Biophysical Studies of the Interaction Between High Molecular Weight Kininogen and gC1q-R : Initiating the Contact Pathway.

Alex Slater, MSci. (Hons).

Thesis submitted to The University of Nottingham
for the degree of Doctor of Philosophy

April 2017

Abstract

The contact pathway is part of the coagulation pathway involved in haemostasis. It is responsible for the deposition of fibrin fibres during blood coagulation as well as the production of the inflammatory mediator bradykinin. This pathway is initiated when the three enzymes prekallikrein, factor XI and factor XII are assembled at the cell surface triggering a cleavage cascade that results in the deposition of fibrin. High molecular weight kininogen is a cofactor responsible for the presentation of Factor XI and prekallikrein to the cell surface where these proteins come into contact with the cell bound FXII. Both FXII and high molecular weight kininogen bind to endothelial cells through the monotrimeric endothelial cell receptor gC1q-R in a zinc-dependent manner.

Domain 5 is the cell surface binding domain of high molecular weight kininogen. Studies were performed on isolated domain 5 which showed it to be a zinc binding, intrinsically disordered domain that binds directly to trimeric gC1q-R. Isothermal titration calorimetry revealed that gC1q-R binding occurs through a sequential binding mechanism where three domain 5 ligands bind to one trimer. The production of N and C-terminal truncations of domain 5 revealed a zinc-dependent N-terminal binding region and a zinc-independent C-terminal binding region that simultaneously bind separate sites on gC1q-R. The C-terminal binding region was further narrowed down to a Lys rich portion of domain 5. Further mutation studies on gC1q-R revealed that the $\beta 6$-$\beta 7$ loop, located towards the centre of the cavity, is crucial for D5 binding.

The findings highlighted within this thesis provide structural and mechanistic detail into the complex interaction between domain 5 and gC1q-R. The dual binding of high molecular weight kininogen and factor XII with this receptor is discussed in order to further understand how gC1q-R assembles the contact initiator proteins at the cell surface.

Acknowledgements

I would firstly like to thank my supervisors Prof. Mark Searle and Prof. Jonas Emsley for giving me the opportunity to work on this project and for their support and guidance throughout my PhD. I would also like to thank Bubacarr Kaira for all his help with the X-Ray crystallography and his constant collaboration throughout this project. I am thankful to Dr Jed Long for all his guidance within the laboratory and Dr Huw Williams for his help regarding all the NMR. I am grateful to Dr. Dan Scott and Jeddidiah Bellamy Carter for all their help with mass spectrometry experiments. An additional thank you to Dr. Susan Wong for her help with binding studies. A big shout out to all members of the MSS, PS, NRT and CL groups who have not only had to put up with my singing for over 3 years but provided an enjoyable and supportive environment to work in. Finally I would like to thank my family and friends for all their support and encouragement.

Abbreviations

aa- Amino acid

APS- Ammonium persulfate

AGF- Analytical gel filtration

β **receptors**- β adrenergic receptor

C1q- Complement component 1q

CD- Circular dichroism

cm- Centimeter

D1-6- Domain 1-6 of high molecular weight kininogen

DNA- Deoxyribonucleic acid

dNTP- Deoxynucleotide triphosphate

E. coli- Escherichia coli

EDTA-Etylenediaminetetraacetic acid

ESI- Electrospray ionisation

FPLC-Fast protein liquid chromatography

FnII- Fibronectin-type II

GPIb-Glycoprotein Ib

GST- Glutathione S-transferase

HF- High Fidelity

HMWK- High molecular weigh kininogen

HSQC- Heteronuclear single quantum correlation

iC3b- Inactive complement component 3b

I-domain- Inserted domain

IPTG- Isopropyl -D-1- thiogalactopyranoside

ITC- Isothermal titration calorimetry

kDa- Kilo Dalton

KLD- Kinase, ligase, Dpn1

LB- Lysogeny broth

m- Meter

MAC-1- Macrophage-1 antigen

M.R.E- Mean residue ellipticity

mAu- Milli absorbance units

mins- Minutes

mS- Millisiemens

MWCO- Molecular weight cut off

NMR- Nuclear magnetic resonance

OD- Optical density

PCR- Polymerase chain reaction

ppm- Parts per million

rcf-Relative centrifugal force

RNA- Ribonucleic acid

rpm- Revolutions per minute

RT- Room temperature

secs- Seconds

SDS PAGE- Sodium dodecyl sulphate poly acrylamide gel electrophoresis

Ta- Association temperature

TEMED- Tetramethylethylenediamine

TOF- Time of flight

Tris- Tris (hydroxymethly) aminomethane

vWF- Von Willebrand factor

U-Units

UV-Ultraviolet

1D-One dimensional

2D-Two dimensional

Amino Acid	Three Letter Code	Single Letter Code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamate	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Table 1: *Three letter and single letter abbreviations for the 20 naturally occurring amino acids.*

Contents

CHAPTER 1

An Introduction to High Molecular Weight Kininogen and the Contact Pathway

The function of living cells requires the interplay and communication between all the individual cellular components. Complex networks of protein-protein interactions (PPIs) are essential for all biological processes [1], whether it be enzyme catalysis, DNA synthesis, metabolism, antigen recognition or signal transduction. The study of PPIs, therefore, is crucial for the understanding of any branch of biology. Furthermore, PPIs play a pivotal role in the fight against disease as the manipulation and control of these interactions enables the prevention of a wide range of diseases and cellular dysfunctions [2].

Haemostasis is reliant on the interaction of proteins from a variety of different cell types and is responsible for the prevention of excessive bleeding after injury through the formation of an insoluble blood clot also known as a haemostatic plug. Thrombosis is where a blood clot begins to obstruct blood flow within a blood vessel, and drugs targeting key events of thrombus formation

are important for the treatment of a range of cardiovascular disorders [3]. The work within this thesis focuses on specific PPIs involved in the activation of haemostasis which may prove invaluable for the design of novel therapeutics for a variety of blood disorders.

1.1 Haemostasis and Thrombosis

Haemostasis is the process that prevents excessive bleeding and tissue infection during vascular injury. It is responsible for the recruitment of platelets to the site of injury; coagulation of the blood, through the depostition of fibrin; and the recruitment of white blood cells to the clot. Upon endothelial cell damage, platelet adhesion to the subendothelium is mediated through their interaction with von Willebrand factor bound to the subendothelium [4]. This is combined with the coagulation pathway where activation of prothrombin to thrombin [5] and subsequent conversion of fibrinogen to fibrin [6] results in the production of fibrin fibres [7]. The combination of fibrin fibres and platelet aggregates results in the formation of a haemostatic plug that prevents excessive bleeding [8]. The production of activated thrombin during the coagulation cascade also further induces the activation of platelets and their aggregation to the site of injury [9]. Another key event during haemostasis is the recruitment of leukocytes to the haemostatic plug preventing infection at the site of injury. This is achieved through the interaction of leukocytes with activated platelets [10] [11] and fibrin [12].

The mechanisms involved in controlling the formation of the haemostatic plug are kept under tight control, but when activated haemostasis overwhelms the regulatory controls, pathologic thrombosis occurs [13] which is the blocking of blood vessels through the formation of a thrombus. The most common form of thrombosis is deep vein thrombosis which is largely caused by endothelial

injury within the deep veins, particularly within the legs and pelvis [14], and this can result in pulmonary embolisms [15]. Other classes of thrombosis include portal vein thrombosis [16], renal vein thrombosis [17], and arterial based thrombosis including thrombotic stroke and myocardial infarction [18].

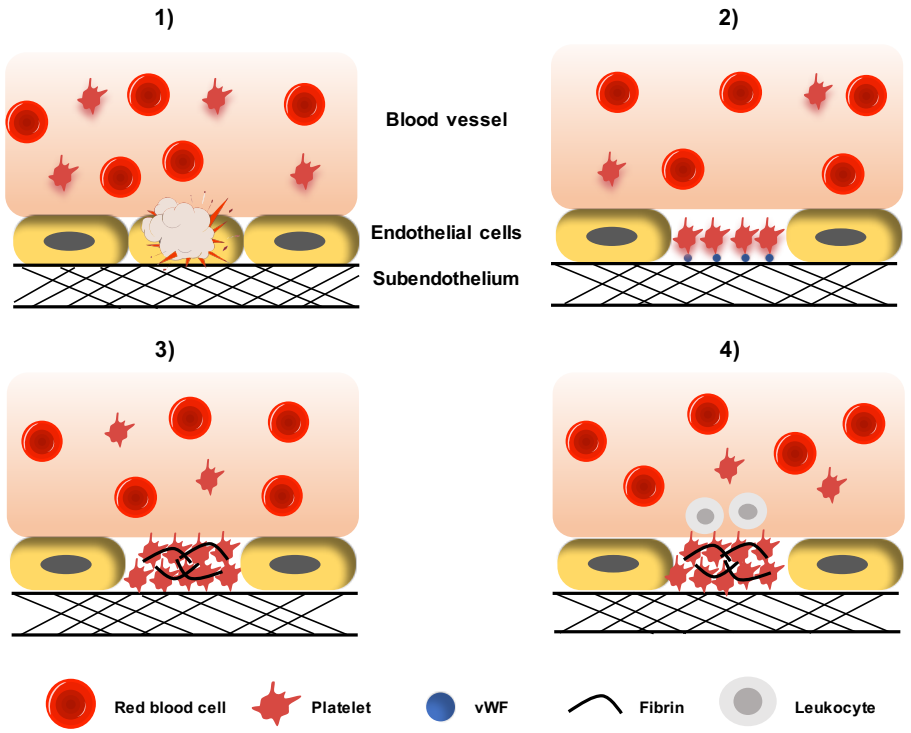


Figure 1.1: *Diagram representing the major steps involved in haemostasis.* ***1)*** *Damage of the endothelium exposes the subendothelium to the inside of the blood vessel.* ***2)*** *Interaction of circulating platelets with vWF results in platelet aggregation at the site of injury.* ***3)*** *Activation of the coagulation pathway results in the formation of fibrin fibres.* ***4)*** *Recruitment of leukocytes to the site of injury through their interaction with both platelets and fibrin.*

1.1.1 The Coagulation Pathway

The coagulation pathway is comprised of the intrinsic and extrinsic pathways. Both pathways involve the cleavage and activation of enzyme precursors, zymogens, resulting in the conversion of fibrinogen to fibrin and subsequent formation of fibrin fibers [19]. The extrinsic pathway is believed to be responsible for the initiation of blood coagulation, whereas the intrinsic pathway is respon-

sible for the amplification of this pathway.

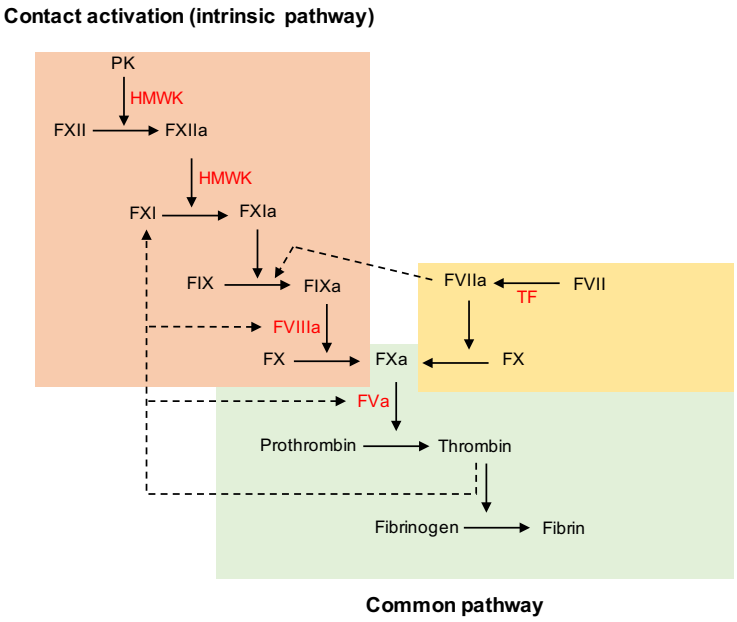


Figure 1.2: *Simplified diagram outlining the key enzymatic steps of the coagulation pathway, resulting in the activation of both thrombin and fibrin. This pathway is separated in to the intrinsic, extrinsic and common branches. Cofactor proteins required for specific activation steps are highlighted in red, and the dotted lines represent additional activation events that occur between the three branches of the coagulation pathway. This image was adapted from Peters et al [20].*

1.1.1.1 The Extrinsic Pathway

The extrinsic pathway, also known as the tissue factor (TF) pathway, involves the activation of FX by FVIIa in a TF dependent manner. TF is expressed within smooth muscle cells [21], fibroblasts [22] and pericytes [23]. TF is exposed upon the disruption of vascular integrity [24] where it forms a complex with FVIIa and activates FX [25] [26], resulting in the subsequent activation of thrombin.

