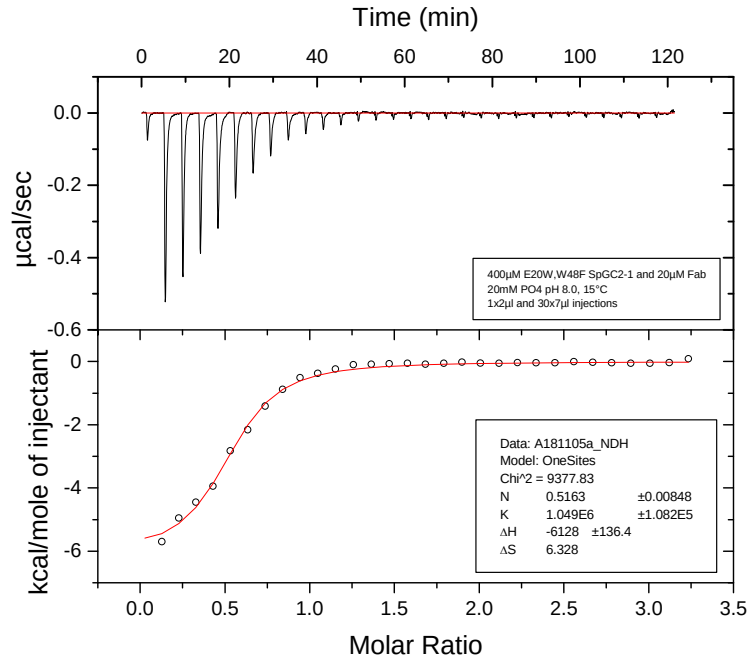


Figure 7.6 – Binding isotherm of E14W-W42F D-SpG$_{C2}$-1 and Fab

The analysis of the binding interaction between E14W-W42F D-SpG$_{C2}$-1 and Fab measured by ITC at 15 °C is presented. A solution of 400 µM SpG was titrated into 20 µM Fab using injections of 1 x 2 µl and 29 x 7 µl. All titrations were carried out in 20 mM potassium phosphate, pH 8.0.

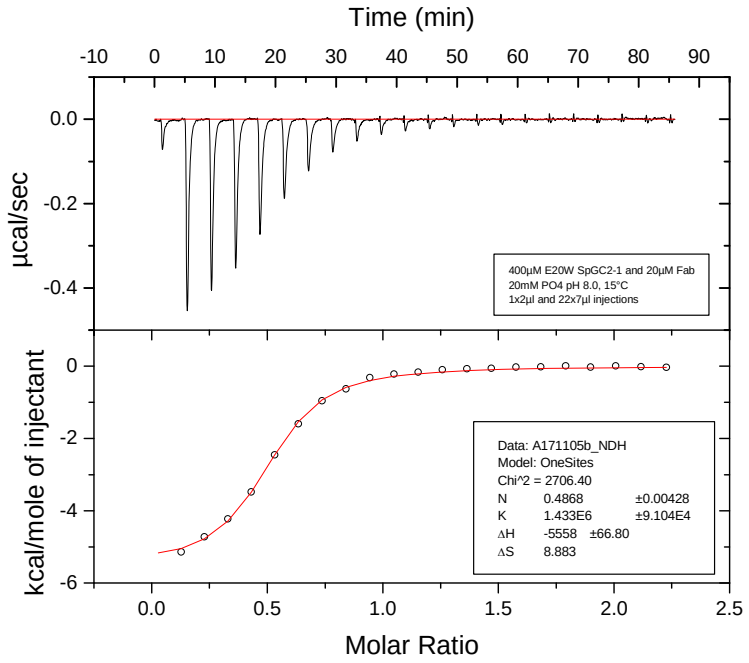


Figure 7.7 – Binding isotherm of E14W D-SpG$_{C2}$-1 and Fab observed by ITC

The analysis of the binding interaction between E14W D-SpG$_{C2}$-1 and Fab measured by ITC at 15 °C is presented. A solution of 400 µM SpG was titrated into 20 µM Fab using injections of 1 x 2 µl and 29 x 7 µl.

These results demonstrate that it is the addition of Fab that produces the heat released upon the interaction. The area underneath each injection peak shown in the top panel is equal to the total heat released for that injection. When these raw data - the total heat per injection - are plotted against the molar ratio of ligand (Fab) added to macromolecule (SpG) in the cell, a complete binding isotherm for the interaction is obtained, presented in the bottom panel. Titrations were carried out in triplicate for the wild-type protein with Fab, and typical isotherms of the data obtained are presented.

Studying the interaction between D-SpG$_{C2}$-1 and Fab has shown that the data obtained were found to fit well to a one binding site model and less rationally to other models. Formation of the SpG•Fab complex is an exothermic reaction, with heat being released upon complex formation, therefore injections of SpG into Fab resulted in a decrease in the electrical energy required to maintain the temperature difference between the sample and reference cells.

7.4.2.2 - Determination of the thermodynamic constants for Fab binding to D-SpG$_{C2}$-1 constructs

Thermodynamic constants, *ΔH* and *ΔS,* and the equilibrium dissociation constant, Kd, for these reactions were obtained by titrating 400 μM D-SpG$_{C2}$-1 constructs into 20 μM Fab, and are presented in table 7.1. Δ*G* was calculated using the Gibbs-Helmholtz equation Δ*G* = Δ*H* – *T* Δ*S*, as stated in equation 6.13 in Chapter 6.

The Kd values obtained for the interaction between wild-type D-SpG$_{C2}$-1, E14W-W42F and E14W with Fab are reasonably close at 1.24 μM ± 0.16, 607.19 nM ± 128.21 and 694.17 nM ± 229.66, respectively. This suggests that the introduction of the tryptophan residue to the alternative position at residue 14, as well as the removal of the Fc reporter tryptophan residue 42, do not greatly affect the interaction of SpG with Fab.

D-SpG$_{C2}$-1 construct	Kd ± std dev. n≥2	Δ*H* (kJ.mol^{-1}) ± std dev. n≥2	Δ*S* (J.mol^{-1}.deg^{-1}) ± std dev. n≥2	Δ*G* (kJ.mol^{-1}) ± std dev. n≥2	N ± std dev. n≥2
Wt	1.24 μM ± 0.16	-26.08 ± 4.30	22.63 ± 15.64	-32.59 ± 0.25	0.40 ± 0.11
E14W	607.19 nM ± 128.21	-24.22 ± 1.36	-35.11 ± 2.93	-34.33 ± 0.37	0.47 ± 0.02
E14W-W42F	694.17 nM ± 229.66	-23.41 ± 2.39	37.03 ±10.04	-34.07 ±0.53	0.51 ± 0.01

Table 7.1 – Summary of the thermodynamic binding constants obtained by titrating D-SpG$_{C2}$-1 constructs into Fab

The thermodynamic values presented in table 7.1 are indicative that the binding of Fab to wild-type SpG is dominated by specific interactions with major enthalpy changes dominating the reaction.

7.5 – Pre-equilibrium binding studies

7.5.1 – Determination of rate constants of D-SpG$_{C2}$-1 and Fab complex formation using stopped-flow fluorescence

All solutions of D-SpG$_{C2}$-1 and human Fab were prepared in 20 mM potassium phosphate buffer at pH 8.0. In addition, 0.5 μl/ml of 1M 2-Mercaptoethanol was added to the solution of Fab in order to prevent dimerisation via C-terminal

disulphide bonds. Both solutions were filtered prior to use to minimise any interference from light scattering, and were also thermo-equilibrated at the operating temperature of 15 °C. Data were analysed as detailed in chapter 5.

7.5.1.1 – Determination of association rate constant k_{ass}

Figure 7.8 shows typical reaction progression curves obtained by using a final concentration of 1.5 μM Fab and final concentrations of E14W-W42F D-SpG$_{C2}$-1 ranging from 30 μM to 50 μM, after mixing had occurred. The association progression curves show a decrease in the fluorescence intensity as the complex between the SpG domain and Fab is formed. By repeating this reaction at various concentrations of E14W-W42F D-SpG$_{C2}$-1, data were collected for the determination of both k_{ass} and k_{diss}.

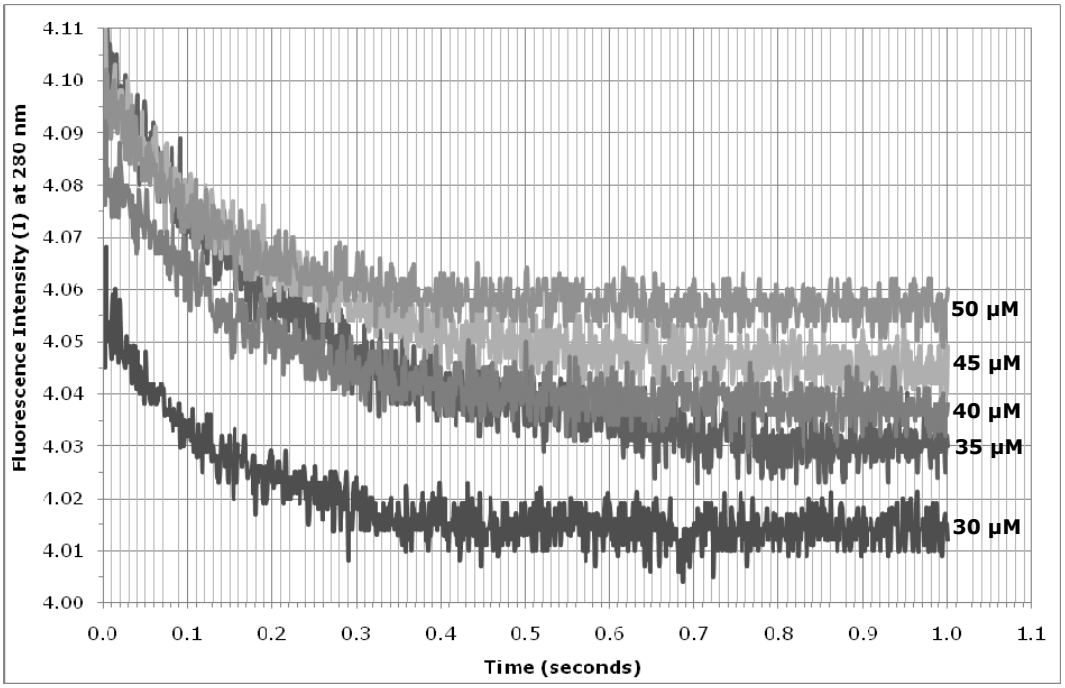


Figure 7.8 – Typical stopped-flow fluorescence profile of the association of varying concentrations of E14W-W42F D-SpG$_{C2}$-1 and Fab

A reaction progression curve for the association of E14W-W42F D-SpG$_{C2}$-1 and Fab is presented. The association curve shows the fluorescence change upon formation of the SpG•Fab complex after mixing 1.5 μM Fab with 30-50 μM SpG, final concentration, at 15˚C, pH 8.0.

7.5.1.2 – Direct determination of the dissociation rate constant k_{diss}

The dissociation rate constant may also be independently measured by a displacement experiment. The rate of dissociation was determined by the 1:1 v/v dilution and displacement of Fab from a pre-formed 1:1 mol:mol ratio (3 μM:3 μM) of the SpG•Fab complex by rapid addition of 30 μM non-fluorescent Fc mutant construct W42F D-SpG$_{C2}$-1.

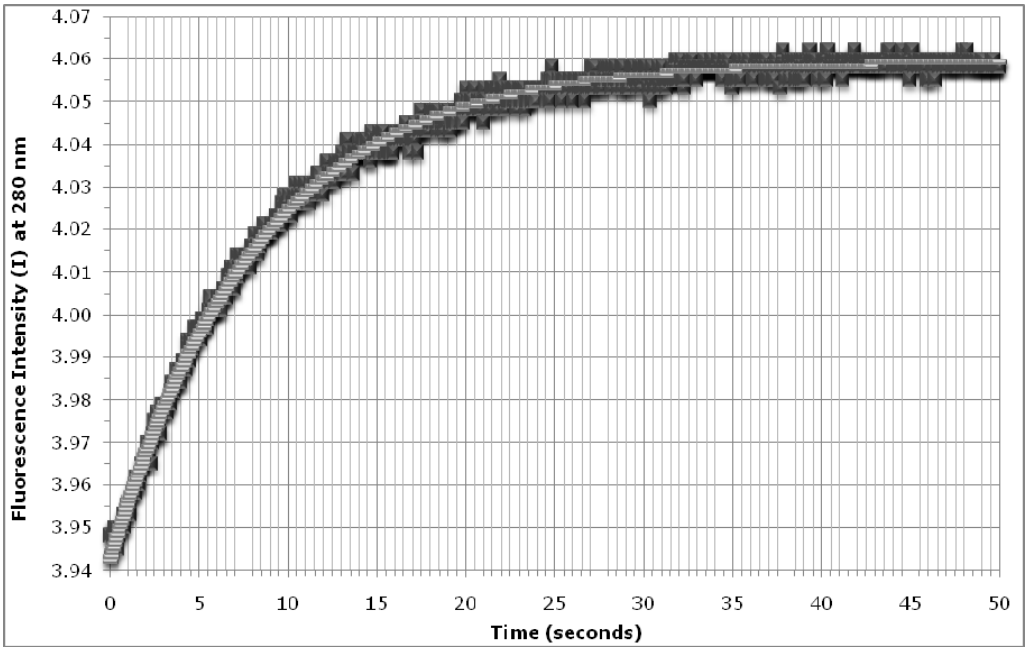


Figure 7.9 – Typical stopped-flow fluorescence profile of the dissociation of E14W-W42F D-SpG$_{C2}$-1 and Fab complex

A reaction progression curve for the dissociation of E14W-W42F D-SpG$_{C2}$-1 and Fab is presented. The dissociation curve shows the fluorescence change upon a 1:1 v/v dilution of 3 µM:3 µM SpG•Fab complex, final concentration, with 30 µM W42F D-SpG$_{C2}$-1, at 15˚C, pH 8.0.

Figure 7.9 presents a typical reaction progression curve for the fluorescence change upon dissociation. The half time for this first order reaction is 5.5 s and the rate of this reaction is therefore 0.12 s^{-1} and is the value of k_{diss} under these conditions.

The dissociation rate constant obtained was used to set the intercept of the y-axis, k_{diss}, to form the gradient of the line and the intercept of the y-axis, respectively.