1.1.1.2 The Contact Pathway

The contact pathway, also referred to as the intrinsic pathway, is responsible for a number of patho-physiological responses such as the inflammatory response [27], fibrinolysis [28] [29] and blood coagulation [30]. It triggers a protein cleavage cascade which, similarly to the extrinsic pathway leads to the activation of thrombin and fibrin to form the haemostatic plug. As well as having roles in the coagulation pathway, the contact pathway is now understood to act as an interface between coagulation, immunity and inflammation [31] [32] which is discussed in more detail throughout this chapter.

The four main proteins involved in the activation of the contact pathway are high molecular weight kininogen (HMWK), prekallikrein (PK), factor XI (FXI) and factor XII (FXII) [33]. These proteins circulate as zymogens in the blood, and contact activation is initiated by the autocatalytic cleavage of FXII to produce the active FXIIa [34]. FXII activation is stimulated upon binding to a range of negative surfaces including collagen, nucleic acids, and platelet and microbial polyphosphate [35], but the physiological negative surface for contact activation is the cell membrane. FXIIa is able to cleave both FXI and PK to the active FXIa and kallikrein, both of which circulate in complex with HMWK [36]. In the presence of HMWK, kallikrein is also able to cleave FXII, in a positive feedback mechanism, that further activates the contact pathway [37]. FXIa triggers the endogenous clotting cascade, whereas activated kallikrein cleaves HMWK [38], resulting in the release of the nonapeptide bradykinin (BK) and the two chained protein HKa [39] . HKa has enhanced negative surface binding compared to uncleaved HMWK and increases the transport of zymogens to the required surface [40].

The assembly of contact initiator proteins has been demonstrated on the membranes of a variety of cell types including endothelial cells [41], neutrophils

[42], bacterial cells [43] [44] and viruses [45]. On the surface of endothelial cells a number of membrane receptors have been identified to be involved in the assembly of both FXII and HMWK [46], and FXII deficient plasma was shown to reduce thrombin formation in cultured endothelial cells, outlining the role of endothelial cells in contact activation [47]. The mechanisms responsible for the activation and control of FXII surface binding are not well understood but is likely due to increased expression and presentation of key receptors at the cell surface [48]. Additionally to cell surfaces, the contact pathway is activated by poly phosphate [49], neutrophil extracellular traps [50], DNA [51], mis-folded protein aggregates [52], heparin [53] and collagen [54]. Due to these activation surfaces, contact activation is amplified upon cell injury or death and during infection.

Despite its role in fibrin formation, the contact pathway is not critical for maintaining haemostasis. This is apparent as deficiencies in contact activator proteins does not present with any bleeding deficiencies [55]. Therefore, the major physiological role of the contact pathway remains unclear, but it has been proposed to be the formation of bradykinin in the inflammatory response [56]. However, the contact pathway is known to play major roles in the pathophysiological response to injury resulting in thrombosis [31]. This is partially due to the large number of surfaces able to activate the contact pathway, and uncontrolled activation of this pathway can result in thrombus formation. Pathophysiological activation of the contact pathway can be initiated by major vessel damage exposing key receptors on the endothelium, bacterial or viral infection and the presence denatured protein aggregates [57]. The link between the contact pathway and thrombus formation has been demonstrated in a variety of animal models, and targeting key contact factors results in reduced thrombosis with minimal effects on haemostasis [58] [59] [60] [61]. Furthermore, contact activation has been linked with venous thrombosis due to the observation that these particular thrombi tend to be low in platelets but rich in fibrin [62]. This

makes the contact pathway a key target for the prevention of thrombosis; therefore, understanding the key events involved in the activation of this pathway is crucial for the development of anti-thrombotic therapies.

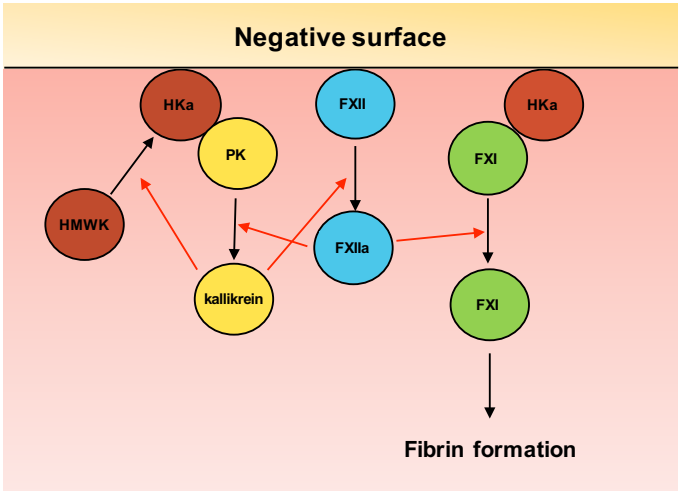


Figure 1.3: *Contact activation involving the four initiator proteins: FXII, FXI, PK and HMWK. Black arrows indicate the conversion of zymogens to the active form, and the red arrows indicate the proteins responsible for activation. The pathway is initiated by the auto activation of FXII at a negative surface. Negative surfaces able to activate FXII including cell membranes, collagen, nucleic acids and polyphosphate. This is followed by the activation of both PK and FXI which are brought to the cell surface through the interaction with HMWK/HKa. Kallikrein cleaves HMWK to HKa which further increases cell surface binding. The activation of FXI triggers a protein cleavage cascade that results is the activation of fibrin.*

1.2 High Molecular Weight Kininogen

This thesis focuses on the interaction of HMWK with the endothelial cell surface and the potential mechanisms involved in contact initiation. Kininogen exists as a high and low molecular weight isoform produced by alternative splicing of the KNG1 gene [63]. HMWK is a non-enzymatic, single chain glycoprotein with a mass of around 120 kDA [64]. It was first identified as a precursor for the nonapeptide bradykinin, but over the past 20 years other biological functions

of HMWK have been explored, including its role in coagulation and innate immunity [31]. HMWK is an abundant plasma protein at 670 nM, and in the blood it circulates in complex with either prekallikrein or Factor XI [65] [66] [36]. The biological roles, structure and interactions of HMWK with other proteins and surfaces will be discussed in this chapter.

1.2.1 Structure of HMWK

There is currently no high resolution structure of HMWK or any of its isolated domains, but electron microscopy has provided insight into the domain orientation and general shape of full length HMWK [67]. HMWK is made up of six domains (D1-6) and forms a ring like structure due to a single dislulphide bond, between Cys28 and Cys614, linking the amino and carboxy terminus . The cleavage of HMWK by kallikrein results in a two chained protein consisting of a heavy chain and light chain. D1-3 are part of the heavy chain and are cystatin-like domains. D4 completes the heavy chain and consists of the small peptide bradykinin [68]. D5 and D6 form the light chain and are involved in surface binding and the binding to PK and activated FXII respectively [69] [70]. Due to its involvement in surface binding and interactions with other proteins in the contact pathway, the light chain of HMWK is the major contributor to the coagulation activity. A cartoon representation of the HMWK shape and domain orientation is shown in Figure 1.4, and the full aa sequence of HMWK can been found in section 9.1.1.1.

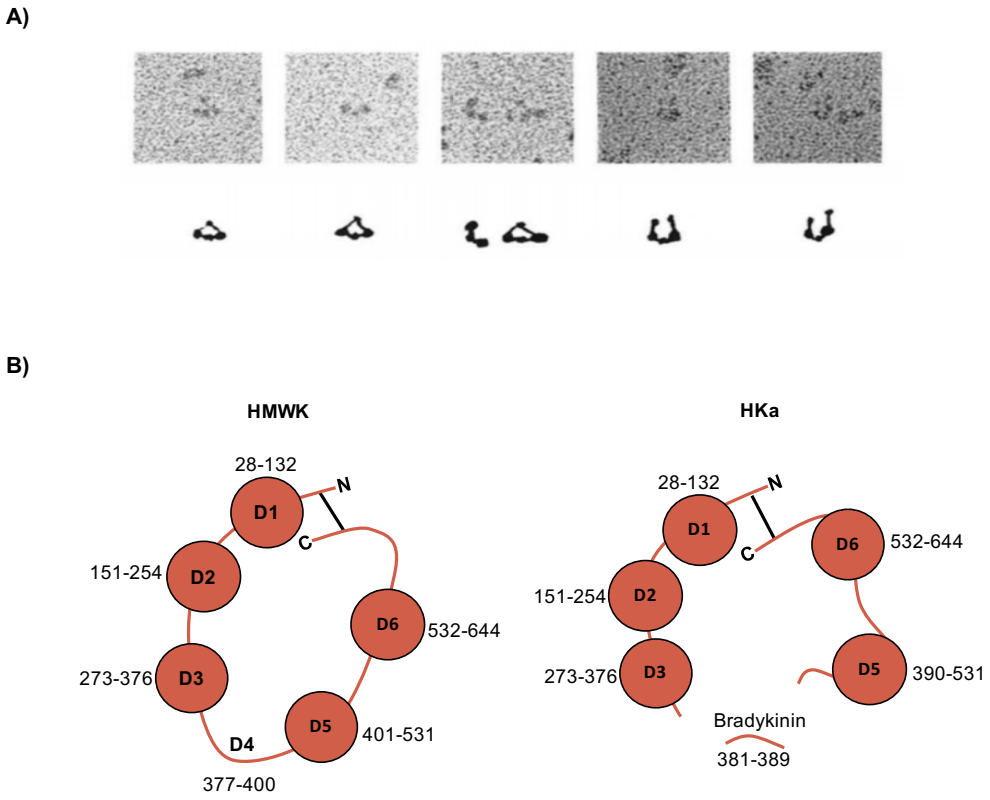


Figure 1.4: ***A)*** *Electron microscopy images, imaged by Weisel et al [67], providing evidence for the shape and orientation of HMWK with its globular domains connected through flexible linker regions in a ring structure.* ***B)*** *A cartoon representation for the structure of HMWK illustrating the ring like structure caused by a disulphide bond between Cys28 and Cys614. Shown on the left is uncleaved HMWK, and shown on the right is the active HKa, produced by cleavage by kallikrein, and the released bradykinin. The domain boundaries are shown next to each domain.*

1.2.2 Activated HMWK

Full length HMWK is cleaved by PK, resulting in the formation of HKa. HKa consists of a 64 kDa heavy chain and a 56 kDa light chain, with the release of the peptide bradykinin from D4 [40]. Cleavage of HMWK by kallikrein is a two-step process [71], with the first cleavage occurring at Lys380-Arg381 and the second at Arg389-Ser390, releasing the small peptide bradykinin. This produces

a disulphide-linked 64 kDa heavy chain and 56 kDa light chain [72]. HKa has been shown to have enhanced surface binding properties [40], but the full significance of HMWK activation is poorly understood, and a more crucial role for HKa formation is likely the release of bradykinin. HMWK can also be cleaved by FXIa, at a much slower rate than by kallikrein, and prolonged exposure with FXIa results in degradation of the light chain producing an inactive product [73]. This acts as a control mechanism for the contact pathway as FXIa can inactivate HMWK and inhibit contact activation. In FXII deficient plasma there is no detectable cleavage of HMWK which is consistent with FXII's involvement in the activation of both PK and FXI [38].

1.2.3 Function of the Heavy Chain

D1-3 of the heavy chain are evolutionarily derived from cystatin. Both D2 and D3 inhibit the cysteine protease papain due to the presence of the conserved QVVAG sequence found within both domains [72]. Additionally, D2 also has the ability to inhibit the calcium dependent cysteine protease calpain [74]. D3 has surface binding properties and has been shown to bind endothelial cells [75], neutrophils [76] and platelets [77].

1.3 Interactions and Function of HMWK D5

D5, similarly to D3, binds to a variety of cell surfaces, including endothelial cells [78] and neutrophils [76]. Despite D3 also having neutrophil and endothelial cell binding properties, D5 is regarded as the major surface binding domain for these surfaces. The additional binding of D3 may serve to increase the binding affinity of HMWK to these cell types. The amino acid sequence of D5 is shown

below:

401 T V S P P H T S M A P A Q D E E R D S G K E Q G H T R R H D W G H E K Q R K H N L G H G H K H E R D Q **G H G H Q R G H G L G H G H E Q Q H G L G H G H** K F K L D D D L E H Q G G H V L D H G **H K H K H G H G H G K H K N K G K K N G K** H N G W K T E H L A S S S E D S 531

It has a His-rich sequence with two distinct regions: a His-Gly-rich region, highlighted in red; and a C-terminal His-Gly-Lys-rich region, highlighted in blue. The Lys-rich region makes D5 a basic domain which is consistent with its ability to bind negative surfaces. D5 shows very little sequence homology with any other protein domains, but the D5 sequence is relatively well conserved between species.

1.3.1 Negative Surface Binding

Contact activation can be stimulated by a number of biologically relevant negative surfaces, including DNA [51] [79] and polyphosphate [80]. Furthermore, a number of different artificial negative surfaces, such as dextran sulphate [81] [82] and kaolin [83], also enhance contact activation. D5 is the negative surface binding domain of HMWK [70] and is therefore a key domain for the recruitment of HMWK to negative surfaces and activation of the contact pathway. Peptide studies, attempting to narrow down the key D5 region involved in cell binding, identified that both the His-Gly and the His-Gly-Lys-rich regions bind negative surfaces [72], providing evidence for multiple binding regions within D5. Within the His-Gly-rich region, the HGLGHGHEQQHGLGHGH sequence serves as a primary structure able to bind negative surfaces, and binding was further increased in the presence of Zn^{2+}[84]. Although the His-Gly-rich regions have been shown to be involved in the negative surface binding ability of HMWK, it is not known how the different regions co-operate or combine together to bind surfaces; this is a key objective throughout this thesis.

1.3.2 Zinc Binding Ability of D5

HMWK's ability to bind negative surfaces and initiate coagulation has been shown to be zinc dependent [85][86]; due to the His-rich sequence of D5, it has been proposed that this domain directly binds Zn^{2+}. The negative surface binding peptide HGLGHGHEQQHGLGHGH was shown to play a role in zinc binding [84], and other His-rich regions throughout the rest of D5 could potentially serve as additional zinc binding sites. Zinc dependence has also been identified in FXII [87], and it is hypothesised the zinc might function to orientate the surface binding domains into an active, folded conformation that is necessary for surface binding. The stoichiometry of zinc binding to D5 and the key residues involved have not been determined, and the role zinc binding plays on the structure of D5 is largely unknown.

1.4 Bradykinin

Apart from its role as the intrinsic arm of the coagulation pathway, another major role of the contact pathway is the production of Bradykinin (BK). BK is a nine amino acid long peptide produced upon the cleavage of HMWK D4 by kallikrein [88]. This peptide has a number of biological roles, including the promotion of angiogenesis [89], inflammatory pain response [90] and innate immunity [91] [92]. The functions of BK are mediated through its interaction with the β1 and β2 receptors [93]. These receptors are pharmaceutical drug targets for inflammatory response [94] and cancer therapy [95]. The production of BK highlights the additional roles of the contact pathway in innate immunity and inflammation [96] [97].

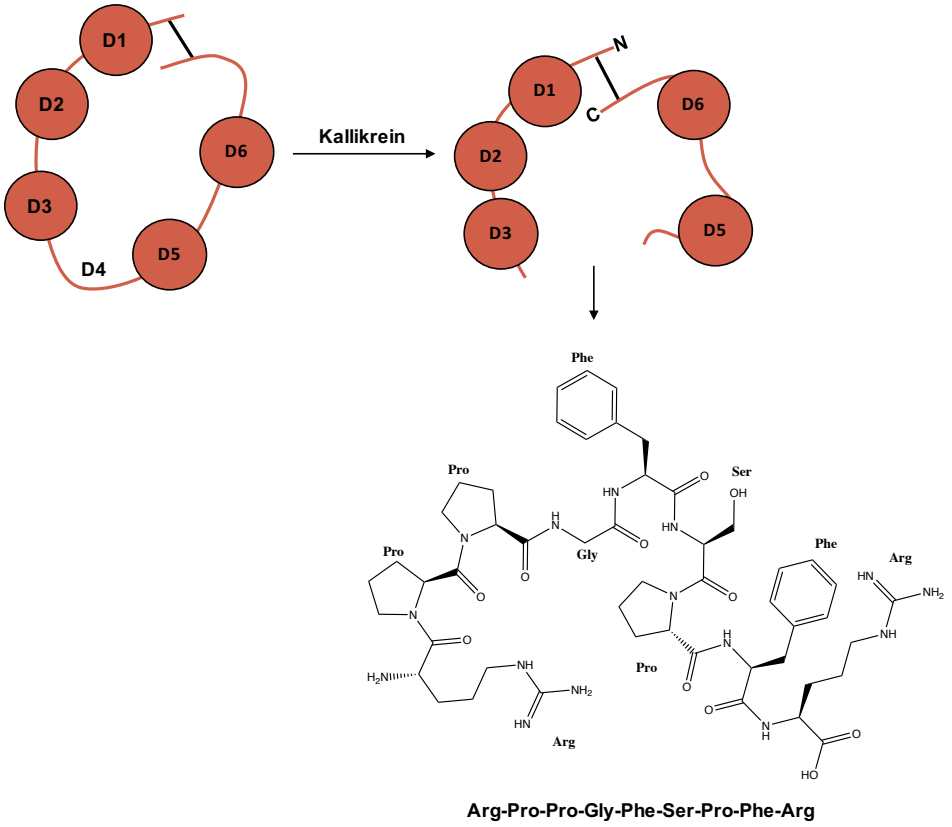


Figure 1.5: *Structure of the nonapeptide bradykinin produced upon the cleavage of HMWK D4 by kallikrein.*

1.5 Interactions of HWMK with Proteins of the Contact Pathway

HMWK binds to both PK and FXI with K_D values of 30 and 2.4 nM respectively [66], and the interaction of both these proteins with HMWK occurs through D6 [98]. Both FXI and PK have 4 conserved apple domains [99]; apple domain 2, and to a lesser extend 1 and 4, provides the binding sites for HMWK [100]. PK does not bind as strongly as FXI, which allows free kallikrein to activate unbound HMWK. The interaction of PK and FXI with HMWK is important for the presentation of these proteins, in the correct orientation, with FXII [37], resulting in the activation of PK and FXI.

1.6 HMWK Function and Receptors at Different Cell Types

Initially, it was thought that the haemostatic pathways and those controlling innate immunity were two separate entities; however, the interplay between both pathways is becoming more clear, and the mechanisms involved in linking both cellular functions are starting to be better understood (reviewed by Delvaeye et al [101]). HMWK is known to bind neutrophils [102], platelets [85] and endothelial cells [103] and plays a major role in the interplay between coagulation and immune response. Although this thesis focuses on HMWK and contact activation, some of the additional haemostatic roles of HMWK at various cell types will be discussed.

1.6.1 Endothelial Cells

At endothelial cells HMWK is involved in contact activation [41] and the regulation of cell proliferation, migration and angiogenesis [104] [105]. One of the major endothelial cell receptors that binds HMWK is gC1q-R, a receptor which binds both HMWK and FXII [106] [107]. This interaction of gC1q-R with both HMWK and FXII is believed to be crucial for the contact initiation on the surface of endothelial cells.

A second endothelial cell receptor shown to bind HMWK is the urokinase-type plasminogen activator receptor (uPAR) [108]. This receptor is responsible for the concentration of active plasmin at the cell surface, through the interaction with urokinase plasminogen activator (uPA) [109], resulting in the degradation of fibrin fibres and other cellular components during blood clot dissolution [110]. Through this mechanism, uPAR has been linked with a number of biological processes involving cell migration, including angiogenesis [111], tumour metastasis [112] and leukocyte migration [113]. HMWK inhibits endothe-

lial cell migration and proliferation, and angiogenesis [104] by disrupting the interaction of uPAR with uPA [114]. This disruption is achieved through the interaction of both D3 and D5 with uPAR [115].

HMWK also binds to cytokeratin 1 (CK1) on endothelial cells through D3, D4 and the Lys-rich D5 region [116]. Cytokeratins are a major component of skin, but this data was the first evidence that this family of proteins may also serve as a family of receptor proteins.

An SPR study looking at the binding of HMWK to all three of these endothelial receptors showed that HMWK binds strongest to gC1q-R (0.8 nM), followed by CK1 (15 nM) and then uPAR (2.3 μM), each in a zinc dependent manner [117]. It was also noted that both gC1q-R and uPAR showed no significant differences in binding affinity for HKa or HMWK, whereas uPAR bound 50 fold tighter to activated HKa. It was therefore proposed that gC1q-R and CK1 are involved in the initial binding of HMWK to the cell surface where it is then cleaved by kallikrein. HKa is then able to selectively bind uPAR and mediate cell migration. Antibodies targeting gC1q-R and CK1 inhibited HMWK binding to endothelial cells by 72 % and 30 % respectively [118] which suggested that gC1q-R is the major endothelial cell receptor for HMWK.

1.6.2 Neutrophils

HWMK has multiple functions at leukocyte surfaces that further outline its overlapping functions in coagulation and immune responses. The complete assembly of HMWK, PK, FXII and FXI has been shown to localise at the surface of neutrophils [42] with the role of creating a circulating platform for contact activation. The release of kinin peptides at the neutrophil surface has been proposed to enhance the passage of neutrophils out of the blood vessels through

the retraction of the endothelial cells; this is important for the mobilisation and exocytosis of neutrophil specific granule products needed for inflammatory responses [119]. HMWK is also essential for the presentation of PK to the neutrophil surface [74], resulting in neutrophil aggregation [27], chemotaxis [120] and the release of elastase through degranulation [121].

The major leukocyte cell receptor for HMWK is the integrin macrophage-1 antigen receptor (Mac-1) [76]. Integrins are receptors that mediate adhesive interactions with other cells and the extracellular matrix. Mac-1 is a multifunctional receptor expressed primarily on monocytes, macrophages, neutrophils and natural killer cells. A variety of different ligands have been documented to bind to this receptor, including fibrinogen [122], intercellular adhesion molecule-1 (ICAM-1) [123] and HMWK [76]. It has functions in immune defence by binding iC3b, resulting in the phagocytosis of erythrocytes [124] and elevated natural killer cell activity against iC3b coated cells [125]. Mac-1 also mediates a variety of cell to cell interactions, including neutrophil-endothelial cells [126], neutrophil-neutrophil adhesion [127] and neutrophil-platelet association [128].

HMWK mediates the adhesion of neutrophils to sites of fibrin formation and endothelial cells by inhibiting the interaction of Mac-1 with fibrinogen and ICAM-1 [129]. This is backed up by additional studies that showed HMWK inhibits the binding of fibrinogen to integrins on both neutrophils and activated platelets [130], preventing the adhesion of neutrophils to fibrinogen surfaces [131]. HMWK binds to Mac-1 via D3 of its heavy chain and D5 of the light chain [76]. By using peptides derived from both D5 and D3, the neutrophil binding site was linked with three non-contiguous peptides: Leu271-Ala277, of D3; and Gly442-Lys458 and Phe459-Lys478, from the His-Gly-rich region of D5 [132].

1.6.3 Platelets

At platelet surfaces HMWK has been shown to inhibit thrombin-induced platelet aggregation by preventing thrombin from binding to the platelet surface [133]. HMWK further mediates platelet activation and aggregation by inhibiting the cysteine protease calpain [74]. The major HMWK endothelial cell receptor gC1q-R is also present on platelet surfaces [134] [135] which has led to speculation that contact activation can also occur on activated platelets as well as endothelial cells.

GPIbα is a platelet cell receptor responsible for platelet recruitment and activation through the interaction of its α subunit with vWF [136] and thrombin [137]. HMWK binds GPIbα, through D3 [138], and functions as a molecular bridge between GPIbα and Mac-1 by simultaneously binding both proteins, through D3 and D5 respectively [139]. Therefore, the binding of HMWK functions to mediate the interaction between Mac-1 and GPIb, contributing to the regulation of leukocyte-platelet interactions. Furthermore, HMWK regulates the binding of thrombin to GPIb [140] and reduces platelet activation.