7.5.1.3 – Calculation of the pre-equilibrium Kd for the SpG•Fab complex

A plot of the observed rate constant (k_{app}) for the association reaction against final SpG concentration will give a straight line with the slope giving the second-order association rate constant (k_{ass} $M^{-1}s^{-1}$*),* and the intercept (or the rate of dissociation determined from the displacement reaction described above) giving the first-order dissociation rate constant (k_{diss} s^{-1}). A linear plot of k_{app} (s^{-1}) against varying E14W-W42F D-SpG$_{C2}$-1, is shown in figure 7.10, allows k_{ass} (0.14 $\mu M^{-1}s^{-1}$) and k_{diss} (0.12 s^{-1}) to be determined from the gradient of the line and the intercept of the y-axis respectively. These values then enable the pre-equilibrium Kd = k_{diss} / k_{ass} to be calculated as 857 nM.

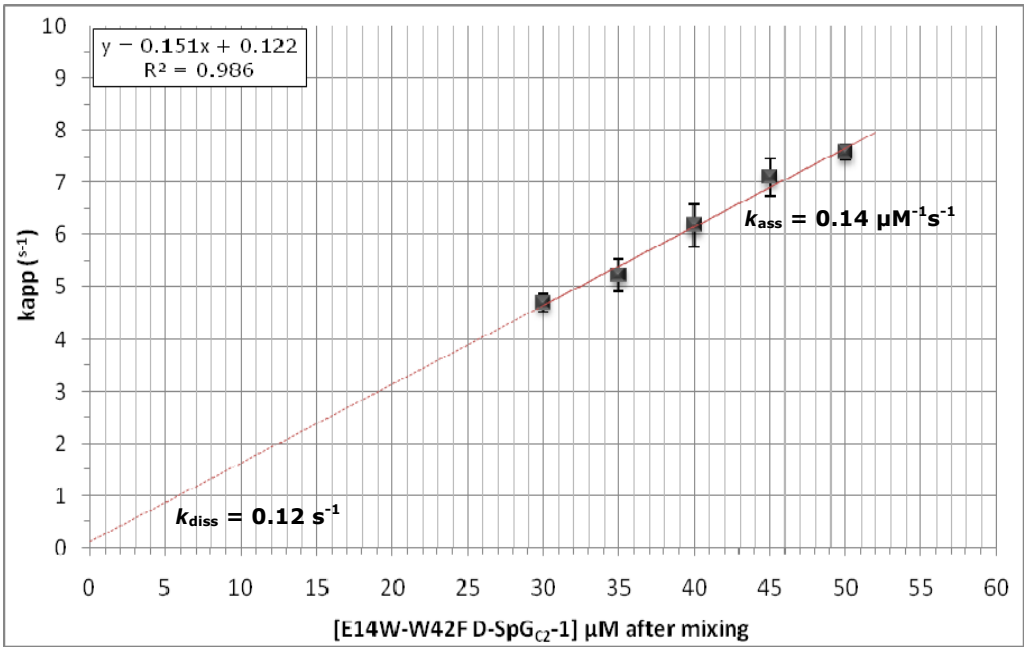


Figure 7.10 – Typical stopped-flow fluorescence analysis of the formation of E14W-W42F D-SpG$_{C2}$-1 and Fab complex

A linear plot of k_{app} (s^{-1}) against varying concentrations of E14W-W42F D-SpG$_{C2}$-1 at 15°C, pH 8.0 is presented.

The reaction progression curves, presented in figures 7.8 and 7.9, fit well to single-exponential equations. Assuming that the fluorescence change occurs on complex formation and dissociation, it suggests that both the association and dissociation of the SpG and Fab complex occur by simple processes. However, a more complex situation could exist in which a rapid, non-rate limiting conformational change may precede or follow the binding reaction. Alternatively, a non-rate limiting or rate limiting structural change may occur before or after the dissociation process. However, these possibilities can almost certainly be ignored since the structure of the domain and the Fab appear to be unchanged before and after complex formation [Gronenborn *et al.*, 1991; Achari *et al.*, 1992; Derrick and Wigley, 1992; Lian *et al.*, 1992; Gallagher *et al.*, 1994; Derrick *et al.*, 1994; Sauer-Eriksson *et al.*, 1995].

By incorporating the E14W mutant close to the proposed Fab binding site it has been shown that this tryptophan residue acts well as a reporter group solely for SpG binding to the C_H1 domain of Fab. Therefore, the resulting double mutant, E14W-W42F D-SpG$_{C2}$-1 *pQE*-30, was then employed as a template for subsequent Fab mutants. The single mutant, E14W (which comprises two tryptophan residues, E14W-W42) was also employed as a control for all experiments undertaken in this study. This was to ensure that the incorporation of the tryptophan residue at the alternative position of 14 has not affected the binding

potential or structure of the SpG construct. A linear plot of k_{app} (s^{-1}) against varying E14W D-SpG$_{C2}$-1 is shown in figure 7.11, and a summary of the rate constants obtained for both E14W-W42F and E14W D-SpG$_{C2}$-1 and Fab complex formation is presented in table 7.2.

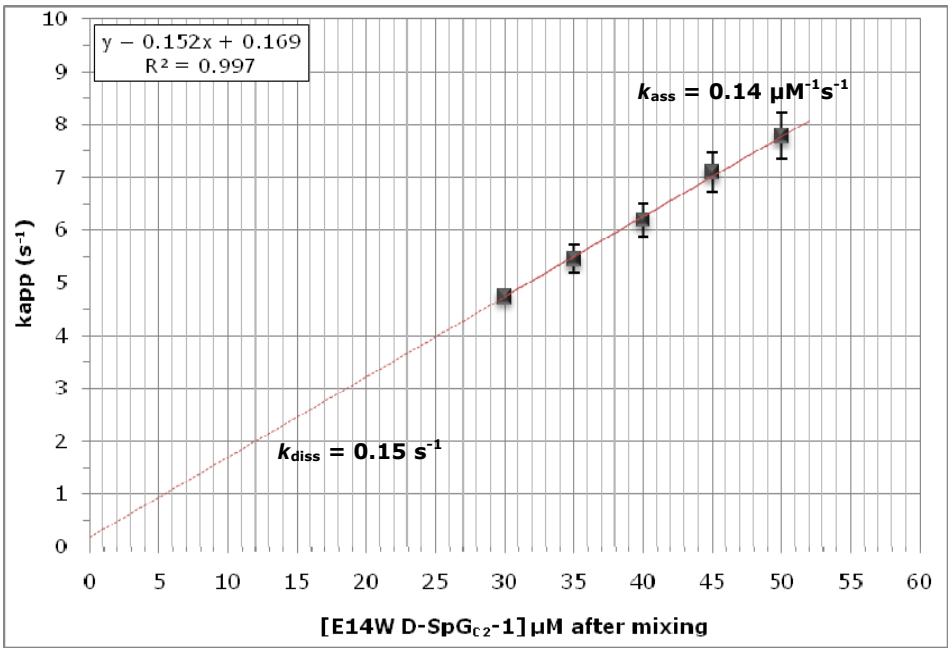


Figure 7.11 – Typical stopped-flow fluorescence analysis of the formation of E14W D-SpG$_{C2}$-1 and Fab complex

This graphical determination represents a linear plot of k_{app} (s^{-1}) against varying concentrations of E14W D-SpG$_{C2}$-1 at 15°C, pH 8.0. The rate constants k_{ass} and k_{diss} can be determined from the gradient of the line and the intercept of the y-axis respectively.

D-SpG$_{C2}$-1 Construct	k_1 ($\mu M^{-1}s^{-1}$) ± std dev. n≥2	k_{-1} (s^{-1}) ± std dev. n≥2	Kd (µM) ± std dev. n≥2
E14W-W42F	0.14 ± 0.02	0.12 ± 0.001	0.86 ± 118.89
E14W	0.14 ± 0.02	0.15 ± 0.02	1.09 ± 0.04

Table 7.2 – Summary of the comparison of rate constants and binding affinity obtained for the interaction of E14W-W42F and E14W D-SpG$_{C2}$-1 with Fab

The association and dissociation rate constants and the Kd values obtained for the binding interaction between both E14W-W42F and E14W D-SpG$_{C2}$-1 constructs with Fab are very similar. This suggests that these tryptophan mutations do not affect the binding of SpG to Fab, and shows that the incorporation of the tryptophan residue at position 14 reports the binding of Fab very well.

7.5.2 – Effect of mutagenesis on the binding constants for the D-SpG$_{C2}$-1 and Fab complex

As previously discussed, mutant constructs of SpG, which have been shown to be critical for the binding interaction with Fab, have been produced to engineer favourable binding properties, as well as to allow specific amino acid residues and their contribution to the biological function of SpG to be evaluated. To achieve these objectives, new mutants were specifically designed to potentially alter the Fab binding affinities from the "wild-type" protein, which is E14W-W42F D-SpG$_{C2}$-1 for these Fab studies. This is important, as it will give a complete overview of the binding interface.

The Kd and rate constants for the complex formed between all D-SpG$_{C2}$-1 mutant constructs were determined by stopped-flow fluorimetry. The binding association and dissociation of Fab to E14W-W42F SpG, presented in figures 7.8 and 7.9, has shown that these reactions each produce a signal that can be easily analysed by stopped-flow. This can then be compared with signals obtained by the mutant constructs, as well as the rate constants, in order to identify residues that are important to the interaction.

7.5.3 – Determination of rate constants for Fab D-SpG$_{C2}$-1 mutants

SpG's interaction with the C_H1 domain of Fab lies predominantly within β-strand 2, but also involves a minor region at the C-terminal end of the α-helix. The site of interaction spans residues Lys9, Thr10, Leu11, Lys12, Gly13, Glu14, Thr15 and Thr16 from β-strand 2 on SpG, with residues Tyr32, Ala33, Asn34, Asp35, Asn36 and Gly37 from the C-terminal end of the α-helix of SpG constituting the minor contact region [Derrick and Wigley, 1992; Lian *et al.*, 1994; Derrick and Wigley, 1994; Sauer-Eriksson *et al.*, 1995]. Six of these fourteen residues, from both the major and minor contact regions, namely residues Thr10, Glu14, Thr15, Thr16, Tyr32 and Asn36 formed the basis of this study and were mutated to different side-chains to characterise their side-chain involvement in the interaction with Fab.

The rates of the association and dissociation of the complex between D-SpG$_{C2}$-1 constructs and Fab, as well as the pre-equilibrium Kd, are summarised below in table 7.3.

D-SpG$_{C2}$-1 construct	k_{ass} ($\mu M^{-1}s^{-1}$) ± std dev. n≥2	k_{diss} (s^{-1}) ± std dev. n≥2	Kd (μM) ± std dev. n≥2
Wt	ND	ND	ND
E14W-W42F	0.14 ± 0.02	0.12 ± 0.01	0.906 ± 118.89
E14W	0.14 ± 0.02	0.15 ± 0.02	1.09 ± 0.04
T10A	ND	ND	ND
T10S	0.09 ± 0.01	0.52 ± 0.04	5.74 ± 0.91
T15A	ND	ND	ND
T15S	0.07 ± 0.01	0.68 ± 0.02	10.20 ± 1.36
T16A	0.08 ± 0.02	0.33 ± 0.02	4.07 ± 0.97
T16S	0.08 ± 0.02	0.30 ± 0.01	3.81 ± 1.03
T15A-T16A	ND	ND	ND
Y32F	0.08 ± 0.01	0.14 ± 0.01	1.75 ± 0.13
N36A	ND	ND	ND
N36D	ND	ND	ND
N36Q	ND	ND	ND

Table 7.3 – Summary of the rate constants and binding affinity obtained for the interaction of D-SpG$_{C2}$-1 mutants and Fab

As previously mentioned, the results obtained for the D-SpG$_{C2}$-1 construct E14W-W42F studied by stopped-flow fluorimetry shows very little difference in its association and dissociation with Fab, as well as the overall binding affinity, within error, to that of E14W SpG. This indicates that the introduction of the tryptophan residue at position 14 reports Fab binding as proposed and the removal of tryptophan residue 42 did not affect the interaction of SpG with Fab. In addition, the overall binding profile for the mutant Y32F with Fab compares quite well with that of E14W-W42F, with only a slight increase in Kd and decrease in the association constant. The dissociation constant of the SpG and Fab complex compares very well to that of the Wt domain and Fab, suggesting that, overall, this tyrosine residue is not significantly involved in the interaction with Fab.

However, six mutant constructs were found to affect the binding affinity of SpG with Fab, namely T10A, T15A, T15A-T16A, N36A, N36D and N36Q. In addition,

during these stopped-flow studies, no significant fluorescence change was observed between Wt SpG_{C2}-1 and Fab, suggesting that, although tryptophan residue 42 reports changes in fluorescence upon Fc binding, it does not report the binding of Fab to SpG.

Three threonine β-strand 2 amino acid residues were selected to be mutated, namely Thr10, Thr15 and Thr16. These threonine amino acids appear to play important roles in the stabilising hydrogen bond network and have been shown to constitute the major contact region. Threonine 10 forms a hydrogen bond with Fab residue Lys215 located on the 7^{th} β-strand on the C_H1 domain. Threonine 15 forms a hydrogen bond with Ser209 and Thr211, with threonine 16 also forming hydrogen bonds with Ser209 as well as with Ser210. Results obtained by stopped-flow studies showed that no apparent binding was observed with Fab when Thr10 and Thr15 were mutated to alanine residues, as shown in figure 7.12. This near total loss of affinity suggests that these residues play significant roles in the interaction and are essential to Fab binding.

The three threonine residues at positions 10, 15 and 16 were also mutated to serine residues, due to their very similar side-chains. The aim of these mutations was to study the effect of replacing a threonine residue with another uncharged, polar residue on its interaction with Fab. Both threonine and serine possess uncharged, polar side-chains, as well as hydroxyl groups, which provide polarity to their otherwise hydrophobic side-chains, and threonine differs only in that it has a hydroxymethyl group in place of a hydrogen group found in serine. It was proposed that by replacing threonine 10 with serine, the hydrogen bond with its corresponding Fab residue will remain intact, but will be in a different orientation. A typical stopped-flow fluorescence analysis of T10S is presented in figure 7.13. Either an increase in k_{diss} or a decrease in k_{ass}, or a combination of both may bring about an increase in Kd of the complex. The calculated Kd value of the complex formed between T10S and Fab is 5.74 ± 0.91 μM, compared well with to the Kd value of 0.906 ± 118.89 μM for the equivalent complex formed by E14W-W42F. This 6-fold increase in Kd is due to the combination a 4-fold increase in the dissociation rate and a 1.5-fold decrease in the association rate.