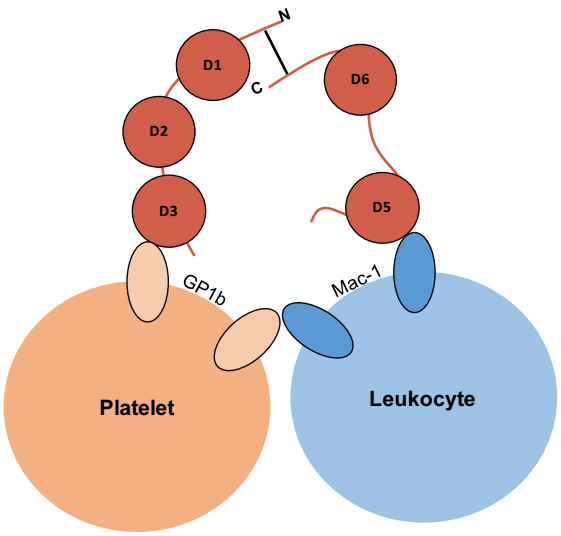


Figure 1.6: *HKa binding the leukocyte integrin Mac-1, through D5; and platelet glycoprotein GPIb, through D3, facilitating leukocyte recruitment at sites of vascular injury. This model can also be applied to full length HMWK.*

1.7 gC1q-R: the Major Endothelial Cell Receptor

gC1q-R was initially discovered to bind the globular heads of the C1q protein [106] but has since been shown to be involved in many different biological pathways, including adipogenesis and insulin signalling [141], thrombosis [142], regulation of RNA splicing [143], and the proliferation of tumour cells [144]. gC1q-R is also known to be involved in contact activation. Experiments have shown that gC1q-R promotes the production of kallikrein in the presence of FXII and HMWK, and the presence of antibodies for gC1q-R inhibited contact activation [48]. Combining experimental data with the fact gC1q-R is a highly negatively charged protein able to bind both HMWK and FXII, it clear that gC1q-R plays a key role in contact activation. Despite this, the interaction between HWMK and gC1q-R is poorly understood; therefore, this is the major interaction studied throughout this thesis.

gC1q-R is primarily localised to the mitochondrial membrane [145] [146] but

also localises at cell surfaces [147] and the nucleus [148]. This receptor contains an N-terminal mitochondrial signalling peptide which when blocked retargets the protein to the nucleus, cytoplasm and cell surface [149]. Potential mechanisms for redirecting gC1q-R localisation include alternative splicing of the primary transcript [150] or redistribution of differentially folded translation products [151]. gC1q-R has also been shown to form a complex with uPAR and CK1 on the surface of endothelial cells [46], and this interaction may be how gC1q-R associates with endothelial cell surfaces. Alternatively, gC1q-R could be released from the mitochondrial membrane upon apoptosis of the cell, and the resulting increase in gC1q-R presentation with plasma proteins would result in contact activation.

Both HMWK and FXII bind to this receptor in a zinc-dependent manner [106], and HMWK binding is believed to occur through D5 [107]. SPR studies calculated a K_D of 0.7-0.8 nM for the interaction of HMWK with gC1q-R, but no difference in binding affinity was detected between activated and non-activated HMWK [117]. It was also observed that HMWK in complex with PK was still able to bind to the receptor, providing evidence that PK is presented at endothelial cell surfaces directly through the interaction of HMWK and gC1q-R. Previous studies suggest that FXII and HMWK compete with one another for the binding to gC1q-R [107]; this is consistent with evidence that shows both these proteins compete for the same surface on HUVEC cells [152] [46]. The interactions of HMWK with FXI and PK, followed by the interaction of both FXII and HMWK with gC1q-R, could potentially be the key initiating steps for contact activation due to the concentration and assembly of the contact initiator proteins at one surface [153].

1.7.1 Structure of gC1q-R

The crystal structure of residues 74-282 of human gC1q-R was solved to 2.25 Å resolution [154] (Figure 1.7). The protein construct used was missing a 73 aa long N-terminal sequence which is involved in targeting the receptor to the mitochondrial membrane. The solved structure revealed that three monomers of gC1q-R form a symmetrical, homotrimeric ring structure with a central cavity of approximately 20 Å. Each monomer consists of seven β-strands (β1-7), which form an anti-parallel β-sheet; and three α helices (αA-C): one at the N-terminus and two at the C-terminus (Figure 1.8). The cavity is partially blocked on one face by loop regions connecting the β6 and β7 strands which reduce the size of the central cavity by 10 Å. There is an asymmetric charge distribution across the gC1q-R surface, with one face being highly negatively charged and the other face, despite containing polar regions, having an overall neutral charge. The difference in charge distribution suggests different functional roles of each face.

The subunit interface between monomers is held by interactions of the αB helix and β-sheet of monomer one with the αA and αC helices of monomer two. The helix-helix interaction is largely hydrophobic and forms a coiled coil structure. There are also intermolecular hydrogen bonds between residues found within disordered loop regions. A visual representation of the monomer-monomer interface interactions that give rise to the trimer is shown in Figure 1.9.

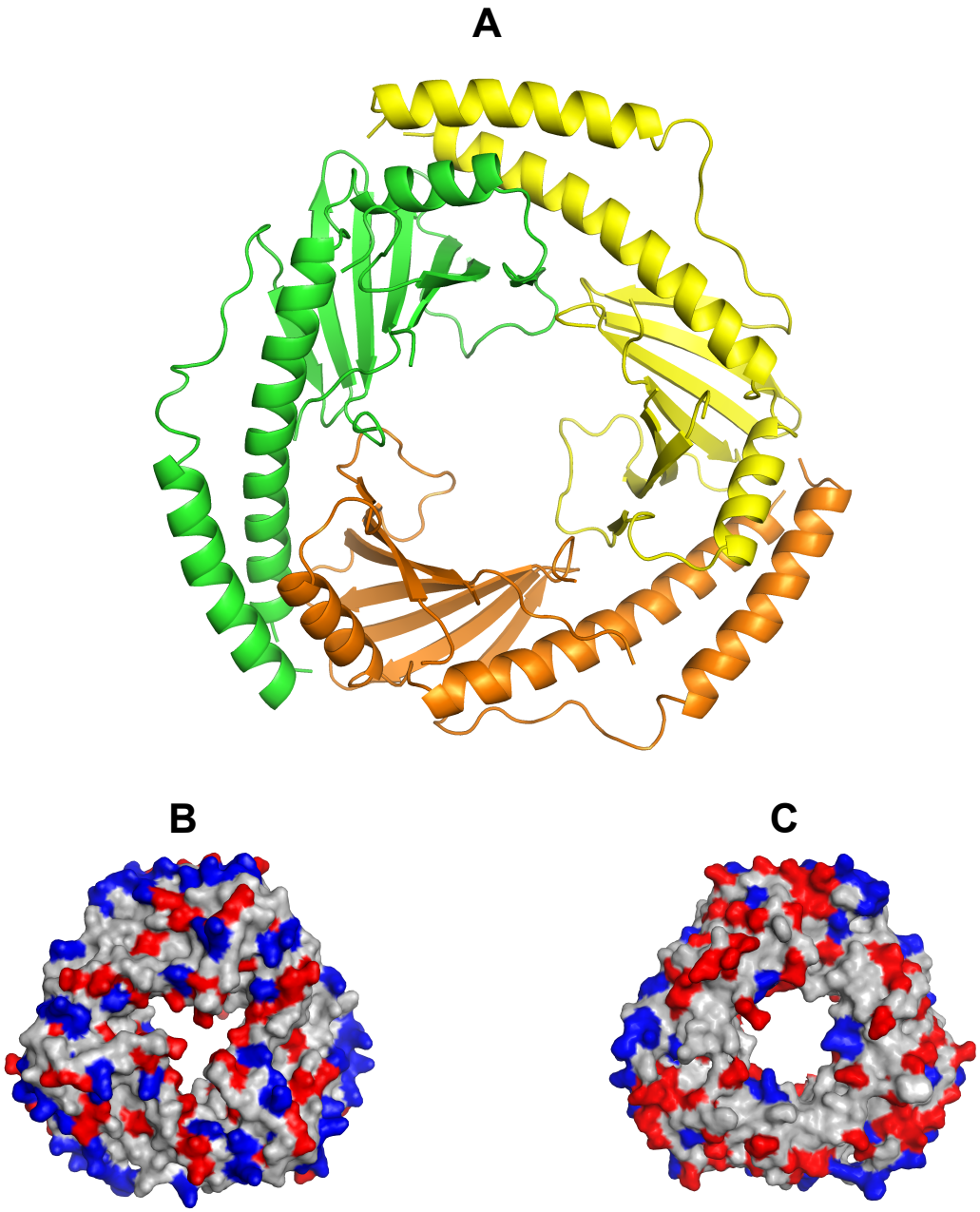


Figure 1.7: *Crystal structure of gC1q-R residues 74-282, PDB:1P32 [154].* ***A)*** *A cartoon representation of the three gC1q-R monomers, coloured green, yellow and orange, forming a mono-trimeric ring structure. Residues 141-161 and 191-196 from monomer 1 (green), 139-163 and 189-202 from monomer 2 (yellow), and 140-160 and 190-201 from monomer 3 (orange) are missing from the crystal structure and correspond to disordered loop regions.* ***B)*** *A surface image of the neutral face of gC1q-R with positive and negatively charged residues coloured in blue and red respectively.* ***C)*** *A surface image of the negative face of gC1q-R with positive and negatively charged residues coloured in blue and red respectively. The central loops present on the neutral face have been removed in this image to better visualise the increased diameter of the central cavity of this face.*

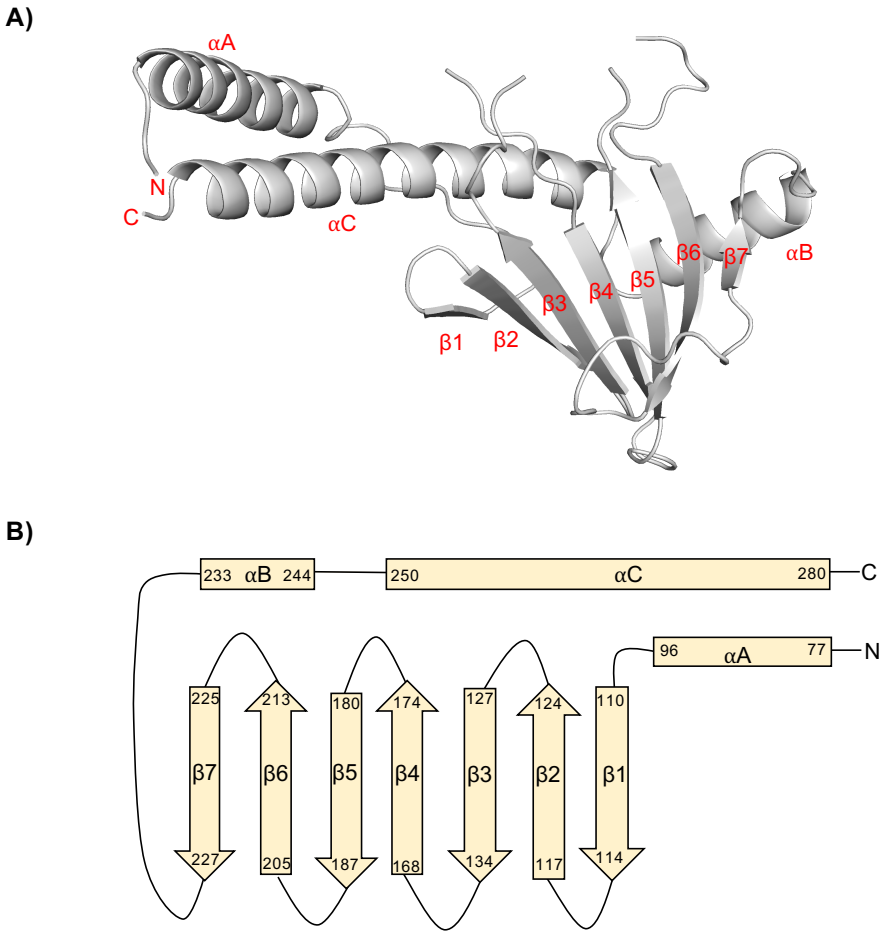


Figure 1.8: *Folding topology of gC1q-R.* ***A)*** *A cartoon representation of one monomer from the gC1q-R crystal structure with labelled secondary structure elements.* ***B)*** *Schematic diagram highlighting the gC1q-R topology and the positions of the major structural elements within peptide chain.*

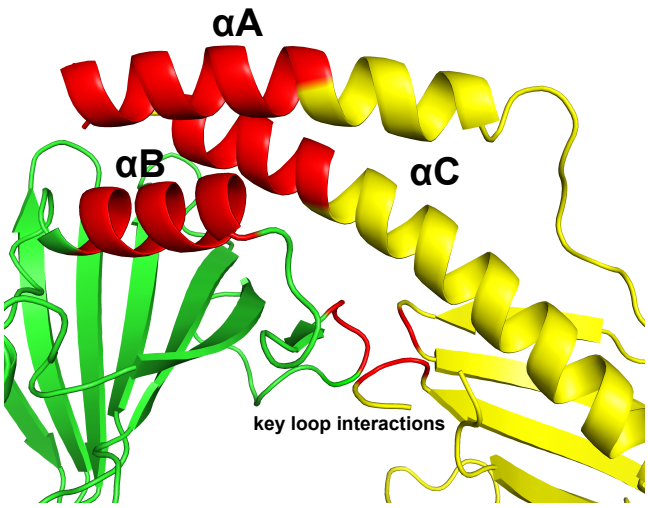


Figure 1.9: *The monomer-monomer interface of the gC1q-R trimer with the major interaction regions highlighted in red. The trimer is stabilised by the interactions of αA and αC of monomer two (yellow) with αB of monomer one (green). Additionally, loop regions from each monomer form intermolecular hydrogen bonds with one another further stabilising the trimer.*

Three mutations have been documented in the literature that prevent the formation of the gC1q-R trimer [155]. The first involved the deletion of the N-terminal αA helix. preventing the formation of the coiled-coil interaction with the αC helix of the adjacent subunit. Secondly, deletions of residues within the β6 strand and the connecting loop between β6 and β7 also prevented trimer formation. There was no observable interaction between HWMK and monomeric mutants suggesting trimer formation is vital for binding.

1.8 Aims and Objectives

There has been substantial research into the role of HMWK during contact activation, and a wide range of binding partners have been documented. Despite this, the mechanistic detail into these interactions is poor, and little is known of the binding interface between HMWK and the various receptors. Both D3 and D5 have been linked with surface binding properties, along with a series

of His-Gly-rich peptides derived from D5, but the key residues involved and the interplay between these binding regions has not been studied in detail. The mechanistic detail of these interactions is crucial for fully understanding contact initiation and the potential development of novel therapeutics targeting the inflammatory, immune and thrombotic responses.

Previous work within the group has focussed on the interaction of HMWK D5 with Mac-1 and the role it plays in mediating platelet leukocyte-interactions. Structural detail into the D5 binding region on Mac-1 provided further mechanistic insight to the role of D5 in leukocyte recruitment to the haemostatic plug. This thesis focuses on one of the alternate functions of HMWK, the assembly of the contact initiators at the cell surface through the interaction with gC1q-R.

The mechanisms for gC1q-R-dependent contact activation have never been fully explored, and it remains unknown how the interaction of FXII and HMWK with this receptor activates the contact pathway. It has been hypothesised that gC1q-R functions to localise both HMWK and FXII at the cell surface and, due to the trimeric nature of this receptor, gC1q-R could potentially bind both proteins simultaneously. This would provide a key multi-protein complex directly involved in contact initiation. Understanding these binding events at a molecular level would aid to further understand the key events of contact initiation and potentially highlight FXII, HMWK or gC1q-R as anti-thrombotic drug targets. Therefore, the overall aim of this research is to confirm the ability of gC1q-R to assemble the contact initiator proteins with major focus on the interaction HMWK with gC1q-R.

The aims of this study were to first detect the interaction between isolated D5 and gC1q-R and narrow down the key regions involved in binding. Although D5 is believed to be responsible for gC1q-R binding [107], the direct interaction of isolated D5 with gC1q-R has never been shown. These experiments also

involved the characterisation of isolated D5 and gC1q-R to understand their structure-function relationship. Attempts to narrow down both the gC1q-R binding site on D5 and the D5 binding site on gC1q-R were made in order to map the interaction sites on both proteins. Finally, the potential mechanisms for the assembly of contact initiators at the cell surface were explored.

CHAPTER 2

Biophysical Techniques

Biophysical techniques have proved one of the most robust ways of studying protein structure [156] [157] and protein interactions [158]. A wide range of biophysical techniques were used to characterise the structure and interactions of D5 and gC1q-R. The techniques used are discussed within this section.

2.1 Circular Dichroism

Circular dichroism (CD) spectroscopy has been widely used for the study of protein structure and folding (reviewed by S.M Kelly [159]) as well as a wide range of other biological macro molecules, including nucleic acids [160] and polysaccharides [161]. It measures the differential absorption of left and right handed circularly polarised light [159] which can be compared to a series of model spectra representing the various secondary structural elements in order to predict protein structure [162].

2.1.1 Principles of Circular Dichroism

Circularly polarised light (CPL) is produced when the two polarisation states of light are out of phase producing a helical light wave. A quarter-wave plate is responsible for the conversion of linearly polarised light to circularly polarised light and works by retarding one of the linear components, resulting in either a right handed or left handed circularly polarised beam (R-CPL or L-CPL).

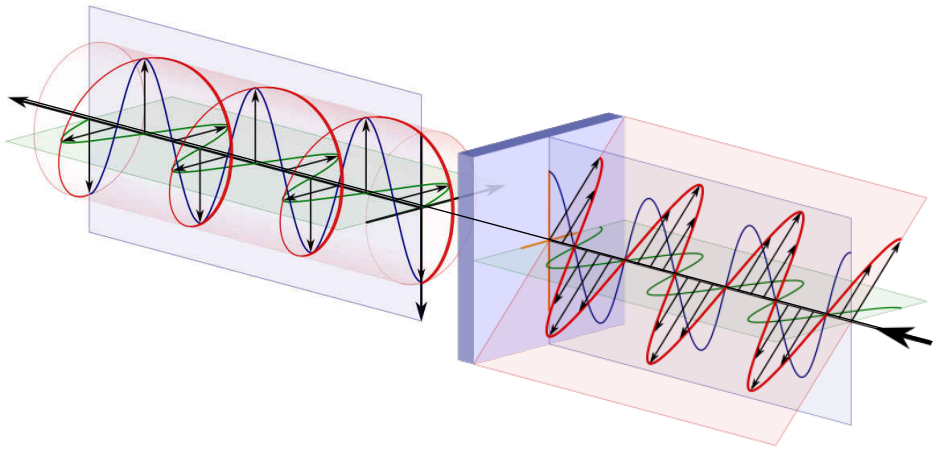


Figure 2.1: *Diagram showing the conversion of linearly polarised light to left handed circularly polarised light after travelling through a quarter wave plate. Blue and green curves represent the individual vertical and horizontal planes of light, whereas the red curve shows the sum of both planes. Image taken from https://en.wikipedia.org/wiki/User:Dave3457/list_of_the_related_images.*

A CD signal will be observed when R-CPL and L-CPL are absorbed at different extents which occurs when a molecule is chiral (optically active). Proteins are optically active for a number of reasons: firstly, 19 of the 20 naturally occurring amino acids have chiral centres at the carbon α, and secondly, and more importantly for CD of proteins, the secondary structural elements found within a protein give rise to asymmetry within the polypeptide chain [163]; therefore, each amino acid is in a unique chemical environment which gives a characteristic CD absorption. The peptide bond is the major chromophore of interest found within proteins and will absorb in the far-UV range of 250-180 nm. Ab-

sorption in this region is due to amide n -π* transitions, resulting in a broad signal centred at 222 nm; and π-π* transitions, resulting in an intense signal at 190 nm. These transitions are dependent on the ϕ and ψ angles, and the different secondary structures, consisting of α helix, β sheet and random coil elements, have well defined values for ϕ and ψ; therefore, far-UV CD profiles can be used to predict secondary structure [164].

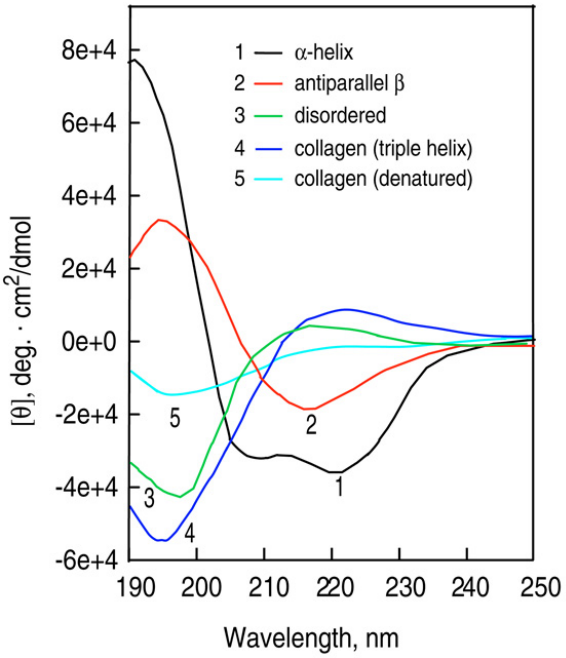


Figure 2.2: *Characteristic far-UV CD spectra for the different secondary structural elements found in proteins. Figure taken from N.J Greenfield [162].*

Largely α helical proteins have characteristic minima at 222 and 208 nm and a maximum at roughly 193 nm [165], whilst proteins with well defined anti-parallel β sheet regions have a minimum at 218 nm and a maximum at 195 nm [166]. Largely disordered proteins have low signal above 210 nm and negative bands at roughly 195 nm [167].

2.1.2 Monitoring Protein Stability by CD

As well as predicting the secondary structure of proteins, CD can be used to analyse the thermal and chemical stability of a protein's fold by running CD spectra at increasing temperatures [168] or in the presence of increasing concentrations of denaturant [169]. By plotting CD at a particular wavelength against temperature or denaturant concentration, protein unfolding curves can be generated. Combining CD spectra with these unfolding experiments allows for easy analysis of protein fold and stability. This is particularly useful when working on protein truncations or mutants, as protein mutations can result in the misfolding of proteins [170]; CD provides a quick and easy technique for confirming that mutant proteins have not misfolded.

2.2 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) allows for the complete characterisation of thermodynamic binding properties for a wide range of molecular interactions [171] and has received a huge degree of development and popularity for the study of protein interactions within the past two decades [172] [173] [174] [175]. ITC is able to detect changes in heat upon the injection of a ligand of a known concentration to a protein solution, and due to the fact heat is an intrinsic property of all biochemical reactions and interactions [176], no protein modifications or immobilisation steps are required; this makes ITC a more favourable technique when compared with other techniques such as SPR [177] and yeast-two-hybrid [178].

2.2.1 Principles of ITC

During an experiment any changes in heat within the sample cell is monitored and compared to a reference cell which is under constant heater power. The power required to keep the sample cell temperature constant with the reference cell is recorded. An exothermic or endothermic binding reaction results in negative or positive power spikes respectively, and these spikes decrease in intensity as binding sites becomes saturated within the cell. The integral of these spikes plotted against molar ratio of injectant produces a binding curve from which the stoichiometry (N), association constant (K_a), enthalpy change (ΔH), entropy change (ΔS) and free energy (ΔG) can be derived.

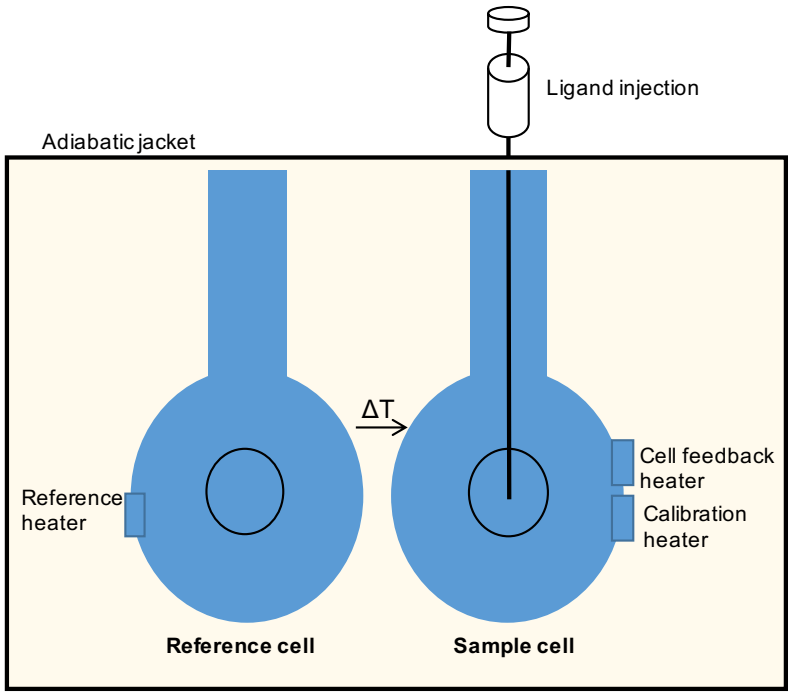


Figure 2.3: *An ITC calorimeter consists of a sample cell and a reference cell thermally coupled within an adiabatic jacket. Ligand is injected into the sample cell, and the heater power required to keep both cells at a constant temperature is monitored.*