These results compare well with the T15S mutant construct, and a typical stopped-flow fluorescence analysis of T15S is presented in figure 7.14. The calculated Kd value of the complex formed between T15S and Fab is 10.20 ± 1.36 μM, which is a 11-fold increase compared to the Kd value for the equivalent complex formed by E14W-W42F, and is nearly double to that obtained for the

T10S mutation. The increase in Kd is due to a 6-fold increase in the dissociation rate and a 2-fold decrease in the association rate.

Characterising the interaction between T16A and T16S mutants and Fab by stopped-flow demonstrated that this amino acid produces a 4-fold increase in Kd regardless of whether it is mutated to an alanine or a serine residue. In addition, both mutations produce a 2.5-fold increase in the dissociation rate and a 1.75-fold decrease in the association rate. This suggests that Thr16 only plays a small role in the interaction with Fab, and is not affected by mutation. Furthermore, no binding was observed with Fab for the double mutant construct T15A-T16A, which was not surprising as a total loss of binding was found with the T15A mutant, even though T16A retained Fab binding.

From calculations based on the Kd of the complexes formed between the T15S and T16S and Fab it can be estimated that the Kd for the double mutant T15S-T16S interaction with Fab would be around 40 μM assuming the presence of two mutations does not perturb the protein fold.

Asparagine residue 36 has been implicated in the minor interfacial interaction of SpG with Fab. It is situated on the very end of the C-terminal end of the α-helix and forms a hydrogen bond with the backbone of Fab residue Val128. Asn36 also forms part of a stabilising hydrogen network, with its side-chain nitrogen atom and its backbone nitrogen each forming a hydrogen bond with the oxygen atom of the backbone of SpG residue Tyr32. This tyrosine residue in turn completes the hydrogen network, with its side-chain oxygen atom forming a hydrogen bond with Fab residue, Pro126. It was decided to mutate this asparagine to alanine (N36A), aspartic acid (N36D) and glutamine (N36Q) residues. However, no significant fluorescence change was observed between each of these asparagine mutations with Fab. These results lend themselves to three possible interpretations. Either that the loss of binding is a direct consequence of the loss in affinity, or that this particular residue can still bind Fab, but the fluorescence change is limited by the event of a structural change upon mutation. However, there is no supporting data available to explain these results, and unfortunately due to the lack of availability of Fab protein, further characterisation could not be achieved by other methods. The third possibility could be that not all Fab molecules have affinity for SpG. Different stocks of Fab protein may exhibit widely varying binding affinity, and it has been stated that only approximately 50% of human Fab fragments are believed to bind to SpG [Erntell *et al*., 1983].

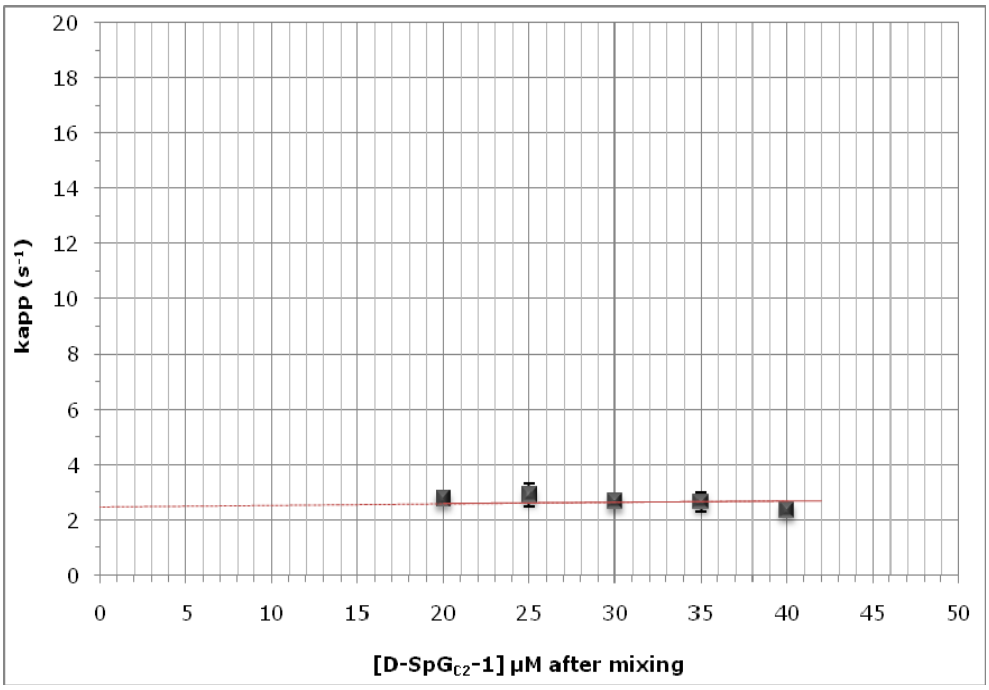


Figure 7.12 – Typical stopped-flow fluorescence analysis of the no Fab complex formation with a selection of D-SpG$_{C2}$-1 mutant constructs

This graphical determination represents a resulting linear plot of k_{app} (s^{-1}) against varying concentrations of D-SpG$_{C2}$-1 at 15°C, pH 8.0, when the formation of the Fab complex cannot be measured.

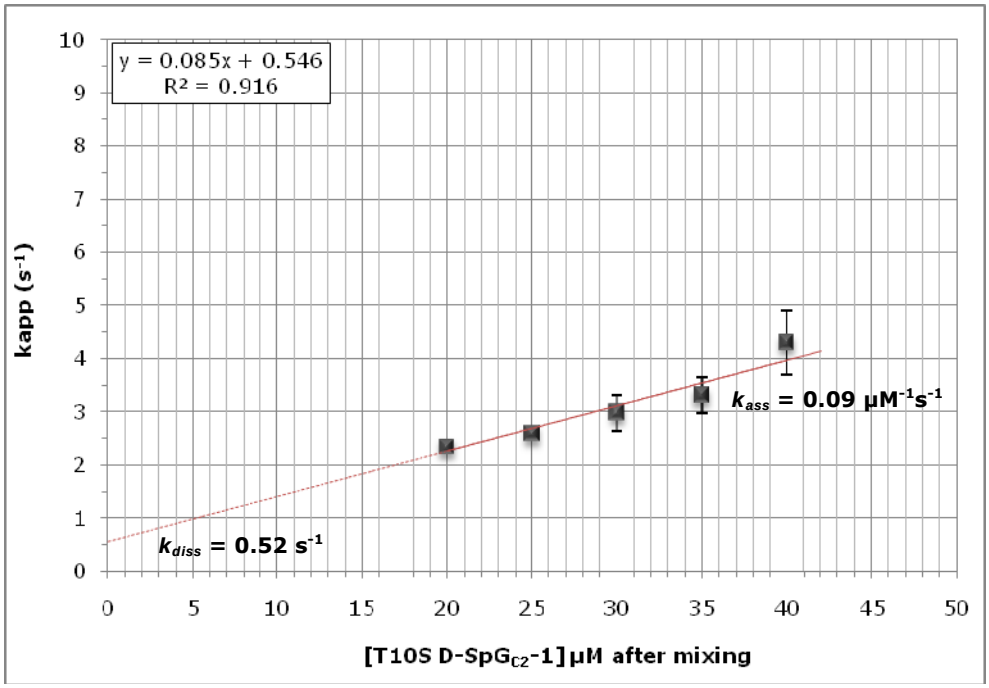


Figure 7.13 – Typical stopped-flow fluorescence analysis of the formation of T10S D-SpG$_{C2}$-1 and Fab complex

This graphical determination represents a linear plot of k_{app} (s^{-1}) against varying concentrations of T10S D-SpG$_{C2}$-1 at 15°C, pH 8.0. The rate constants k_{ass} and k_{diss} can be determined from the gradient of the line and the intercept of the y-axis, respectively.

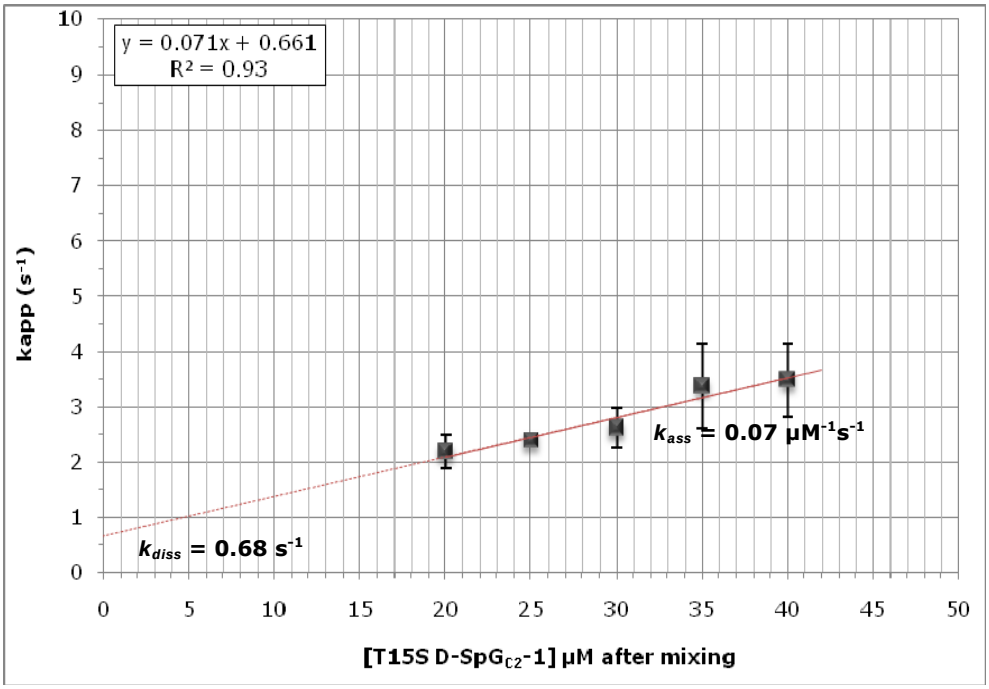


Figure 7.14 – Typical stopped-flow fluorescence analysis of the formation of T15S D-SpG$_{C2}$-1 and Fab complex

This graphical determination represents a linear plot of k_{app} (s^{-1}) against varying concentrations of T15S D-SpG$_{C2}$-1 at 15°C, pH 8.0. The rate constants k_{ass} and k_{diss} can be determined from the gradient of the line and the intercept of the y-axis, respectively.

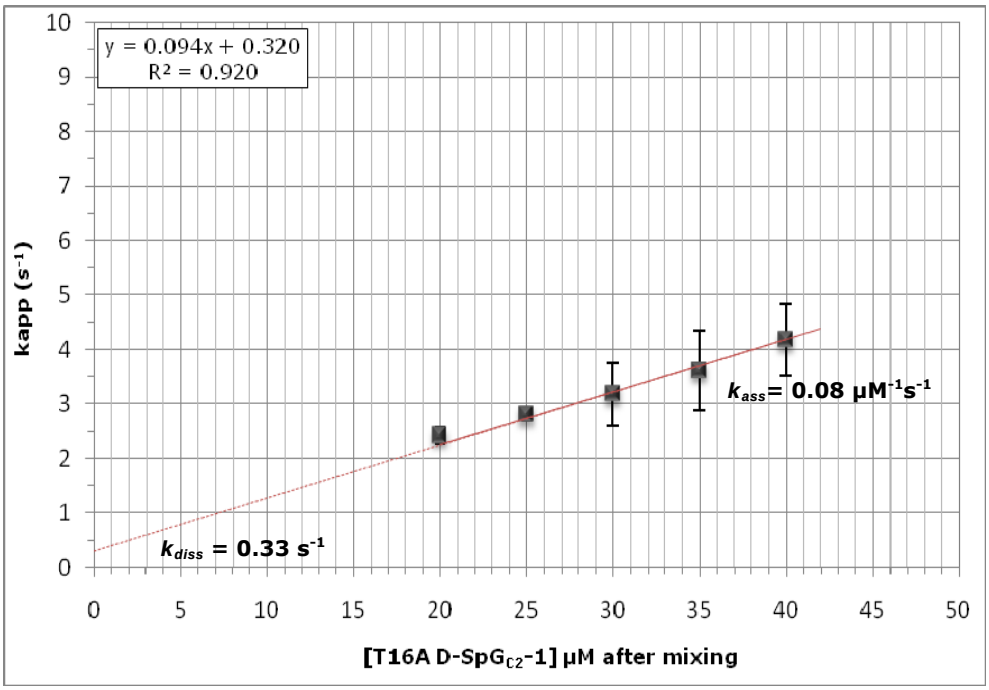


Figure 7.15 – Typical stopped-flow fluorescence analysis of the formation of T16A D-SpG$_{C2}$-1 and Fab complex

This graphical determination represents a linear plot of k_{app} (s^{-1}) against varying concentrations of T16A D-SpG$_{C2}$-1 at 15°C, pH 8.0. The rate constants k_{ass} and k_{diss} can be determined from the gradient of the line and the intercept of the y-axis, respectively.

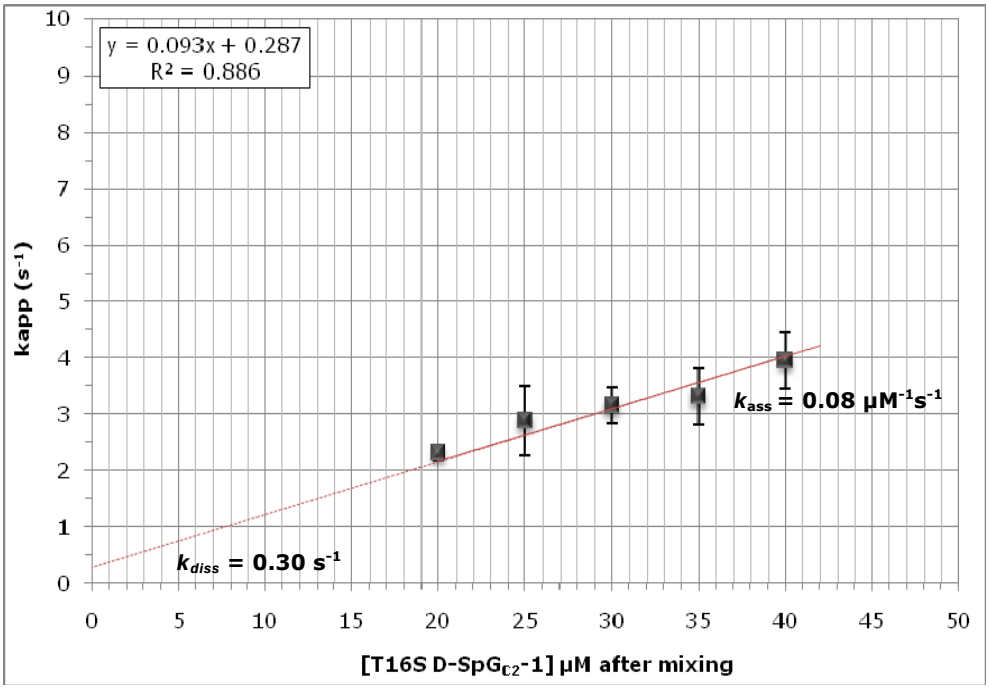


Figure 7.16 – Typical stopped-flow fluorescence analysis of the formation of T16S D-SpG$_{C2}$-1 and Fab complex

This graphical determination represents a linear plot of k_{app} (s^{-1}) against varying concentrations of T16S D-SpG$_{C2}$-1 at 15°C, pH 8.0. The rate constants k_{ass} and k_{diss} can be determined from the gradient of the line and the intercept of the y-axis, respectively.

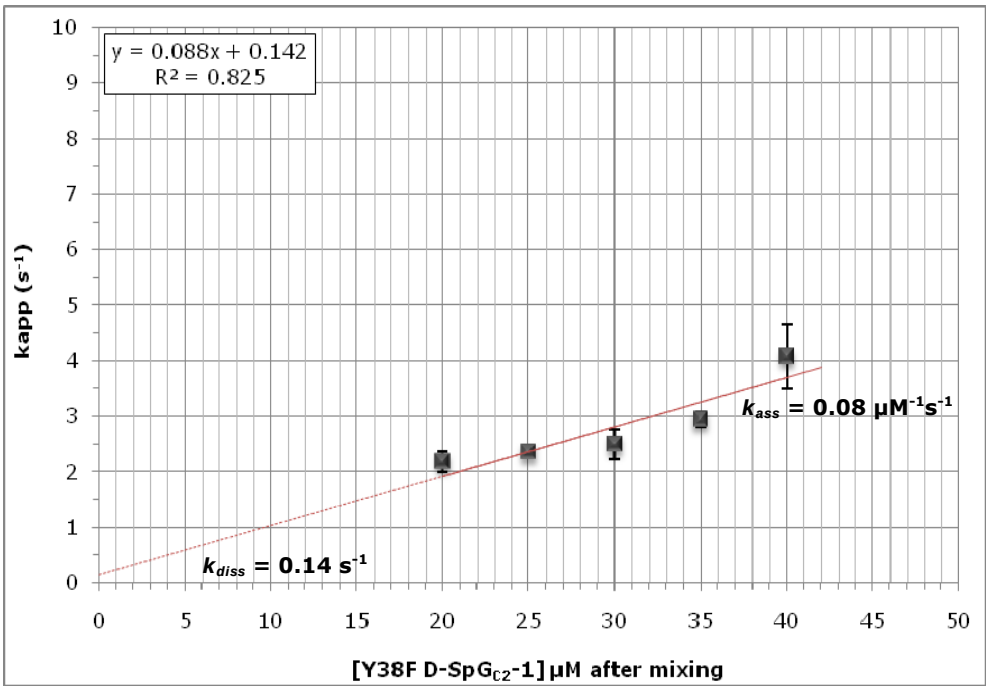


Figure 7.17 – Typical stopped-flow fluorescence analysis of the formation of Y32F D-SpG$_{C2}$-1 and Fab complex

This graphical determination represents a linear plot of k_{app} (s^{-1}) against varying concentrations of Y32F D-SpG$_{C2}$-1 at 15°C, pH 8.0. The rate constants k_{ass} and k_{diss} can be determined from the gradient of the line and the intercept of the y-axis, respectively.

7.6 – Discussion

In this chapter fluorescence has been employed to study the binding interaction between SpG and Fab. The introduction of a tryptophan residue at position 14 in D-SpG_{C2}-1 has allowed this residue to be exploited as a reporter group for exclusively monitoring Fab binding. In addition, this binding interaction has permitted the characterisation of the reversible interactions in terms of binding, stoichiometry, kinetics and thermodynamics, and enabled the reliable determination of these parameters.

Equilibrium studies using fluorescence titrations provided a simple and effective method for determining the true dissociation constant for the complex and allowed the number of binding sites to be calculated. Pre-equilibrium studies employed fast kinetic stopped-flow fluorimetry to determine the rates of association and dissociation of the complex between SpG and Fab. Further characterisation of this interaction was achieved by investigating the role of individual amino-acids that have been implicated in the binding interface between SpG and Fab.

The region of Fab interaction with SpG has been defined by structural studies and shown to lie within the C_H1 domain of the heavy chain [Derrick and Wigley, 1992]. The Fab binding site is distinct from the Fc binding site through which the binding of IgG is mediated, and previous studies have shown that SpG interacts with the Fab regions of human IgG with a 10-fold lower affinity than determined for the human Fc region [Bjorck and Kronvall, 1984]. However, there is no data available to support any of the results obtained, and unfortunately due to the lack of availability of Fab protein, further characterisation cannot be achieved by other methods.

Method @ 15°C, pH 8.0	SpG + Fc	SpG + Fab
Titrations	0.5 μM	3 μM
ITC	0.7 μM	0.6-1 μM
Stopped-Flow	2.5-3 μM	1 μM

Table 7.4 – Comparison of Kd values obtained for SpG with Fc and Fab obtained by various measuring techniques

CHAPTER 8.0

GENERAL DISCUSSION

8.1 - Introduction and aims

Protein G is a multi-domain protein associated with the cell wall from certain Streptococci that functions as a binding unit to several proteins including IgG. It has a high affinity for the interface between the C_H2-C_H3 domains of Fc of IgG, where it blocks the binding to the Fc receptor. Protein G (and SpG) also binds IgG at the C_H1 domain, which explains its cross-reactivity with Fab without affecting the antigen-antibody interaction. The 208 bp single Ig-binding domain gene construct (SpG) based on the wild-type SpG_{C2}-1 gene encoding the C2 domain, was successfully cloned into the N-terminal 6xHis tagged *pQE*-30 expression vector. Specific amino acid residues that contribute to the Fc and Fab binding interactions via a network of hydrogen bonds were identified and effectively modified by a programme of two-sided SOE-ing site-directed mutagenesis.

Generating a library of constructs that offers a range of binding affinities enables various biotechnological uses. Serum antibodies have many applications, such as for the construction of affinity columns, as a coating reagent or for labelling in an immunoassay. However, the majority of these applications require some degree of antibody purification, and to obtain the accepted level of >80% purity requires a large amount of the antibody. As a result, the discovery and isolation of Ig-binding proteins, such as protein A, G and L, has led to the use of these proteins in biotechnology as immunochemical reagents for identifying and purifying antibodies. Furthermore, as the use of antibodies in research and medicine has increased dramatically over recent years, the development of rapid and efficient techniques for Ig purification has become increasingly important. Ig-binding proteins are already heavily exploited immunochemical reagents in the screening and isolation of monoclonal antibodies, but they all possess limitations in their application. Thus, further investigations on these proteins, such as SpG and its interaction with Fc and Fab undertaken in this study, should increase our knowledge and understanding of this Ig-binding protein, as well as to increase its use in immunochemistry and biotechnology. For example, to remove Ig or Ig complexes from sera, a high affinity SpG ligand would be beneficial. Alternatively, to purify Ig from solution, an affinity ligand with a lower affinity would offer advantages for the elution of the Ig from the column. Therefore, it was for these reasons that an extensive series of mutant proteins were produced. A high level of expression for the wild-type, Fab and all but one of Fc mutant constructs, was achieved to evaluate the contribution of specific amino acids to the biological function of SpG with the aim of engineering proteins with different IgG subclass affinities or pH dependencies.

To enable the binding interaction between SpG with Fc and Fab to be fully characterised, knowledge of the contribution from each amino-acid involved is required. The solvent exposure and polarity of the environment surrounding several residues located on both binding sites for both Fc and Fab has been investigated in this study using fluorescence spectroscopy, as has the stability of the Ig-binding domain when various mutants are introduced into the protein. These studies have been achieved by exploiting unique tryptophan residues at key positions within each binding interface. The presence of a native tryptophan (W42) buried in the folded state, and the introduction of a tryptophan residue at position 14 (E14W) allowed these residues to be exploited as reporter groups and have proved instrumental in elucidating the polarity of the environment at different positions within each binding site. The spacial arrangements of the secondary-structure elements, in which an α-helix is positioned diagonally across a β-sheet, means that SpG is highly stable and therefore, provided an excellent model for conformation stability studies using fluorimetry and circular dichroism (CD), as well as by guanidine hydrochloride solvent denaturation experiments. Thus, polarity and stability changes in the mutant constructs surrounding these fluorophores, designed to probe the interaction with Fc or Fab respectively, were exclusively monitored. Certain residues will change the susceptibility of each tryptophan residue and alter the environment, as will the binding of either Fc or Fab fragments.

Since an ideal protein for biotechnology applications must be not only thermodynamically stable, but also offer the appropriate affinity, the binding properties of SpG and its mutant constructs were characterised thoroughly by a combination of fluorescence titration and isothermal titration calorimetry (ITC), and kinetic measurements of real-time binding using stopped-flow fluorescence spectroscopy. The binding interactions of SpG with Fc and Fab permitted the characterisation of the reversible interactions in terms of binding, stoichiometry, kinetics and thermodynamics, and enabled the reliable determination of these parameters.

8.2 - Analysis of wild-type SpG binding interactions

Studies performed on wild-type SpG were carried out to study the native properties of the binding interaction with both Fc and Fab fragments. Equilibrium studies using fluorescence titrations provided a simple and effective method for determining the true dissociation constant for the complex and allowed the number of binding sites to be calculated. The Kd value for Wt D-SpG_{C2}-1 binding

to Fc was able to be described by a single value of 506 ± 0.12 nM at 15°C, pH 8.0 and 250 ± 0.22 nM at 25°C, pH 6.0, and the stoichiometry of binding confirmed that Fc has two identical binding sites for SpG. Previous data obtained under the same conditions to determine the equilibrium dissociation constant for the interaction of D-SpG_{C2}-1 with Fc are similar to the latter data obtained in these studies. Walker obtained a value of 306 ± 65 nM and Gallagher also obtained a Kd value of 333 nM for the B1 (C1) domain of SpG using titration calorimetry. This compares well, but could indicate that the introduction of a 6x His-tag at the N-terminal end of the protein has slightly increased the binding affinity of D-SpG_{C2}-1 for Fc.

Conversely, fluorescence titration results showed that E14W-W42F D-SpG_{C2}-1 binding to Fab was able to be described by a single Kd value of 3.33 ± 1.40 μM, and the stoichiometry of binding confirmed that Fab has one binding site for SpG. This is expected as Fab fragments have one heavy C_H1 domain with a potential binding site for SpG. By incorporating the E14W mutant close to the proposed Fab binding site it has been successfully shown that this tryptophan residue acts well as a reporter group solely for SpG binding to the C_H1 domain of Fab. The association and dissociation rate constants and the Kd values obtained for the binding interaction between both E14W-W42F and E14W D-SpG_{C2}-1 constructs with Fab are very similar. This suggests that these tryptophan mutations do not affect the binding of SpG to Fab, and shows that the incorporation of the tryptophan residue at position 14 reports the binding of Fab very well. Moreover, the spectra obtained for E14W-W42F D-SpG_{C2}-1 suggests that tryptophan 14, like tryptophan 42, is in a solvent exposed position, and is being quenched on complex formation.

ITC experiments were also performed and demonstrated that it is the addition of both Fc and Fab that produces the heat released upon the formation of each complex in an exothermic reaction. Further analysis indicated that these binding reactions occur spontaneously and are driven by negative enthalpic changes, although entropic changes are unfavourable for the reaction in this direction. A Kd of 677 nM ± 154.1 was obtained for the SpG•Fc complex under these conditions. This value is very similar to that determined by fluorescence titrations performed under the same experimental conditions in which the Kd value for Wt D-SpG_{C2}-1 binding to Fc was found to be 506 nM ± 0.1. This also compares reasonably well with previous data obtained by Gallagher [Gallagher, 1994] for the B1 domain of a different SpG construct using titration calorimetry, which gave a Kd value of

333 nM. Although these values are slightly different, the slight increase in Kd found for the construct described in this thesis could indicate that the introduction of a 6x His-tag at the N-terminal end of the protein has slightly decreased the binding affinity of SpG for Fc**.** Similarly, the stoichiometry of binding for the interaction between SpG and Fc compares well for the two different experiments. Data from the ITC experiments gave a stoichiometry of 1.70 ± 0.05 equivalents of SpG to 1 equivalent of Fc, whereas the stoichiometry of binding from the fluorescence titrations was 2.1 ± 0.4 mol SpG: mol Fc. In contrast, the Kd values obtained for the formation of the wild-type D-SpG_{C2}-1, E14W-W42F and E14W SpG•Fab complex are reasonably close at 1.24 μM ± 0.16, 607.19 nM ± 128.21 and 694.17 nM ± 229.66, respectively. This suggests that the introduction of the tryptophan residue at the alternative position at residue 14, as well as the removal of the Fc reporter tryptophan residue 42, does not greatly affect the interaction of SpG with Fab. The thermodynamic values are indicative that the binding of Fab to wild-type SpG is dominated by specific interactions with major enthalpy changes dominating the reaction.