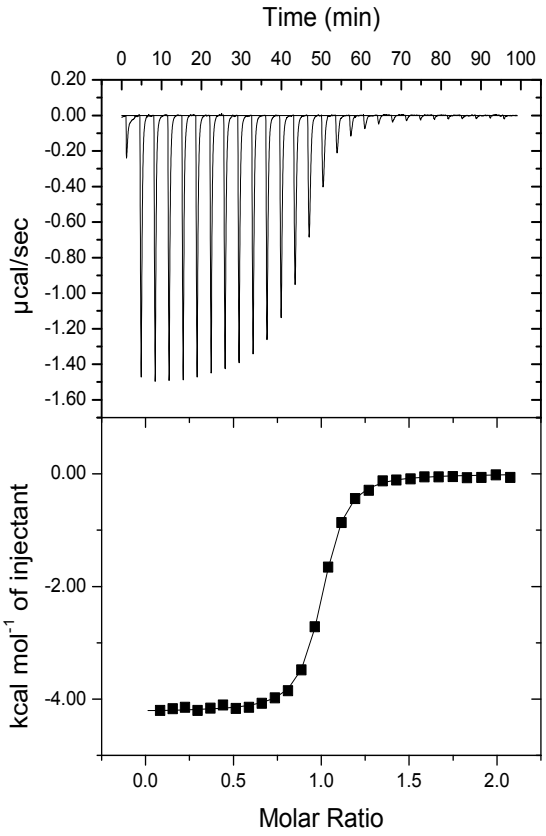


Figure 2.4: *ITC data for the titration of 1 mM $CaCl_2$ to 0.1 mM EDTA. The top panel represents the raw ITC data where heater power is plotted against time. Each negative spike represents an exothermic binding reaction, and as the species in the cell starts to become saturated the spikes get smaller until only buffer-buffer spikes are observed. The bottom panel shows the binding curve produced when each spike is integrated and energy is plotted against molar ratio. This binding curve is used to directly calculate ΔH, N and K_a.*

2.2.2 Deriving Thermodynamic Parameters

ΔH, N and K_a are derived directly from the binding curve. Negative and positive ΔH values correspond to exothermic or endothermic reactions respectively. The dissociation constant K_D is derived from K_a using equation 2.2.1. From these parameters ΔG and ΔS are derived using equations 2.2.2 and 2.2.3 respectively.

$$K_D = 1/K_a = \frac{[L] \times [M]}{[ML]} \tag{2.2.1}$$

Where [L] and [M] are the concentrations of unbound ligand and receptor respectively, and [ML] is the concentration of the bound species.

$$\Delta G = RT\ ln\ K_D \tag{2.2.2}$$

Where R is the gas constant, 8.314 J mol^{-1} K^{-1}, and T is the temperature in Kelvin. For a spontaneous reaction $\Delta G < 0$.

$$\Delta G = \Delta H - T\Delta S \tag{2.2.3}$$

Entropy represents the degree of disorder within a system; a positive ΔS is consistent with the release of solvent, whereas a negative ΔS shows structural confinement. ΔS and ΔH values can be used to understand the nature of a particular protein-ligand interaction [179]. Furthermore, ITC offers a sensitive method for calculating binding affinities which, when combined with protein mutagenesis, can be used to locate key binding regions involved in protein interactions.

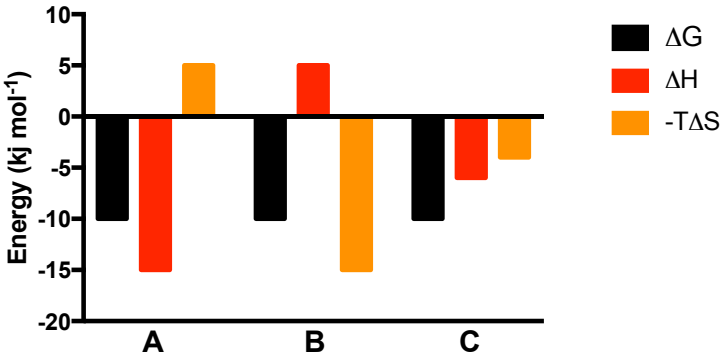


Figure 2.5: *Three different thermodynamic profiles corresponding to protein interactions with the same ΔG.* ***A)*** *Favourable hydrogen and electrostatic bonds result in a negative ΔH which compensates for an unfavourable conformational change and negative ΔS.* ***B)*** *A large increase in entropy, due to hydrophobic interactions, compensates for the increase in ΔH produced by the formation of unfavourable hydrogen and electrostatic bonds.* ***C)*** *Both favourable hydrogen and electrostatic bonds, and hydrophobic interactions result in a decrease in enthalpy and increase in entropy.*

2.3 X-Ray Crystallography

X-Ray crystallography is the most successful biophysical technique for atomic-level structural determination of proteins. The first protein solved was myoglobin in 1958 [180], and the setup of the Protein Data Bank (PDB) allowed for a single repository for all solved crystal structures [181]. The PDB has expanded greatly since its release and now also includes structures solved using other techniques such as NMR [182].

X-ray crysallography involves the crystallisation of a protein or protein-ligand complex followed by bombardment with X-rays, resulting in an X-ray diffraction pattern. This pattern can be used to calculate electron density within the protein crystal which can then be used to calculate protein structure.

2.3.1 Protein Crystallisation

Crystallising a protein produces a repeating unit cell that is packed throughout the entire crystal. The most common method of crystallisation is vapour diffusion [183] where crystals are produced by mixing small volumes of protein and precipitant to form a drop which is surrounded by a reservoir containing a lager volume of precipitant. Within the drop, water will diffuse out overtime, and both the protein and precipitant concentration will increase until the precipitant concentration within the drop is equal to that within the well and an equilibrium between the reservoir and the drop is formed. If crystallisation conditions are favourable the protein will become supersaturated and forced out of solution in the form of protein crystals. Factors that influence the likelihood of crystal formation are protein purity and concentration, temperature, pH, and the specific precipitant or additives used [184].

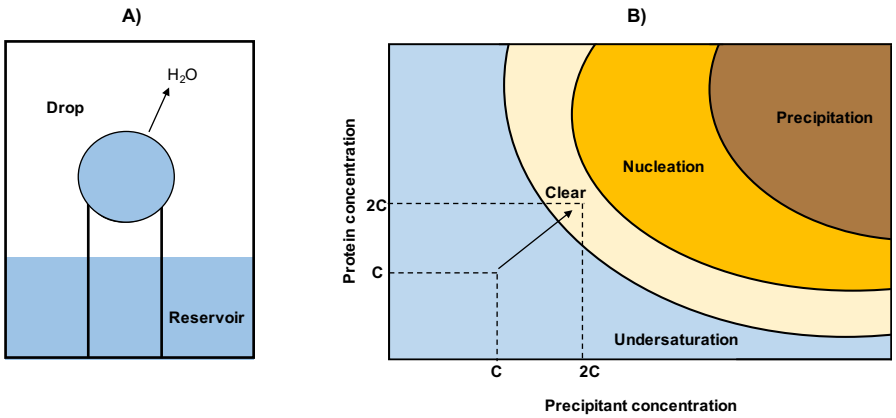


Figure 2.6: ***A)*** *Cartoon representation of the sitting drop vapour diffusion technique where the drop, consisting of an equal mixture of protein and precipitant, sits above the reservoir which is filled with a larger volume of precipitant.* ***B)*** *Graph, where protein concentration is plotted against precipitant concentration, highlighting the various stages of crystal formation by vapour diffusion. Water diffusion out of the drop concentrates both protein and precipitant until the protein begins to precipitate. Crystals will eventually form assuming the crystallisation conditions favour the process. Image adapted from Dessau et al [185].*

2.3.2 Crystal Diffraction

Once proteins crystals have been produced, an X-ray beam is fired at the crystal and X-ray diffraction is detected. The most common X-ray sources used are synchrotron sources; these have a much more intense X-ray beam allowing for shorter exposure times and a better signal to noise ratio [186]. X-ray diffraction is detected using a charged coupled device which transforms X-ray photons to an electrical signal that can then be sent to a computer [187]. The X-ray diffraction profile is dependent on the crystal lattice structure and is used to calculate electron density.

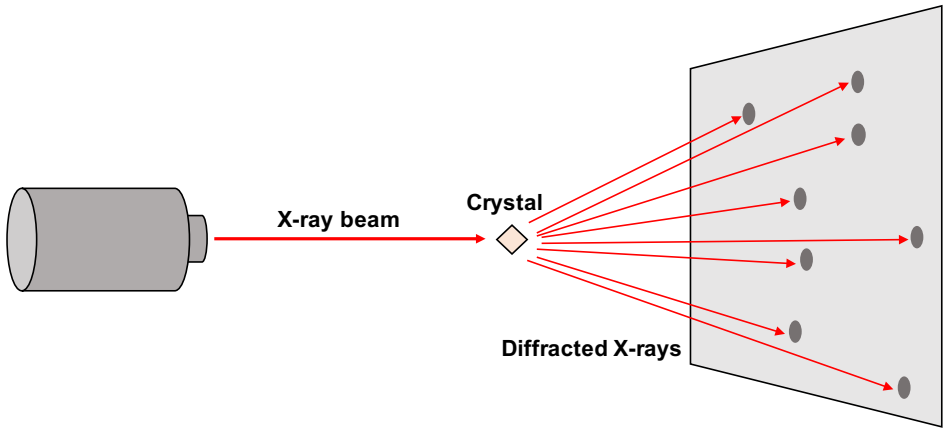


Figure 2.7: *Simplified diagram for crystal X-ray diffraction. A frozen crystal is mounted in a loop and an X-ray beam is shot through the crystal. X-ray beam diffraction is detected, and diffraction patterns can be analysed to determine the atomic structure of the crystal.*

2.3.3 Solving the Phase

The diffraction pattern must be Fourier transformed in order to produce an electron density map of the protein crystal. X-rays have both an amplitude and a phase, and both parameters for each diffraction point must be known in order to produce an electron density map [188]. The amplitude is recorded directly

during the experiment but all information about the phase is lost. The phase can be solved using isomorphous replacement: a technique which was first used for solving protein structures in the 1950's [189] [180]. It involves the incorporation of a heavy atom to the crystal which greatly influences the X-ray diffraction, and by comparing the crystal patterns of the native and the heavy atom bound crystal, the position of the heavy atom can be determined and the phase can be deducted [190]. Both single isomorphous replacement (SIR) and multiple isomorphous replacement (MIR), where one or multiple heavy atoms are incorporated into the crystal, have been used for phase determination.

An alternate method for solving the phase is multi-wavelength anomalous diffraction (MAD) [188]. This technique is used when the presence of the heavy metal causes non-isomorphism between the native and metal bound crystal. This technique involves data collection from a single crystal at a series of different wavelengths which approach the absorption edge of the incorporated heavy atom. At this wavelength anomalous scattering will occur which can then be used to determine the phase. Single-wavelength anomalous diffraction (SAD) also exists and is similar to MAD but only requires the collection of one data set, reducing the radiation damage to the crystal [190].

If a crystal structure of a similar protein has been solved, molecular replacement can be used to solve the phase without the use of isomorphous replacement. This technique involves using co-ordinates from solved structure as a template for fitting similar crystal data [191].

2.3.4 Model Building and Refinement

Once an electron density map has been produced, the protein sequence must be fit to the electron density co-ordinates. Usually the protein backbone is fit first

followed by side chain atoms if the resolution is high enough. The preliminary model is refined against the experimental data which improves phase calculations, and the process of model refinement against experimental data can be cycled until there are no further improvements to the model [192]. The R-factor is a measure of the agreement between experimental data and the model [193], and a good model should have a value of $\leqslant 0.2$ %.

2.4 Mass Spectrometry

Mass spectrometry has become an extremely effective technique in the study of proteins and is now widely used for the investigation of large protein complexes [194], protein dynamics [195], proteomics [196] and protein-metal or protein-drug interactions [197]. Mass spectrometry is able to determine accurate masses of a chemical species by measuring the mass/charge (m/z) ratio of gas-phase ions. It consists of an ion source, which converts the solution based sample to gaseous ions; a mass analyser, that separates different species based on their m/z ratio; and an ion detector [198].