Stopped-flow experiments indicated that the fluorescence intensity decreases as the complex between the Wt SpG domain with either Fc or Fab is formed. These experimental results also suggest that both the association and dissociation of the SpG complexes occur by simple processes. However, a more complex situation could exist in which a rapid, non-rate limiting conformational change may precede or follow the binding reaction in either Wt D-SpG_{C2}-1 or the ligand. This would explain the discrepancy in results obtained for the SpG and Fc complex formation, with a 4-fold higher Kd value of 2.52 μM being obtained using pre-equilibrium techniques compared to a Kd value of 600 nM obtained in equilibrium studies. Alternatively, a non-rate limiting or rate limiting structural change may occur before or after the dissociation process or it is possible that small changes in Kd may occur as side chains become more immobilised in the complex. However, it is thought that this structural change is relatively small and these possibilities can almost certainly be ignored since the structure of the domain and its ligands appear to be unchanged before and after complex formation. Furthermore, the X-ray structure of the backbone of SpG in free solution and in the complex is super-imposable.

A detailed pre-equilibrium binding study of the electrostatic nature of the interaction was performed at a wide range of pH conditions, namely pH 4.0 to pH 9.0, at which SpG is considered stable. Results demonstrated that the interaction

between wild-type D-SpG$_{C2}$-1 and Fc is pH dependent, exhibiting a maximum binding affinity at pH 6.0, which compares well with results obtained by Walker. This also agrees with results obtained by immobilisation techniques for SpG. The results obtained for the association and dissociation rate constants show that the effect of pH on the Kd appears to be mediated through changes in the rate of association of the complex, as the dissociation is virtually unchanged over the same pH range. This seems to indicate that charged residues are involved in the formation of the complex between wild-type D-SpG$_{C2}$-1 and Fc, but not with the dissociation of the complex. There are a number of ionisable residues in wild-type SpG that are involved in the interfacial interaction with Fc, namely Glu26, Lys27, Lys30, Asp39 and Glu41. At high pH, tyrosines on the Fc, which have been implicated in binding, may also begin to ionise and disrupt the interaction. The p*K* of around 6.3 suggests that one or more of the acidic residues Glu26, Asp39 or Glu41 may be responsible for this pH effect.

8.3 - Analysis of D-SpG$_{C2}$-1 mutant constructs designed to perturb Fc binding interactions

SpG binds Fc at the hinge region that is situated in the cleft between C_H2 and C_H3 domains of Fc. Eight residues within the α-helix and β-strand 3 of SpG have been identified to be involved in the interfacial interactions with the loop regions of Fc, namely Glu26, Lys27, Lys30, Gln31, Asn34, Asp39, Glu41 and Trp42. Six of these residues, Glu26, Lys27, Gln31, Asn34, Glu41 and Trp42, formed the basis of this study, and were produced to potentially alter the Fc binding affinity from that of the wild-type protein. Engineering favourable binding properties allowed specific amino acid residues and their contribution to the biological function of SpG to be evaluated, providing a complete overview of the binding interface. However, Lys27A was not successfully expressed so could not be characterised any further.

The exclusive tryptophan residue in the D-SpG$_{C2}$-1 gene, W42, has been implicated in the interfacial interaction with Fc, and is situated on the N-terminal end of β-strand 3. Although this tryptophan residue is not directly involved in binding to Fc through a hydrogen bond network, it was decided to make W42F mutant in order to collaborate the work performed by Walker. Results obtained the pre-equilibrium studies showed that even though this mutation is in a slightly different construct, the effect was still the same and a loss of binding was found. It is also clear from the ITC experiments that W42 is essential for binding activity, which further supports the previous studies, indicating that tryptophan 42 is very important for the binding interaction of SpG with Fc. However, it appears that

some restoration of activity is brought about by the placement of a tryptophan at residue 14 (E14W) and a Kd of 4.5 μM was observed when in complex with Fc. In practice this may be a useful protein for affinity chromatographic applications as a Kd in the micromolar range is ideal for affinity ligands where easy elution of the bound target protein is required.

These studies go further to show that in addition to the importance of W42, residues E26, Q31 and N34 play an important role in the binding. A dramatic reduction in the affinity of SpG for Fc was obtained for the E26A mutation. This finding is fully supported by results obtained by Sloan and Hellinga [1999] and Strop *et al*., [2000], who found that a single alanine mutation of E26 was found to virtually abolish the stable complex formation, with a >4,000-fold decrease in binding affinity. Their studies also proposed that the binding interface between SpG and Fc is polar in character and involves two well-defined "knobs-into-holes" interactions. Of the residues contributing to the hydrogen bonds across the interface, the most critical interaction was found to be formed by E26, which acts as a charged knob on the surface of SpG, inserting into a polar hole on the Fc fragment.

The second "knob-into-hole" interaction proposed by Sloan and Hellinga is formed by the protrusion of Fc surface residues into a hole of SpG bordered by residues N34, D35, D39, E41 and W42. Aside from the W42 residue, as previously discussed, only the N34 residue was found to significantly affect binding, with a 50-fold increase in affinity for the N34A mutation. Asparagine 34 is implicated in the interfacial interaction of SpG with Fc, and is located at the C-terminal end of the α-helix. This residue's side-chain forms a hydrogen bond with the side-chain of Fc residue Asn434, and continues this vast hydrogen bond network by also forming hydrogen bonds with two other Fc residues, Tyr436 and His433. Again, data obtained by Sloan and Hellinga, as well as studies by Gülich *et al*., [2002] fully support the combined results presented here from conformational stability, equilibrium and pre-equilibrium temperature and pH binding studies. The N34 substitutions were found to be the only mutant constructs that led to major consequences on protein stability. This may be significant as both residue N34 mutations also led to changes in binding characteristics during equilibrium and pre-equilibrium studies, whereby the N34A substitution increased Kd by about 10-fold and the N34D substitution effectively eliminates all apparent binding activity compared to the binding affinity for the equivalent complex formed by wild-type. These results also strongly suggest that this residue plays a significant

role in SpG's interaction with Fc, and the close proximity of the bulky tryptophan residue may also play a part.

Mutants were also investigated to determine the electrostatic nature of the interactions and their sensitivity to changes in pH performed at a wide range of pH conditions. Results obtained demonstrated that the interaction between SpG and Fc was dependent on these environmental conditions. The results also show that mutations Q31E and N34D have both altered the pH vs log Kd profile, thus indicating that ionisation of the replacement is affecting the Kd of the complexes. The glutamine residue at position 31 is of particular interest for its possible use in biotechnological applications. The substitution by alanine resulted in no significant change in the Kd value of the complex, whereas the substitution by a glutamate residue increased Kd by approximately 6-fold, demonstrating that a negative charge at this position is detrimental to binding. Reduced binding affinities were also obtained, with values of 4.5 μM from ITC equilibrium studies and 9 μM from pre-equilibrium binding studies, suggesting that this would be an ideal protein to purify Ig from solution. The lower affinity and the influence from the ionisable group would offer advantages for the elution of the Ig from the column.

8.4 - Analysis of D-SpG$_{C2}$-1 mutant constructs designed to perturb Fab binding interactions

An additional aim of this research was to study the role of various amino acids that form the major contact regions in the binding of SpG with Fab, as very little is known of the binding properties and characteristics of this interaction. Single Ig-binding domains of SpG have also been shown to bind to Fab by forming an anti-parallel interaction with the last β-strand of the C_H1 domain of Fab. SpG's binding site for the interaction with Fab lies predominantly within β-strand 2, but also involves a minor region at the C-terminal end of the α-helix. The amino acid residues that form the major contact regions have been identified, and the site of interaction spans residues Lys9 to Thr16 from β-strand 2 on SpG and Ser209 to Lys216 from the 7th β-strand of the C_H1 domain of Fab. In addition, residues Tyr32 to Gly37 from the C-terminal end of the α-helix of SpG and Pro125 to Tyr129 on the 1st β-strand of the C_H1 domain of Fab constitute the minor contact region. A selection of six residues from both the major and minor contact regions, namely residues Thr10, Glu14, Thr15, Thr16, Tyr32 and Asn36, were chosen for their side-chain involvement in the interaction with Fab and selected to be mutated and used to study and characterise the binding affinity of Fab to SpG. This complex is stabilised through a network of hydrogen bonds, and the

importance and effect of these bonds to the stability and nature of the complex by removing these key residues was the prime aim of the studies.

As previously mentioned, SpG possesses a unique tryptophan residue at position 42, located on β-strand 3, which is known for its important role in Fc binding. Its substitution by phenylalanine reduces the binding interaction of SpG with Fc approximately 300-fold. Therefore, using this W42F D-SpG$_{C2}$-1 construct enabled IgG Fab binding to be studied without the interference from any Fc interactions. However, the removal of the tryptophan in the W42F mutation created a non-fluorescent protein, resulting in Fab binding studies using fluorimetry being impossible. Therefore, a tryptophan residue was re-introduced at an alternative position (E14W) to allow the effect of Fab mutations to be probed in fluorescence studies. By incorporating the E14W mutant close to the proposed Fab binding site this tryptophan residue successfully acted as a reporter group solely for SpG binding to the C_H1 domain of Fab and results compares very well to that of wild-type D-SpG$_{C2}$-1. Therefore, the mutant W42F D-SpG$_{C2}$-1 construct was used to make the E14W mutant and the resulting double mutant, E14W-W42F D-SpG$_{C2}$-1 *pQE*-30, was then employed as a template for subsequent mutants to allow the exclusive study of the anti-parallel interaction that takes place in forming the SpG•Fab complex. The single mutant, E14W, was also employed as a control for all experiments undertaken in this study.

The region of Fab interaction with SpG has been well defined by structural studies and shown to lie within the C_H1 domain of the heavy chain. The Fab binding site is distinct from the Fc binding site through which the binding of IgG is mediated and previous studies have shown that SpG interacts with the Fab regions of human IgG with a 10-fold lower affinity than determined for the human Fc region. For the mutations designed to explore Fab binding interactions none led to significantly reduced stabilities, indeed mutants with decreased affinities for Fab had, if anything, slightly higher conformational stability than wild-type. However, there is no data available to support any of the stability results obtained, and unfortunately due to the lack of availability of Fab protein, further characterisation cannot be achieved by other methods.

However, during pre-equilibrium studies six mutant constructs were found to affect the binding affinity of SpG with Fab, namely T10A, T15A, T15A-T16A, N36A, N36D and N36Q, with no significant fluorescence change being observed. From the three threonine β-strand 2 amino acid residues shown to constitute the major contact region due to their important roles in the stabilising hydrogen bond

network, namely Thr10, Thr15 and Thr16, results obtained showed that no apparent binding was observed with Fab when Thr10 and Thr15 were mutated to alanine residues. This near total loss of affinity suggests that these residues play significant roles in the interaction and are essential to Fab binding. Furthermore, no binding was observed with Fab for the double mutant construct T15A-T16A, which was not surprising as a total loss of binding was found with the T15A mutant, even though T16A retained Fab binding.

Asparagine residue 36 has been implicated in the minor interfacial interaction of SpG with Fab. It is situated on the very end of the C-terminal end of the α-helix and forms a hydrogen bond with the backbone of Fab residue Val128. Asn36 also forms part of a stabilising hydrogen network, with its side-chain nitrogen atom and its backbone nitrogen each forming a hydrogen bond with the oxygen atom of the backbone of SpG residue Tyr32. This tyrosine residue in turn completes the hydrogen network, with its side-chain oxygen atom forming a hydrogen bond with Fab residue, Pro126. It was decided to mutate this asparagine to alanine (N36A), aspartic acid (N36D) and glutamine (N36Q) residues. However, no significant fluorescence change was observed when each of these asparagine mutants were reacted with Fab. These results lend themselves to two possible interpretations. Either that the loss of binding is a direct consequence of the loss in affinity or that this particular mutant can still bind Fab, but the fluorescence change does not take place as usual.

8.5 - Future work

Hybrid proteins of PpL having the ability to bind to light chains of all Ig and also to bind to light and heavy chains of immunoglobulin G can assist the formation of a more stable Ig-binding domain with binding sites of comparable affinity for exploitation as versatile biotechnology tools. One common feature of proteins G and L is that they both possess multiple copies of Ig-binding domains, each containing approximately 60–70 residues. Within each protein these domains are highly conserved, although some sequence differences can exist that will give rise to a range of binding affinities that causes Igs to be eluted from affinity columns in non-symmetrical peaks. Furthermore, the interactions are so strong that the pH at which the elution occurs is very low and thus these native proteins are not ideal ligands for affinity purification. However, the Ig-binding properties of PpL and SpG have just been recently successfully combined through the construction of a novel hybrid protein, LG, consisting of a single Ig-binding domain of L produced in this lab and the domain of SpG used in these studies [Harrison *et al*.,

2008]. These studies have shown that despite the small size of the protein, each domain exhibits all complexes to bind (Fc, C_H1 to SpG and V_L1 to PpL simultaneously). The binding affinity of each participant L and G can be engineered such that each releases its complex partner under similar conditions. Thus a library of such hybrids can be produced for various uses.

Despite their limited binding properties, such Ig-binding proteins have found worldwide use in the purification of antibodies, in immunoassays and in some therapies to remove immune complexes from serum to minimise autoimmune disease and renal transplant rejection. Overall, these studies have improved our knowledge of SpG and its interactions with both Fc and Fab fragments. However, further development of SpG as an immunological tool is still necessary to fully understand the nature of its interaction with Fc and Fab. It is hoped that the significant results obtained from the work carried out on the SpG mutants described in this study, can be exploited to create a universal Ig-receptor binding protein, increasing its use in immunochemistry and biotechnology applications.