2.4.1 Ionisation

One major hurdle for the study of proteins in their native state using this technique was that the harsh ionising methods would denature the protein structure. Therefore, the development of soft ionising techniques such as electrospray ionisation (ESI) [199] and matrix-assisted laser desorption/ionisation (MALDI) [200] revolutionised the study of proteins by mass spectrometry. ESI was the ionisation technique used for all the mass spectrometry data presented in this thesis. ESI uses electrical energy to convert solution based ions to the gas phase; since this does not require heating or bombardment with high energy particles,

ESI is considered a soft ionisation technique and can be used for the study of native proteins with a high degree of sensitivity. The transfer of ionic species to the gas phase occurs in three steps [201]:

1) The dispersal of a fine spray of charged droplets.

2) Solvent evaporation.

3) Ionic ejection from the charged droplets.

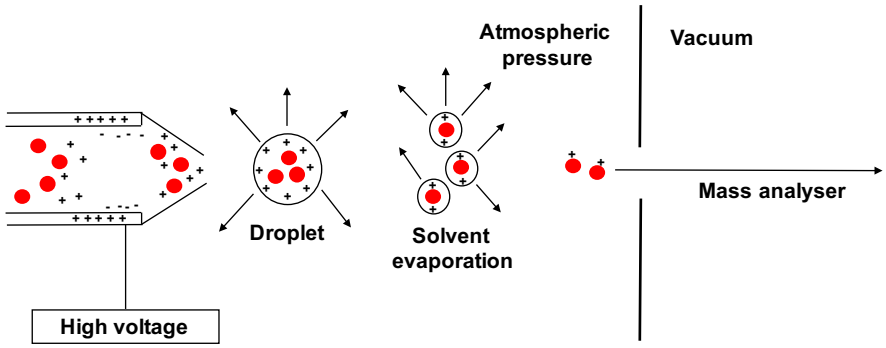


Figure 2.8: *The formation of ions by ESI involves the application of high voltage to the tip of the spray needle, resulting in the dispersion of the sample solution into an aerosol of highly charged droplets. Solvent evaporation from the charged droplet causes a "coolombic explosion" producing smaller droplets, and this process repeats itself until the charge is transferred to the analyte in the gas phase. Image adapted from Mano et al [202].*

2.4.2 Ion Detection

Once in the gas phase, the ions are accelerated through a mass analyser, and a time of flight (TOF) analyser was used to separate the various ions for detection. This method accelerates the ions with an electric field of known strength to a detector [203]. The ion travel speed is dependent on its m/z ratio and can be calculated based on how long it takes to reach the detector at a known distance, and the larger the m/z ratio the longer it takes each ion to travel to the detector. A number of different charge states will be visible for each protein, and the more

disordered a protein the larger the distribution of charge states [204].

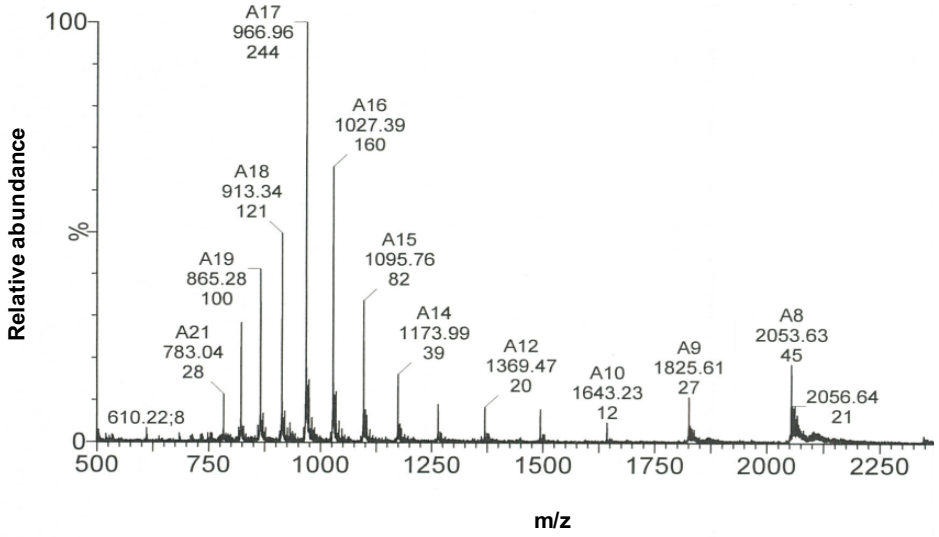


Figure 2.9: *Native ESI-TOF mass spectrum for D5 of HMWK. Each peak corresponds to a different charge state of D5 with the same calculated mass but a different m/z ratio. The wide charge state distribution is a characteristic feature for a disordered protein.*

2.5 Nuclear Magnetic Resonance

Nuclear magnetic resonance (NMR) is an extremely effective technique for the study of protein structure, dynamics and interactions. Experiments are usually carried out in solution giving NMR an advantage over X-ray crystallography as additional details, including protein dynamics and flexibility, can be determined. The major disadvantages of studying proteins by NMR include increased difficulties when studying larger protein structures, long data collection causes issues with unstable proteins and expensive labelling techniques.

2.5.1 Principle of NMR

Atomic nuclei with odd mass numbers exhibit a property known as spin and will rotate around a given axis. The most common nuclei used in NMR is the ^{1}H nuclei which has a spin of 1/2. Due to the spinning charge of this nucleus, it generates a very small magnetic field which, when applied to an external magnetic field, will precess around the external field axis. By applying a radio pulse which is both perpendicular to the applied magnetic field and at the resonant frequency of the precessing nuclei the precession will be flipped 90 °to the magnetic field. The rotating magnetic field induces an electric current which decays over time, and this decaying signal generates a peak for that particular nucleus. The frequency at which a peak appears is dependent on the chemical environment of that particular nuclei, resulting in a range of peaks from nuclei in different chemical environments.

2.5.2 2D Heteronuclear NMR Experiments

Due to the size and chemical complexity of proteins, 1D NMR cannot be used for sophisticated analysis of protein structure. 2D NMR methods have been developed which allow for the study of proteins in far more detail than could be achieved with conventional 1D methods. Heteronuclear experiments utilise more than one spin active nuclei in one experiment, and its development allowed for the study of larger proteins, over 100 amino acids, compared with homonuclear 2D NMR techniques. Additionally to ^{1}H, the most common nuclei used for protein NMR studies are ^{15}N and ^{13}C; these isotopes of nitrogen and carbon are not naturally abundant but, by growing proteins in a minimal media containing a source of ^{15}N or ^{13}C, can be fully incorporated into the protein [205].

The most common heteronuclear experiment for protein NMR is the heteronuclear single quantum correlation (HSQC) experiment. This correlates the nitrogen atom of each amine group with the directly bound hydrogen atom [206]; therefore, there should be a single amine peak for every amino acid within a particular protein, excluding proline, plus additional side chain amine peaks from Arg, Lys, Gln and Asn. A HSQC experiment has ^{1}H frequency plotted against ^{15}N or ^{13}C frequency. This experiment is often the first 2D experiment used for proteins and gives a good indication to whether a full backbone assignment will be possible. By combining the HSQC with other 2D and 3D experiments, it is possible to assign each HSQC peak to the correct amino acid.

Each peak is extremely sensitive to the chemical environment and can be used to map protein interactions to key residues or binding patches. Upon the titration of an unlabelled binding partner, key residues involved in binding will experience a greater change in chemical environment, and the corresponding HSQC peaks will shift in the spectrum [207]. As a general rule, the bigger the chemical shift perturbation the closer to the binding pocket that residue lies.

CHAPTER 3

Materials and Methods

3.1 Biological Materials

The His-tagged D5 construct DNA (section 9.1.1.1) was supplied in pET28a by Keith McCrae (Department of Cellular and Molecular Medicine, Cleveland Clinic). The gC1q-R construct DNA (section 9.1.1.2) was supplied in a pT7 A.A-32 plasmid, originally described by Krainer et al [208], by Bubacar Kaira (Emsley Group, Centre of Biomedical Sciences, Nottingham University). pGEX vectors used to clone D5 truncations were supplied by Jed Long (Searle Group, Centre of Biomedical Sciences, Nottingham University). Other biological reagents used and their suppliers are listed in Table 3.1.

Biological Reagent	Supplier
BamHI	New England Biolabs
BL21 (DE3)	Agilent
DNAse I	New England Biolabs
NdeI	New England Biolabs
Thrombin	Enzyme Research Laboratories
XhoI	New England Biolabs
XL1-Blue	Agilent

Table 3.1: *List of used biological reagents and their suppliers.*

3.2 Chemical Materials

All DNA primers were purchased from Sigma-Aldrich. Other chemicals used and their suppliers are listed in Table 3.2. A breakdown of all the buffers utilised is documented in Table 3.3.

Chemical	Supplier
^{15}N Ammonium Chloride	Sigma-Aldrich
30 % Protogel	National Diagnostics
Agarose	Fisher
Ammonium Acetate	Fisher
Ammonium Molybdate Tetrahydrate	Sigma-Aldrich
Ampicillin Sodium Salt	Fisher
APS	Sigma-Aldrich
Biotin	Sigma-Aldrich
Bromophenol Blue	Fisher
Calcium Chloride	Sigma-Aldrich
Copper(II) Chloride Dihydrate	Sigma-Aldrich
D_2O	Sigma-Aldrich
Disodium Hydrogen Orthophosphate	Fisher
Dipotassium Hydrogen Orthophosphate	Fisher
DTT	Sigma-Aldrich
EDTA	Sigma-Aldrich
Ethidium Bromide	Fisher
Glucose	Fisher
Glutathione Sepharose 4B	GE Healthcare
Glycerol	Fisher
Glycine	Fisher
Imidazole	Sigma-Aldrich
IPTG	VWR
Iron(III) Chloride Hexahydrate	Sigma-Aldrich
Kanamycin Sulphate	Fisher
LB Agar	Sigma-Aldrich
LB Broth	Sigma-Aldrich
M 2-Mercaptoethanol	Sigma-Aldrich
Magnesium Sulphate	Sigma-Aldrich
Manganese(II) Chloride Tetrahydrate	Sigma Aldrich
Nickel(II) Sulphate	Fisher
Potassium Dihydrogen Orthophosphate	Fisher
Sodium Azide	Sigma-Aldrich
Sodium Chloride	Fisher
Sodium Dodecyl Sulphate	Fisher
Sodium Tetraborate Decahydrate	Sigma-Aldrich
TEMED	Sigma-Aldrich
Thiamine Hydrochloride	Sigma-Aldrich
Tris	Fisher
MEM Vitamin Solution	Sigma-Aldrich
Zinc Chloride	Sigma-Aldrich

Table 3.2: *List of used chemicals and their suppliers.*

Solutions	Chemical Concentrations
15 % Acrylamide Buffer	375 mM Tris HCl, 15 % protogel (w/v), 0.1 % SDS (w/v), pH 8.8
2 X SDS Loading Dye	100 mM Tris HCl, 4 % SDS (w/v), 20 % Glycerol (v/v), 0.2 % Bromophenol Blue (w/v), 200 mM DTT, pH 6.8
AGF Buffer	50 mM Tris HCl, 100 mM Sodium Chloride, pH 7.5
AGF EDTA Buffer	50 mM Tris HCl, 100 mM Sodium Chloride, 5 mM EDTA, pH 7.5
AGF Zinc Buffer	50 mM Tris HCl, 100 mM Sodium Chloride, 50 µM $ZnCl_2$, pH 7.5
CD Buffer	50 mM Tris (SO4), pH 7.5
Cell Storage Buffer	LB Broth 25 g/L, 8 % Glycerol (v/v)
Cleavage Buffer	50 mM Tris HCl, 150 mM Sodium Chloride, 25 mM Calcium Chloride, pH 8.4
Denaturing AGF Buffer	50 mM Tris HCl, 100 mM Sodium Chloride, 0.1 M 2-mercaptoethanol and 0.1 % SDS (w/v), pH 7.5
Desalt Buffer	100 mM Ammonium Acetate
Gel Filtration Buffer	30 mM Potassium Phosphate, 100 mM Sodium Chloride, pH 7.0
His Trap Binding Buffer	20 mM Potassium Phosphate, 100 mM Sodium Chloride, pH 7.5
His Trap Elution Buffer	20 mM Potassium Phosphate, 100 mM Sodium Chloride, 1 M Imidazole, pH 7.5
Ion Exchange High Saltzdfz Buffer	10 mM Potassium Phosphate, 2 M Sodium Chloride, pH 7.0
Ion Exchange High Salt Buffer B	10 mM Potassium Phosphate, 2 M Sodium Chloride, pH 6.0
Ion Exchange Low Salt Buffer	10 mM Potassium Phosphate, pH 7.0
Ion Exchange Low Salt Buffer B	10 mM Potassium Phosphate, pH 6.0
ITC Buffer	20 mM Potassium Phosphate, 100 mM Sodium Chloride, pH 7.4
ITC EDTA Buffer	20 mM Potassium Phosphate, 100 mM Sodium Chloride, 5 mM EDTA, pH 7.4
ITC Zinc Buffer	20 mM Potassium Phosphate, 100 mM Sodium Chloride, 50 µM Zinc Chloride, pH 7.4
Lysis Buffer	10 mM Tris HCl, 150 mM Sodium Chloride, pH 7.5
NMR Buffer	10 mM Potassium Phosphate, 25 mM Sodium Chloride, 0.02 % Sodium Azide (w/v), 10 % D_2O (v/v), pH 7.0
Protein Storage Buffer	20 mM Potassium Phosphate, 100 mM Sodium Chloride, 10 % Glycerol (v/v), pH 7.4
SDS Running Buffer	25 mM Tris HCl, 250 mM Glycine, 0.1 % SDS (w/v)
Stacking Acrylamide Buffer	125 mM Tris HCl, 4 % protogel (v/v), 0.1 % SDS (w/v), pH 6.8
TAE	40 mM Tris HCl, 50 mM EDTA, 50 mM acetic acid, pH 8.0

Table 3.3: *List of buffers used and their individual component concentrations. All buffers were made using miliQ water.*

3.3 Molecular Biology Techniques

3.3.1 Sterilisation and Sterile Technique.

All buffers, glassware and plastic consumables used for the molecular biology techniques described were autoclaved for 20 mins at 121 °C. All work involving live cells was performed next to a bunsen flame to decrease any chance of contamination.