CHAPTER 9.0

REFERENCES

Achari A, Hale SP, Howard AJ, Clore GM, Gronenborn AM, Hardman KD and Whitlow M. 1.67-Å X-ray structure of the B2 immunoglobulin-binding domain of streptococcal protein G and comparison to the NMR structure of the B1 domain. *Biochemistry* 1992; **31**: 10449-10457

Ahmad F, Yadav S and Taneja S. Determining stability of proteins from guanidinium chloride transition curves. *Biochemistry Journal* 1992; **287**: 481-485

Åkerström B and Björck L. A physicochemical study of protein G, a molecule with unique immunoglobulin G-binding properties. *The Journal of Biological Chemistry* 1986; **261(22)**: 10240-10247

Åkerström B, Nielsen E and Björck L. Definition of IgG- and albumin-binding regions of streptococcal protein G. *The Journal of Biological Chemistry* 1987; **262(28)**: 13388-13391

Åkerström B and Björck L. Protein L: An immunoglobulin light chain-binding bacterial protein. Characterisation of binding and physicochemical properties. *The Journal of Biological Chemistry* 1989; **264**: 19740-19746

Åkerström B, Nilson BHK, Hoogenboom HR and Björck L. On the interaction between single chain Fv antibodies and bacterial immunoglobulin-binding proteins. *Journal of Immunological Methods* 1994; **177**: 151-163

Åkesson P, Cooney J, Kishimoto F, Björck L. Protein H. A novel IgG binding bacterial protein. *Molecular Immunology* 1990; **27**: 523-531

Åkesson P, Schmidt KH, Cooney J and Björck L. M1 protein and protein H: IgGFc- and albumin-binding streptococcal surface proteins encoded by adjacent genes. *Biochemical Journal* 1994; **300**: 877-886

Alexander P, Fahnestock S, Lee T, Orban J and Bryan P. Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2: Why small proteins tend to have high denaturation temperatures. *Biochemistry* 1992; **31**: 3597-3603

Alston RW, Urbanikova L, Sevcik J, Lasagna M, Reinhart GD, Scholtz JM and Pace CN. Contribution of single tryptophan residue to the fluorescence and stability of ribonuclease Sa. *Biophysical Journal* 2004; **87:** 4036–4047

Alzari PM, Lascombe MB and Poljak RJ. Three-dimensional structure of antibodies. *Annual Review of Immunology* 1988; **6**: 555-558

Anfinsen CB. Principles that govern the folding of protein chains. *Science* 1973; **181**: 223-230

Arai M and Kuwajima K. Role of the molten globule state in protein folding. *Advances in Protein Chemistry* 2000; **53**: 209-210

Armen RS, Mari L, DeMarco, DO, Alonso V, Daggett V and Fersht A. A pleated sheet structure may define the pre-fibrillar amyloidogenic intermediate in amyloid disease. *Proceedings of the National Academy of Sciences (USA)* 2004; **101(32)**: 11622-11627

Baker D and Agard DA. Kinetics versus thermodynamics in protein folding. *Biochemistry* 1994; **33**:7505-7509

Baker D. A surprising simplicity to protein folding. *Nature* 2000; **405**: 39-42

Baldwin RL. The nature of protein folding pathways: The classical versus the new view. *Journal of Biomolecular NMR* 1995; **5(2)**: 103-109

Baldwin RL. How Hofmeister ion interactions affect protein stability. *Biophysical Journal* 1996: **71**: 2056–2063

Baldwin RL. Protein folding from 1961-1982. *Nature Structural Biology* 1999; **6**: 814-817

Ban M. DNA sequencing and PCR methods. *Advances in Clinical Neuroscience Rehabilitation* 2006; **5(6)**: 18-19

Basharov MA. Protein folding. *Journal of Cellular and Molecular Medicine* 2003; **7**: 223–237

Bell CE and Lewis M. A closer view of the conformation of the Lac repressor bound to operator. *Nature Structural Biology* 2000; **7(3)**: 209-214

Berge A, Kihlberg BM, Sjöholm AG and Björck L. Streptococcal protein H forms soluble complement-activating complexes with IgG, but inhibits complement activation by IgG-coated targets. *The American Society for Biochemistry and Molecular Biology* 1997; **272(33)**: 20774-20781

Bird RE, Hardman KD, Jacobson JW, Johnson S, Kaufman BM, Lee SM, Lee T, Pope SH, Riordan GS and Whitlow M. Single-chain antigen-binding proteins. *Science* 1988; **242**: 423-426

Birnboim HC. A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods in Enzymology* 1993; **100**: 243-255

Birnboim HC and Doly J. A rapid alkaline lysis procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* 1979; **7**: 1513-1522

Björck L and Kronvall G. Purification and some properties of streptococcal protein G, a novel IgG-binding reagent. *The Journal of Immunology* 1984; **133(2)**: 969-974

Björck L. A novel bacterial cell wall protein with affinity for Ig L chains. *Journal of Immunology* 1988; **140**: 1194-1197

Borrebaeck CA, Malmborg AC, Furebring C, Michaelsson A, Ward S, Danielsson L and Ohlin M. Kinetic analysis of recombinant antibody-antigen interactions: Relationship between structural domains and antigen binding. *Biotechnology (N.Y)* 1992; **6**: 697-698

Bouvignies G, Bernado P, Meier S, Cho K, Grzesiek S, Brüschweiler R and Blackledge M. Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. *Proceedings of the National Academy of Sciences (USA)* 2005; **102**: 13885-13890

Boyle MDP and Reis KJ. Bacterial Fc receptors. Biotechnology 1987; **5**: 697-703

Brown S and Head-Gordon T. Intermediates and the folding of proteins L and G. *Protein Science* 2004; **13**: 958-970

Calloni, G., N. Taddei. Comparison of the folding processes of distantly related proteins - Importance of hydrophobic content in folding. *Journal of Molecular Biology* 2003; **330(3)**: 577-91

Chellani M. Isothermal titration calorimetry: Biological applications. Application Note 1999: 14-17

Chen Y, Li YH, Chen XP, Gong LM, Zhang SP, Chang ZJ, Zhang XF, Fu XY and Liu L. Point mutation at single tyrosine residue of novel oncogene NOK abrogates tumorigenesis in nude mice. *Cancer Research* 2005; **65**: 10838-10846

Chou, PY and Fasman GD. Conformational parameters for amino acids in helical β-sheet, and random coil regions calculated from proteins. *Biochemistry* 1974; **13**: 211-221

Cline J, Braman JC and Hogrefe HH. PCR fidelity of *pfu* DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Research* 1996; **24(18)**: 3546-3551

Cohen LA. The enzymes. 3rd ed. New York: John Wiley & Sons; 1970: 147-211

Cossins AJ. Characterisation of the binding interaction between Peptostreptococcal protein L and a recombinant domain antibody. PhD Thesis 2006; University of Southampton

Davies DR, Padlan EA and Sheriff S. Antibody-antigen complexes. *Annual Review of Biochemistry* 1990; **59**: 439-473

Daggett V and Fersht AR. The present view of the mechanism of protein folding. *Nature Reviews: Molecular Cell Biology* 2003a; **4:** 497–502

Debarbieux L and Beckwith J. The reductive enzyme thioredoxin 1 acts as an oxidant when it is exported to the *Escherichia coli* periplasm. *Proceedings in the National Academy of Science (USA)* 1998; **95**: 10751-10756

De Château M, Nilson BH, Erntell M, Myhre E, Magnusson CG, Åkerström B, and Björck L. On the interaction between protein L and immunoglobulins of various mammalian species. *Scandinavian Journal of Immunology* 1993; **37**: 399-405

Deisenhofer J. 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* 1981; **20(9)**: 2361-2370

Derrick JP and Wigley DB. Crystal structure of a streptococcal protein G domain bound to a Fab fragment. *Nature* 1992; **359**: 752-754

Derrick JP and Wigley DB. The third IgG-binding domain from streptococcal protein G: An analysis by x-ray crystallography of the structure alone and in a complex with Fab. *Journal of Molecular Biology* 1994; **243**: 906-918

Derrick JP, Maiden MCJ and Feavers IM. Crystal structure of a Fab fragment in complex with a Meningococcal Serosubtype antigen and a protein G domain. *Journal of Molecular Biology* 1999; **293**: 81-91

Dill KA and Chan HS. From Levinthal to pathways to funnels. *Nature Structural Biology* 1997; **4**: 10-19

Dinner AR, Sali A and Karplus M. The folding mechanism of larger model proteins: Role of native structure. *Proceedings of the National Academy of Science (USA)* 1996; **93:** 8356-8361

Dinner AR, Sali A, Smith LJ, Dobson CM and Karplus M. Understanding protein folding via free energy surfaces from theory and experiment. *Trends in Biochemical Science* 2000; **25**: 331-339

Dixon M and Webb EC. *Enzymes* 3rd Edition 1979. London: Longman

Dolk E, van der Vaart M, Hulsik DL, Vriend G, de Haard H, Spinelli S, Cambillau C, Frenken L and Verrips T. Isolation of Llama antibody fragments for prevention of dandruff by phage display in shampoo. *Applied and Environmental Microbiology* 2005; **71(1)**: 442-450

Doyle ML. Characterisation of binding interactions by isothermal titration calorimetry. *Current Opinion in Biotechnology* 1997; **8**: 31-35

Eaton WA, Munoz V, Hagen SJ, Jas GS, Lapidus LJ, Henry ER and Hofrichter. Fast kinetics and mechanisms in protein folding. *Annual Review of Biophyscial Biomolecular Structure* 2000; **29**: 327-359

Eckert KA and Kunkel TA. High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucleic Acids Research* 1990; **18(13)**: 3739-3744

Engvall E and Perlmann P. Enzyme-linked immunosorbent assay (ELISA). III. Quantitation of specific antibodies by enzyme-labelled anti-immunoglobulin in antigen-coated tubes. *Journal of Immunology* 1972; **109**: 129-135

Engvall E and Perlmann P. Enzyme-linked immunosorbent assay (ELISA). III. Quantitative assay of immunoglobulin G. *Immunochemistry* 1974; **8**: 871-874

Enokizono J, Wikström M, Sjöbring U, Björck L, Forsén S, Arata Y, Kato K and Shimada I. NMR analysis of the interaction between protein L and Ig light chains. *Journal of Molecular Biology* 1997; **270**: 8-13

Erlich HA, Gelfand DH and Saiki RK. Specific DNA amplification. *Nature* 1988; **331**: 461-462

Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, Matthews BW. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* 1992; **255**: 178-183

Essen LO and Skerra A. The *de novo* design of an antibody combining site - crystallographic analysis of the V_L domain confirms the structural model. *Journal of Molecular Biology* 1994; **238**: 226-544

Fahnestock SR, Alexander P, Nagle J and Filpula D. Gene for an immunoglobulin-binding protein from a Group G streptococcus. *Journal of Bacteriology* 1986; **167(3)**: 870-880

Evans PE and Radford SE. Probing the structure of folding intermediates. *Current Opinions in Structural Biology* 1994; **4:** 100–106

Falkenberg C, Björck L and Åkerström B. Localisation of the binding site for streptococcal protein G on human serum albumin. Identification of a 5.5-kilodalton protein G binding albumin fragment. *Biochemistry* 1992; **31**: 1451-1457

Feeney RE. Chemical modification of proteins: Comments and perspectives. *International Journal of Peptide and Protein research* 1987; **27**: 145-161

Fersht A. Structure and mechanism in protein science. Freeman. New York 1999

Fedorov AN and Baldwin TO Cotranslational protein folding. *Journal of Biological Chemistry* 1997; **272(52)**: 32715-32718

Foote J and Winter G. Antibody framework residues affecting the conformational of the hypervariable loops. *Journal of Molecular Biology* 1992; **224**: 487-499

Forgsren A, Grubb AO. Many bacterial species bind to human IgD. *Journal of Immunology* 1979; **22**: 1468-1472

Frick IM, Wikström M, Forsén S, Drakenberg T, Gomi H, Sjöbring U and Björck L. Convergent evolution among immunoglobulin G-binding bacterial proteins. *Proceedings in the National Academy of Sciences (USA)* 1992; **89**: 8532-8536

Fujii T, Ohkuri T, Onodera R and Ueda T. Stable supply of large amounts of human Fab from the inclusion bodies in *E.coli*. *Journal of Biochemistry* 2007;

141: 699-707

Galloway CA, Sowden MP and Smith HC. Increasing the yield of soluble protein expressed in *E.coli* by induction during late log phase. *Biotechniques* 2003; **34(3)**: 524-530

Genovese A, Bouvet JP, Florio G, Lamparter-Schummert B, Björck L and Marone G. Bacterial immunoglobulin superantigen proteins A and L activate human mast cells by interacting with immunoglobulin E. *Infection and Immunity* 2000: **68(10)**: 5517-5524

Glasgow JN, Everts M and Curiel DT. Transductional targeting of adenovirus vectors for gene therapy. *Cancer Gene Therapy* 2006; **13(9)**: 830-844

Goding JW. Use of staphylococcal protein A as an immunological reagent. *Journal of Immunological Methods* 1978; **20**: 241-253

Gomi H, Hozumi T, Hattori S, Tagawa C, Kishimoto F and Björck L. The gene sequence and some properties of protein H. A novel IgG-binding protein. *The Journal of Immunology* 1990; **144**: 4046-4052

Gore MG. Spectrophotometry and spectrofluorimetry. First Edition. Oxford Publishing 2000

Gore MG and Bottomley SP. Stopped-flow fluorescence spectroscopy. Spectrophotometry and spectrofluorimetry. First Edition. Oxford Publishing 2000; Chapter 9: 241-264

Gouda H, Torigoe M, Saito A, Sato M, Arata Y and Shimada I. Three-dimensional solution structure of the B domain of Staphylococcal protein A: Comparisons of the solution and crystal structures. *Biochemistry* 1992; **31**: 9665-9672

Gouda H, Shiraishi M, Takahashi H, Kato K, Torigoe H, Arata Y and Shimada I. NMR study of the interaction between the B domain of Staphylococcal protein A and the Fc portion of immunoglobulin G. *Biochemistry* 1998; **37**: 129-136

Goulding NJ, Knight SM, Godolphin JL and Guyre PM. Increase in neurophil Fc receptor I expression following interferon gamma treatment in rheumatoid arthritis. *Annals of the Rheumatic Diseases* 1992; **51**: 465-468

Goward CR, Murphy JP, Atkinson T and Barstow DA. Expression and purification of a truncated recombinant streptococcal protein G. *Biochemistry Journal* 1990; **267**: 171-177

Goward CR, Scawen MD, Murphy JP, Atkinson T. Molecular evolution of bacterial cell-surface proteins. *TIBS* 1993; **18**: 136-140

Graille M, Stura E, Corper AL, Sutton BJ, Taussig MJ, Charbonnier JB and Silverman GJ. Crystal structure of a *Staphylococcus aureus* protein A domain complexed with the Fab fragment of a human IgM antibody: Structural basis for

recognition of B-cell receptors and superantigen activity. *Proceedings in the National Academy of Sciences (USA)* 2000; **97**: 5399-5404

Graille M, Stura EA, Housden NG, Beckingham JA, Bottomley SP, Beale D, Taussig MJ, Sutton BJ, Gore MG and Charbonnier JB. Complex between Peptostreptococcus magnus protein L and a human antibody reveals structural convergence in the interaction modes of Fab binding proteins. *Structure* 2001; **9(8)**: 679-687

Gronenborn AM, Filpula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT and Clore GM. A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* 1991; **253**: 657-661

Gronenborn AM and Clore GM. Identification of the contact surface of a streptococcal protein G domain complexed with a human Fc fragment. *Journal of Molecular Biology* 1993; **233**: 331-335

Gronenborn AM, Frank MK and Clore GM. Core mutants of the immunoglobulin binding domain of streptococcal protein G: Stability and structural integrity. *Federation of European Biochemical Societies* 1996; **398**: 312-346

Guss B, Eliasson M, Olsson A, Uhlén M, Frej AK, Jörnvall H, Flock J and Lingberg M. Structure of the IgG-binding regions of streptococcal protein G. *The EMBO Journal* 1986; **5(7)**: 1567-1575

Harris LJ, Larson SB, Hasel KW, Day J, Greenwood A and McPherson A. The three-dimensional structure of an intact monoclonal antibody for canine lymphoma. *Nature* 1992; **360**: 369-372

Ho SN, Hunt HD, Horton RM, Pullen JK and Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 1989; **77**: 51-59

Holdgate GA and Ward WHJ. Measurements of binding thermodynamics in drug discovery. *Drug Discovery Today* 2005; **10(22)**: 1543-1550

Holliger P and Hudson PJ. Engineered antibody fragments and the rise of single domains. *Nature Biotechnology* 2005; **23(9)**: 1126-1136

Hoozemans JJM, Veerhuis R, Van Haastert ES, Rozemuller JM, Baas F, Eikelenboom P and Scheper W. The unfolded protein response is activated in Alzheimer's disease. *Acta Neuropathologica* 2005; **110(2)**: 165-172

Horton RM, Hunt HD, Ho SN, Pullen and Pease LR. Engineering hybrid genes without the use of restriction enzymes: Gene splicing by overlap extension. *Gene* 1989; **77**: 61-68
Horwich AL and Weissman JS. Deadly conformations - protein misfolding in prion disease. *Cell* 1997; **89**: 499-510

Hudson PJ. Recombinant antibody constructs in caner therapy. *Current Opinion in Immunology* 1999; **11**: 548-557

Humphreys DP, Carrington B, Bowering LC, Ganesh R, Sehdev M, Smith BJ, King LM, Reeks DG, Lawson A and Popplewell AG. A plasmid system for optimisation of Fab' production in *Escherichia coli*: Importance of heavy and light chain synthesis. *Protein Expression and Purification* 2002; **26**: 309-320

Humphrey DP, Heywood SP, King LM, Bowering LC, Turner JP and Lane SE. Engineering *Escherichia coli* to improve the purification of Periplasmic Fab' fragments: Changing the p*I* of the chromosomally encoded PhoS/PstS protein. *Protein Expression and Purification* 2004; **37**: 109-118

Hurst M, Jostock T, Menzel C, Voedisch B, Mohr A, Brennesi M, Kirsch MI, Meier and Dübel S. Single chain Fab (scFab) fragment. *BMC Biotechnology* 2007; **7**: 14-29

Huyghues-Despointes BM, Qu X, Tsai J and Scholtz JM. Terminal ion pairs stabilise the second beat-hairpin of the B1 domain of protein G. *Proteins* 2006; **63**: 1005-1017

Ibrahim S, SeppäläI and Mäkelä O. V-region-mediated binding of human Ig by protein A. *The Journal of Immunology* 1993a; **151(7)**: 3597-3603

Ibrahim S, Kaartinen M, Seppälä I, Matoso-Ferreira A and Mäkelä O. The alternative binding site for protein A in the Fab fragment of immunoglobulins. *Scandinavian Journal of Immunology* 1993b; **37**: 257-264

James EL, Whisstock JC, Gore MG and Bottomley SP. Probing the unfolding pathway of alpha1-antitrypsin. *Journal of Biological Chemistry* 1999; **274(14)**: 9482-9488

Jansson B, Uhlén M and Nygren PÅ. All individual domains of staphylococcal protein A show Fab binding. *FEMS Immunology and Medical Microbiology* 1998; **20**: 69-78

Jelesarov I and Bosshard HR. Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *Journal of Molecular Recognition* 1999; **12**: 3-18

Johansson MU, de Château M, Wikström M, Forsén S, Drakenberg T and Björck L. Solution structure of the albumin-binding GA module: A versatile bacterial protein domain. *Journal of Molecular Biology* 1997; **266**: 859-865

Johnson WC Jr. Secondary structure of proteins through circular dichroism spectroscopy. *Annual Review of Biophysical Chemistry* 1988; **17**: 145-166

Johnson WC jnr. Protein secondary structure and circular dichroism. A practical guide. *Proteins* 1990; **7**: 205-214

Joosten V, Lokman C, vanden Hondel CAMJJ and Punt PJ. The production of antibody fragments and antibody fusion proteins by yeasts and filamentous fungi. *Microbial Cell Factories* 2003; **2(1)**: 1-15

Kahn PC. The interpretation of near-ultraviolet circular dichroism. *Methods Enzymology* 1979; **61***:* 339-378.

Kajander T and Khan PC. Buried charged surface in proteins. *Structure with Folding and Design* 2000; **8(11)**: 1203-1214

Kastern W, Holst E, Nielsen E, Sjöbring U and Björck L. Protein L, a bacterial immunoglobulin-binding protein and possible virulence determinant. *Infection and Immunity* 1990; **58(5)**: 1217-1222

Kastern W, Sjöbring U, Björck L. Structure of peptostreptococcal protein L and identification of a repeated immunoglobulin light chain binding domain. *Journal of Biological Chemistry* 1992; **267**: 12820-12825

Kahn PC. The interpretation of near-ultraviolet circular dichroism. *Methods in Enzymology* 1979; **61**: 339-378

Kellis JT Jnr, Nyberg K, Fersht AR. Energetics of complementary side-chain packing in a protein hydrophobic core. *Biochemistry* 1989; **28**: 4914-4922

Kelly SM, Jess TJ and Price NC. How to study proteins by circular dichroism. *Biochimica et Biophysica acta* 2005; **1751(2)**: 119-139

Kihlberg BM, Collin M, Olsén A and Björck L. protein H, an antiphagocytic surface protein in *Streptococcus pyogenes*. *Infection and Immunity* 1999; **67(4)**: 1708-1714

Kim PS and Baldwin RL. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annual Review of Biochemistry* 1982; **51**: 459–489

Kim PS and Baldwin RL. Intermediates in the folding reactions of small proteins. *Annual Review of Biochemistry* 1990; **59**: 631-66

Klapper MH. On the nature of the protein interior. *Biochimica et Biophysica Acta* 1971; **229(3)**: 557-566

Köhler G and Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 1975; **256**: 495-497

Kronvall G. A surface component in group A, C and G streptococci with non-immune reactivity for immunoglobulin G. *Journal of Immunology* 1973; **111**: 1401-1406

Kuipers BJH and Gruppen H. Prediction of molar extinction coefficients of proteins and peptides using UV absorption of the constituent amino acids at 214 nm to enable quantitative reverse phase high-performance liquid chromatography-mass spectrometry analysis. *Journal of Agricultural Food Chemistry* 2007: **55(14)**: 5445-5451

Kuszewski J, Clore GM and Gronenborn AM. Fast folding of a prototypic polypeptide: The immunoglobulin binding domain of streptococcal protein G. *Protein Science* 1994; **3(11)**: 1945-1952

Kuwajima K. Protein folding *in vitro*. *Current Opinions in Biotechnology* 1992; **3(5)**: 462-467

Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970; **227**: 680-692.

Laidler KJ. Physical Chemistry with Biological Applications. California: Benjamin/Cummings; 1978

Lakowicz JR. Principles of Fluorescence Spectroscopy. 2nd Edition. Plenum Publishing. New York; 1999

Lancet D, Isenman D, Sjödahl j, Sjöquist J and Pecht I. Interactions between staphylococcal protein A and immunoglobulin domains. *Biochemical and Biophysical Research Communications* 1978; **85(2)**: 608-614

Lander ES, Linton LM and Birren B. Initial sequencing and analysis of the human genome. *Nature* 2001; **409**: 860-921

Langone JJ. Use of labelled protein A in quantitative immunochemical analysis of antigens and antibodies. *Journal of Immunological Methods* 1982; **51**: 3-22

Langone JJ. Protein A of *Staphylococcus aureus* and related immunoglobulin receptors produced by Streptococci and Pneumonococci. *Advances in Immunoglobulin* 1982a; **32**: 157-252

Lian LY, Yang JC, Derrick JP, Sutcliffe MJ, Roberts GCK, Murphy JP, Goward CR and Atkinson T. Sequential ^{1}H NMR assignments and secondary structure of an IgG-binding domain from protein G. *Biochemistry* 1991; **30**: 5335-5340

Lian LY, Derrick JP, Sutcliffe MJ, Yang JC and Roberts GCK. Determination of the solution structures of domains II and III of protein G from *Streptococcus* by ^{1}H Nuclear Magnetic Resonance. *Journal of Molecular Biology* 1992; **228**: 1219-1234

Lian LY, Barsukov IL, Derrick JP and Roberts GCK. Mapping the interactions between streptococcal protein G and the Fab fragment of IgG in solution. *Structural Biology* 1994; **1(6)**: 355-357

Lin J, Zhou M, Zhao X, Luo S and Lu Y. Extractive fermentation of L-lactic acid with immobilised *Rhizopus oryzae* in a three-phase fluidised bed. *Journal of Chemical Engineering and Processing* 2006: 304-307

Lindberg MJ and Tangrot. Crystal structure of the ribosomal protein S6 from Thermus thermophilus. *The European Molecular Biology Organisation Journal* 2002; **13(6)**: 1249-1254

Lindmark R, Thorén-Tolling K and Sjöquist J. Binding of Immunoglobulins to protein A and immunoglobulin levels in mammalian sera. *Journal of Iimmunological Methods* 1983; **62**: 1-13

Löfdahl S, Guss B, Uhlén M, Philipson L and Lindberg M. Gene for staphylococcal protein A. *Proceedings in the National Academy of Sciences (USA)* 1983; **50**: 697-701

Lundberg KS, Shoemaker DD, Adams MWW, Short JM, Storge JA and Mathur EJ. High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* 1991; **108**: 1-6

Malissard M and Berger EG. Improving solubility of catalytic domain of human beta-1,4-galatosyltransferase 1 through rationally designed amino acid replacements. *European Journal of Biochemistry* 2001; **268(15)**: 4352-4358

Mariuzza RA and Poljak RJ. The basics of binding: mechanisms of antigen recognition and mimicry by antibodies. *Current Opinions in Immunology* 1993; **5**: 50-55

McCallister EL and Alm E. Critical role of beta-hairpin formation in protein G folding. *Nature Structural Biology* 2000; **7(8)**: 669-73

Messing J, Crea R and Seeburg PH. A system for shotgun DNA sequencing. *Nucleic Acids Research* 1981; **24(9)**: 309-321

Moks T, Abrahmsén L, Nilsson B, Hellman U, Sjöquist J and Uhlén M. Staphylococcal protein A consists of five IgG-binding domains. *European Journal of Biochemistry* 1986; **156**: 637-643

Munson M, Balasubramanian S, Fleming KG, Nagi AD, O'Brien R, Sturtevant JM and Regan L. What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties. *Protein Science* 1996; **5**: 1584-1593

Murphy JP, Duggleby CJ, Atkinson MA, Trowen AR, Atkinson T and Goward CR. The functional units of a peptostreptococcal protein L. *Molecular Microbiology* 1994; **12(6)**: 911-920

Myhre EB, Erntell M. A non-immune interaction between the light chain of human immunoglobulin and a surface component of a *Peptococcus magnus* stain. *Molecular Immunology* 1985; **22(8)**: 879-885

Myers JK and Pace CN. Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Science* 1996; **4(10)**: 2138-2148

Nathans D and Smith HO. Restriction endonucleases in the analysis and restructuring of DNA molecules. *Annual Review of Biochemistry* 1975; **44**: 273-293

Neu HC and Heppel LA. The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts. *The Journal of Biological Chemistry* 1965; **240(9)**: 3685-3692

Nilson BHK, Solomon A, Björck L and Åkerström B. Protein L from *Peptostreptococcus magnus* binds to the κ light chain variable domain. *The Journal of Biological Chemistry* 1992; **4**: 2234-2239

Nilson BHK, Lögdberg, Kastern W, Björck L and Åkerström B. Purification of antibodies using protein L-binding framework structures in the light chain variable domain. *Journal of Immunological Methods* 1993; **164**: 33-40

Nilsson B and Anderson S. Proper and improper folding of a protein in the cellular environment. *Annual Review of Microbiology* 1991; **45**: 607-635

Nilsson B, Moks T, Jansson B, Abrahmsén L, Elmblad A, Holmgren E, Henrichson C, Jones TA and Uhlén M. A synthetic IgG-binding domain based on staphylococcal protein A. *Protein Engineering* 1987; **1(2)**: 107-113

Nilsson B, Forsberg G, Moks T, Hartmaris M, Uhlén M. Fusion proteins in biotechnology and structural biology. *Current Opinions in Structural Biology* 1992; **2**; 569-575

Northey JGB, Di Nardo AA and Davidson AR. Hydrophobic core packing in the SH3 domain folding transition state. *Nature* 2002; **9(2)**: 126-130

Nyemeyer H, Garcia AE and Onuchic JN. Folding funnels and frustration in off-lattice minimalist protein landscapes. *Proceedings in the National Academy of Science (USA)* 1998; **95**: 5921-5928

Oeding P, Grov A and Myklestad B. Immunochemical studies on antigen preparations from *Staphylococcus aureus*. Precipitating and erythrocyte-sensitising properties of protein A (antigen A) and related substances. *Acta Pathologica et Microbiologica Scandenavica* 1964; **62**: 117-127

Olsson A, Eliasson M, Guss B, Nilsson B, Hellman U, Lindberg M and Uhlén M. Structure and evolution of the repetitive gene encoding streptococcal protein G. *European Journal of Biochemistry* 1987; **168**: 319-324

Pace CN. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in Enzymology* 1986; **131**: 266-280

Pace CN, Shirley BA, McNutt M and Gajiwala K. Forces contributing to the conformational stability of proteins. *Federation of American Societies for Experimental Biology Journal* 1996; **10(1)**: 75-83

Pace CN and Trevino S. Protein structure, stability and solubility in water and other solvents. *Philosophical Transactions of the Royal Society B: Biological Sciences* 2004; **359(1448)**: 1225-1234

Pace CN and Vanderburg KE. Determining globular protein stability: guanidine hydrochloride denaturation of myoglobin. *Biochemistry* 1979; **18(2)**: 288-292

Palmer A, Welsh K, Gjörstrup P, Taube D, Bewick M, Thick M. Removal of anti-HLA antibodies by extracoporeal immunoadsorption to enable renal transplantations. *Lancet* 1989; **1**: 10-12

Padlan EA. Anatomy of the antibody molecule. *Molecular Immunology* 1994; **31(3)**: 169-217

Park H, Bradrick TD and Howell EE. A glutamine 67 to histidine mutation in homotetrameric R67 dihdrofolate reductase results in four mutations per single active site pore and causes substantial substrate and cofactor inhibition. *Protein Engineering* 1997; **10(12)**: 1415-1424

Park SH, Shastry MCR and Roder H. Folding dynamics of the B1 domain of protein G explored by ultrarapid mixing. *Nature Structural Biology* 1999; **6**: 943-947

Patella V, Casolaro V, Björck L and Marone G. Protein L: A bacterial Ig-binding protein that activates human basophils and mast cells. *The Journal of Immunology* 1990; **145**: 3054-3061

Patnaik PR. Investigation of induction effect on the steady state performance of a continuous fermentation for recombinant beta-galactosidase. *Process Biochemistry* 2001; **36(11)**: 1069-1074

Perozzo R, Folers G and Scapozza L. Thermodynamics of protein-ligand interactions: History, presence and future aspects. *Journal of Receptors and Signal Transduction* 2004; **24(1&2)**: 1-52

Pierce MM, Raman CS and Nall BT. Isothermal titration calorimetry of protein-protein interactions. *Methods* 1999; **19**: 213-221

Pilcher JB, Tsang VCW, Zhou W, Black CM and Sidman C. Optimisation of binding capacity and specificity of protein G on various solid matrices for immunoglobulins. *Journal of Immunological Methods* 1991; **136(2)**: 279-286

Playfair J. Infection and Immunity. First Edition. Oxford University Press; 1995

Popplewell AP, Gore MG, Scawen M and Atkinson T. Synthesis and mutagenesis of an IgG-binding protein based upon protein A of *Staphylococcus aureus*. *Protein Engineering* 1991; **4(8)**: 963-970
Porter RR. Structural studies of immunoglobulins. *Science* 1973; **180**: 713-716

Porter P, Coley J and Gani M. Immunochemical criteria for successful matching of monoclonal antibodies to immunoassays of peptide hormones for assessment of pregnancy and ovulation. *Progress in Clinical Biological Research* 1988; **285**: 181-200

Ptitsyn OB. Molten globule and protein folding. *Advances in Protein Chemistry* 1995; **47**: 83-229

Pugsley AP. The complete general secretory pathway in Gram-negative bacteria. *Microbiological Reviews* 1993; **57(1)**: 50-108

Radford SE and Dobson CM. From computer stimulations to human disease: emerging themes in protein folding. *Cell* 1999; **97**: 291-298

Ramakrishnan V. Ribosome structure and the mechanism of translation. *Cell* 2002; **108(4)**: 557-572

Ramirez-Alvarado M, Cocco MJ and Regan L. Mutations in the B1 domain of protein G that delay the onset of amyloid fibril formation *in vitro*. *Protein Science* 2003; **12**: 567-576

Richards FM and Lim WA. An analysis of packing in the protein folding problem. *Quarterly Review of Biophysics* 1993; **26(4)**: 423-498

Roben PW, Salem AN and Silverman GJ. V_H3 family antibodies bind domain D of staphylococcal protein A. *The Journal of Immunology* 1995; **154**: 6437-6445

Roder H and Colon W. Kinetic role of early intermediates in protein folding. *Current Opinion in Structural Biology* 1997; **7**: 15-28

Ruan M, Akkoyunlu M, Grubb A and Forsgren A. Protein D of *Haemophilus influenzae*: A novel bacterial surface with affinity for human IgD. *The Journal of Immunity* 1990; **145**: 3379-3384

Saiki RK, Gelfand DH, Stoffer S, Scharf SJ, Higuchi R, Horn GT. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 1988b; **239**: 487-491

Sambrook J, Fritsch EF and Maniatis T. Molecular Cloning: A laboratory manual. 2nd Ed. New York: Cold Spring Harbour Laboratory Press; 1989: 6.3-6.23.

Sanger F, Nicklen S and Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proceedings in the National Academy of Sciences (USA)* 1977; **74**: 5463-5467

Sasaki K and Munson RS. Protein D of *Haemophilus influenzae* is not a universal immunoglobulin D-binding protein. *Infection and Immunity* 1993; **61(7)**: 3026-3031

Sasso EH, Silverman GJ and Mannik M. Human IgA and IgG $F(ab')_2$ that bind to staphylococcal protein A belong to the V_HIII subgroup. *The Journal of Immunology* 1991; **147(6)**: 1877-1883

Sato S, Ward CL, Krouse ME, Wine JJ and Kopito RR. Glycerol reverses the misfolding phenotype of the most common cystic fibrosis mutation. *Journal of Biological Chemistry* 1996; **271(2)**: 635-638

Sauer-Eriksson AE, Kleywegt GJ, Uhlén and Jones TA. Crystal structure of the C2 fragment of streptococcal protein G in complex with the Fc domain of human IgG. *Structure* 1995; **3**: 265-278

Schindler T, Herder M, Marahiel MA, Schmid FX: Extremely rapid protein folding in the absence of intermediates. *Nature Structural Biology* 1995; **2**: 663-673

Schreiber G and Seizer T. Predicting the rate enhancement of protein complex formation from the electrostatic energy of interaction. *Journal of Molecular Biology* 1999; **287**: 409-419

Schutz CN and Warshel A. What are the "dielectric constants" of proteins and how to validate electrostatic models? *Proteins* 2001; **44**: 400-417

Sharp PA, Sugden B and Sambrook J. Detection of two restriction endonucleases activities in *Haemophilus parainfluenzae* using analytical agarose ethidium bromide electrophoresis. *Journal of Biochemistry* 1973; **12**: 3.55-3.63

Shaw KL, Grimsley GR, Yakovlev GI, Makarov AA and Pace CN. The effect of net charge on the solubility, activity, and stability of ribonuclease Sa. *Protein Science* 2001; **10**: 1206–1215

Sheinerman FB, Norel R and Honig B. Electrostatic aspects of protein-protein interactions. *Current Opinions in Structural Biology* 2000; **10**: 153-159

Shortle D. Probing the determinants of protein folding and stability with amino acid substitutions. *Journal of Biological Chemistry* 1989; **264**: 5315-5318

Sjöbring U, Trojnar J, Grubb A, Åkerström B and Björck L. Ig-binding bacterial proteins also bind proteinase inhibitors. *The Journal of Immunology* 1989; **143(9)**: 2948-2954

Sjöbring U, Björck L and Kastern W. Streptococcal protein G: Gene structure and protein binding properties. *The Journal of Biological Chemistry* 1991; **266(1)**: 399-405.

Sjödahl J. Repetitive sequences in protein A from *Staphylococcus aureus*. *European Journal of Biochemistry* 1977a; **73**: 343-351

Sjödahl J. Structural studies on the four repetitive Fc-binding regions in protein A from *Staphylococcus aureus*. *European Journal of Biochemistry* 1977b; **78**: 471-490

Sjöquist J, Movitz J, Johansson IB and Hjelm H. Localisation of protein A in the bacteria. *European Journal of Biochemistry* 1972; **30**: 190-194

Sloan DJ and Hellinga HW. Structure-based engineering of environmentally sensitive fluorophores for monitoring protein-protein interactions. *Protein Engineering* 1998; **11(9)**: 819-823

Smith PK, Kroh RI, Hermanson GY, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ and Klent DC. Measurement of protein using Bicinchoninic acid. *Analytical Biochemistry* 1985; **150**: 76-85

Sohi MK, Wan T, Sutton BJ, Atkinson T, Atkinson MA, Murphy JP, Bottomley SP and Gore MG. Crystallisation and X-ray analysis of a single Fab binding domain from protein L of *Peptostreptococcus magnus*. *Proteins: Structure, Function and Genetics* 1995; **23**: 610-612

Sreerama N and Woody RW. A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Analytical Biochemistry* 1993; **209**: 32-44

Starovasnik MA, O'Connell MP and Fairbrother WJ. Antibody variable region binding by *Staphylococcal* protein A: Thermodynamic analysis and location of the Fv binding site on E-domain. *Protein Science* 1999; **8**: 1423-1431

Surolia A, Pain D and Khan MI. Protein A; Nature's universal anti-antibody. *TIBS* 1982: 74-76

Swaminathan R, Krishnamoorthy G and Periasamy N. Similarity of fluorescence lifetime distributions for single tryptophan proteins in the random coil state. *Biophysical Journal* 1994; **67(5)**: 2013-2023

Tanford C. Protein denaturation. *Advances in Protein Chemistry* 1968; **23**: 121-282

Tashiro M and Montelione GT. Structures of bacterial immunoglobulin-binding domains and their complexes with immunoglobulins. *Current Opinion in Structural Biology* 1995; **5**: 471-481

Thurlkill RL, Grimsley GR, Scholtz M and Pace CN. pK values of the ionisable groups of proteins. *Protein Science* 2006; **15**: 1214-1218

Tonegawa S. The molecules of the immune system. *Scientific American Journal* 1985; **253**: 122-130

Trevor P. Creamer, Rajgopal Srinivasan, and George D. Rose Modeling Unfolded States of Proteins and Peptides. II. Backbone Solvent Accessibility. *Biochemistry* 1997; **36**: 2832-2835

Torigoe H, Shimada, Saito A, Sato M and Arata Y. Sequential ^{1}H NMR assignments and secondary structure of the B domain of staphylococcal protein A: Structural changes between the free B domain in solution and the Fc-bound B domain in crystal. *Biochemistry* 1990; **29**; 8787-8793

Uhlén M, Lindberg M and Philipson L. The gene for staphylococcal protein A. *Immunology Today* 1984a; **5(8)**: 244-249

Uhlén M, Guss B, Nilsson B, Gatenbeck S, Philipson L and Lindberg M. Complete sequence of the staphylococcal gene encoding protein A. A gene evolved through

multiple duplications. *The Journal of Biological Chemistry* 1984b; **259(3)**: 1695-1702

Ullrich A, Shine J, Chirgwin J, Pictet R, Tischer E and Rutter WJ. Rat insulin genes: Construction of plasmids containing the coding sequences. *Science* 1977; **196**: 1313-1319

Verway WF. A type-specific antigenic protein derived from the staphylococcus. *Journal of Experimental Medicine* 1940; **71**: 635-644

Viguera AR, Vega C and Serrano L. Unspecific hydrophobic stabilization of folding transition states. *Proceedings of the National Academy of Science (USA)* 2002; **99(8)**: 5349-5354

Vogelstein B and Gillespie D. Preparative and analytical purification of DNA from agarose. *Proceedings in the National Academy of Sciences (USA)* 1979; **76**: 615-619

Walker KN. Protein engineering and characterisation of an IgG-binding domain based upon protein G from *Streptococcus* Group G. PhD thesis 1994; University of Southampton

Walker KN, Bottomley SP, Popplewell AP, Sutton BJ and Gore MG. Equilibrium and pre-equilibrium fluorescence spectroscopic studies of the binding of a single-immunoglobulin-binding domain derived from Protein G to the Fc fragment from human IgG1. *Biochemistry Journal* 1995; **310**: 177-184

Ward ES, Güssow D, Griffiths AD, Jones PT and Winter G. Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. *Nature* 1989; **341**: 544-546

Weber G and Teale FWJ. Ultraviolet fluorescence of aromatic amino acids. *Biochemistry Journal* 1957; **65**:476-482

Wen-Chuan K, Kuo-Yu L, Gwo-Jen J, Hui-Kang T and Chien C. Simultaneous measurement of phase retardation and fast-axis angle of phase retardation plate. *Japanese Journal of Applied Physics* 2004; **44**: 1095-1100

Wikström M, Sjöbring U, Kastern W, Björck L, Drakenberg T and Forsén S. Proton nuclear magnetic resonance sequential assignments and secondary structure of an immunoglobulin light chain-binding domain of protein L. *Biochemistry* 1993; **32**: 3381-3386

Wikström M, Drakenberg T, Forsen S, Sjöbring U and Björck L. Three-dimensional solution structure of an immunoglobulin light chain-binding domain of protein L. Comparison with the IgG-binding domains of protein G. *Biochemistry* 1994; **33**: 14011-14017

Wikström M, Sjöbring U, Drakenberg T, Forsen S and Björck L. Mapping of the immunoglobulin light chain-binding site of protein L. *Journal of Molecular Biology* 1995; **250**: 128-133

Wikström M, Forsén S and Drakenberg T. Backbone dynamics of a domain of protein L which binds to immunoglobulin light chains. *European Journal of Biochemistry* 1996; **235**: 543-548

Woody RW. Aromatic side-chain contributions to the far-ultraviolet circular dichroism of peptides and proteins. *Biopolymers* 1978; **17**: 1451-1467

Woody RW. Contribution of tryptophan side-chains to the far-ultraviolet circular dichroism of proteins. *European Biophysical Journal* 1994; **23**: 253-262

Woody RW. Circular dichroism. *Methods in Enzymology* 1995; **246**: 34-71

Worrall AF and Connolly BA. *Journal of Biological Chemistry* 1990; **265(35)**: 1889 -21895

Yon JM. Protein folding: Concepts and perspectives. *Cellular and Molecular Life Science* 1997; **53(7)**: 557-567

Xu D, Tsai CJ and Nussinov R. Hydrogen bonds and salt bridges across protein-protein interfaces. *Protein Engineering* 1997; **10(9)**: 999-1012

Zimmerman SB and Pheiffer BH. Macromolecular crowding allows blunt-end ligation by DNA ligases from rat liver or *Escherichia coli*. *Proceedings in the National Academy of Sciences (USA)* 1983; **80**: 5852-5856