3.3.2 Overnight Cultures

Overnight cultures were required for both the purification of plasmid DNA and protein expression. Cells were scraped from glycerol stocks or single colonies from agar plates into 10 mL LB Broth with the appropriate antibiotic resistance. The cell resistance was dependent on the plasmid the cells had been transformed with, and a summary of plasmid resistance is shown in Table 3.4. Cultures were left to grow for approximately 12-16 hours at 37 °C with 180 rpm agitation.

Antibiotic resistance	Antibiotic Concentration (μg/mL)	Resistant Plasmids
Ampicillin	100	gC1q-R T7, pGEX 4T-1
Kanamycin	50	pET28-a

Table 3.4: *Antibiotic resistance of the different bacterial expression plasmids and their working concentrations.*

3.3.3 Agar Plates

Bacterial suspensions were plated onto LB agar plates in order to get individual bacterial colonies. 40 g/L of LB agar was suspended in water and autoclaved. On the day of use, the solid agar was heated in a microwave until fully melted and was left to cool before adding the appropriate antibiotic (Table 3.4). The agar was poured into petri dishes, using sterile technique, and left to set at RT next to flame. Plates were stored at 4 °C.

3.3.4 Glycerol Stocks

Cells were stored as 8 % glycerol (v/v) stocks for long term storage. 1 mL of overnight culture from a freshly transformed agar plate was centrifuged at 1,000 rcf, and the pellet was gently resuspended in Cell Storage Buffer. Cell suspensions were snap frozen and stored at - 80 °C.

3.3.5 Plasmid DNA Purification

Plasmids were purified using the Qiaprep plasmid miniprep kit from Qiagen. This technique involves alkaline lysis of E. coli cells followed by the binding of DNA to a silica membrane under high salt conditions and elution of plasmid DNA under low salt conditions [209]. 5 mL of overnight culture from E.coli XL1-Blue glycerol stocks were grown overnight and centrifuged at 3,000 g. The pellets were treated according to the Qiaprep protocol, and plasmid concentrations were calculated by nanodrop as described in section 3.8.1.

3.3.6 DNA Transformation.

Plasmids were transformed into appropriate *E. coli* strains using the calcium dependent heat shock method first identified by Mandel et al. [210]. Treatment of cells with a source of Ca^{2+} followed by a heat-cold shock allows for the uptake of exogenous DNA into the cell [211] [212]. Overnight cultures of XL1-Blue or BL21 (DE3) cells were inoculated, 1 in 50, into fresh LB Broth and grown for another 30 mins at 37 °C before being centrifuged at 3,000 g. The pellets were re-suspended in 4 mL of ice cold calcium chloride (50 mM) and left on ice for 1 hour. Cells were once again centrifuged at 3,000 g, and the pellet was re-suspended in 800 µL of ice cold calcium chloride (50 mM) and stored on ice. 100 µL of competent cells were mixed with 1 µL of purified plasmid DNA and then left on ice for 30 mins. The cells were then heat shocked at 42 °C for 45 seconds and then placed on ice for a further 5 mins. 400 µL of fresh LB was added, and transformed cells were grown at 37 °C with 180 rpm agitation for 30 mins before being plated on agar plates as described in section 3.3.3.

3.3.7 Agarose Gels

Agarose gel electrophoresis was used in order to observe and analyse DNA fragments from cloning experiments. This is an effective method of separating DNA based on size and works through the application of an electric current which pulls the negatively charged DNA through the gel [213]. The distance traveled is dependent on size, and smaller DNA fragments travel quicker through the gel.

1 % w/v agarose gels were made by dissolving 0.5 g of agarose in 50 mL of TAE Buffer and heated until fully dissolved. The liquid agarose was left to cool at RT before 1 µL of ethidium bromide was added; the mixture was then poured

into the gel caster (BioRad Sub-cell GT UV-Transparent Mini-Gel Tray) fitted with the appropriate well comb. Once set, the gel was placed in the gel tank (Bio-Rad Mini-Sub Cell GT) and submerged in TAE Buffer.

1 µL of 6X DNA loading dye, ThermoFisher Scientific, was added to the 5 µL of DNA prior to loading the sample into the gel. Gels were run at 80 V for 30 mins using a BioRad PowerPac Basic, and DNA bands were visualised under a UV light source. Quick-Load Purple 2-Log DNA Ladder (NEB) was also loaded onto the gel and was used to determine the relative sizes of DNA.

3.3.8 Cloning of Protein Mutants Using Q5 Kit.

The Q5 kit (NEB) was used to produce the majority of mutants used throughout this thesis due to the ease and high success rate of this technique. All protocols were taken from the NEB Q5 Site-Directed Mutagenesis Kit webpage. In the presence of appropriate DNA primers, the Q5 polymerase is able to amplify the complete plasmid template DNA followed by circularisation of the recombinant plasmid containing the mutations introduced by the DNA primers.

3.3.8.1 Q5 Polymerase PCR Site Directed Mutagenesis

All primers were designed using the NEBaseChanger online tool. This protocol uses Q5 polymerase in combination with non-overlapping forward and reverse primers for exponential amplification. This polymerase has high fidelity, processivity and speed of replication [214] making it a quick and reliable enzyme for the production of mutant proteins. The primers used in conjunction with Q5 polymerase to produce protein mutants are shown in Table 3.5.

Protein Construct	Forward and Reverse Primers	Ta (°C)
HMWK 401-443	**F** - 5′ tagCATAATCTTGGCCATGGC 3′ **R** - 5′ TTTTCTTTGTTTTTCATGCC 3′	60
gC1q-R ΔEDEAE	**F** - AGTGACATCTTCTCTATCAG 3′ **R** - 3′ TTGTCCAACCTCATCCTC 5′	58
gC1q-R Δβ6β7	**F** - 5′ ACACTCAACACAGATTCC 3′ **R** - 3′ GGACTGAAAGCTAACTTC 5′	58
gC1q-R ΔGlu step 1	**F** - 5′ tgctCCCTCGCAAGGGCAGAAG 3′ **R** - 3′ gcagcACCATCAAATGTTGGTGGGATG 5′	66
gC1q-R ΔGlu step 2	**F** - 5′ GCAGAAGGTTgctgctCAGGAGCCTGAACTGAC 3′ **R** - 3′ CCTTGCGAGGGTTCCTCC 5′	64

Table 3.5: *PCR primers used for the site directed mutagenesis of D5 and gC1q-R using Q5 polymerase. The D5-1 pGEX 4T1 plasmid (described in section 3.3.10) was used as template DNA for the production of HMWK 401-443, whereas the gC1-R WT plasmid was used as a template for the production of gC1q-R mutants. The Ta values for each set of primers, as calculated by the NEBaseChanger tool, are also shown. gC1q-R ΔGlu was produced in two steps: the production of Glu156Ala and Glu157Ala, then the additional mutations of Glu146Ala, 147Ala and 148Ala.*

Reagent	Final Concentration
Q5 Hot Start High-Fidelity 2X Master Mix	1X
Forward Primer	0.5 µM
Reverse Primer	0.5 µM
Template DNA	1 ng/µL

Table 3.6: *Reagents for the production of protein mutant DNA using Q5 polymerase in a 25 µL reaction.*

Step	Cycles	Temperature (°C)	Time (secs)
1: Initial denaturation	1	98	30
2a: Denaturing		98	10
2b: Annealing	25	Ta	30
2c: Extension		72	180
3: Final extension	1	72	120
Hold	1	8	-

Table 3.7: *Thermocycler settings for Q5 polymerase site directed mutagenesis PCR reactions.*

3.3.8.2 KLD Treatment

The PCR products were treated with KLD mix from NEB. KLD consists of DpnI, kinase and ligase. DpnI is an endonuclease that selectively digests methylated DNA [215] and ensures specific removal of the original template DNA, and the kinase and ligase circularises the plasmid DNA. KLD products were stored a 4 °C and transformed into XL1-Blue cells using the protocol described in section 3.3.6.

Reagent	Volume (µL)	Final Concentration
PCR product	1	-
2 X KLD reaction buffer	5	1 X
10X KLD Enzyme Mix	1	1 X

Table 3.8: *Reagents used for the KLD treatment of PCR products in a 10* µL *reaction.*

3.3.9 Cloning of Protein Truncations Using High Fidelity Phusion Polymerase

PCR using high fidelity (HF) phusion polymerase was used for the production of D5-1 (section 9.1.1.3), D5-2 (section 9.1.1.4) and gC1q-R ΔN (section 9.1.1.5) protein constructs. The Q5 protocol was not used for these mutations, as both D5 fragments would be sub cloned into new plasmids and only small PCR fragments, corresponding to the D5-1 and D5-2 genes, were required for the cloning. Q5 continuously failed to successfully generate gC1q-R ΔN so this alternate method was also used for this construct.

3.3.9.1 HF Phusion Polymerase PCR Site Directed Mutagenesis

Phusion polymerase is a high fidelity polymerase with error rates over 10 fold less than those found with *Taq* polymerase [216]. Primers for D5 truncations were designed to produce two D5 fragments, consisting of residues T401-H473 (D5-1) and G474-S531 (D5-2), using the full length D5 plasmid as a template. 5′ *BamHI* and 3′ *XhoI* restriction digest sites were incorporated using these primers, and these fragments were inserted into a pGEX 4T-1 plasmid digested with the same restriction enzymes.

Primers were also designed for the production of gC1q-R ΔN, which consisted of residues Ile96-Gln282. This mutant lacked the N-terminal helix involved in trimer formation. The WT gC1q-R plasmid was used as a template and 5′ *NdeI* and 3′ *BamHI* restriction sites were incorporated. These restriction sites were also present at either end of the WT gC1q-R gene; therefore, the WT gene could be removed and replaced with the gC1q-R ΔN PCR product.

Protein Construct	Forward and Reverse Primers	Restriction Site	Ta (°C)
D5-1	**F** - 5′ GGCggatccACTGTAAGTCC 3′ **R** - 5′ CCGCctcgagCTAATGACCAAGACCATGC 5′	*BamHI* *XhoI*	52
D5-2	**F** - 5′ GGCggattcGGACATAAGTTCAAACTTG 3′ **R** - 5′ CCGCctcgagCTAACTGGCTTC 3′	*BamHI* *XhoI*	52
gC1q-R ΔN	**F** - 5′ GGAATTCcatatgATTCAGAAGCATAAAACC 3′ **R** - 5′ CGCggatccCTACTGGCTCTTGACAAAACT 3′	*NdeI* *BamHI*	48

Table 3.9: *PCR primers used for the production of D5 and gC1q-R truncations using HF phusion polymerase. The pET28-a D5 plasmid was used as a template for the production of D5-1 and D5-2, and the gC1q-R WT plasmid was used as a template for gC1q-R ΔN. The Ta values for each set of primers, as calculated by ThermoFisher Scientific online Ta calculator, are also shown. The incorporated restriction sites are represented in lowercase.*

Reagent	Final Concentration
HF Phusion buffer	1 X
Forward Primer	0.5 µM
Reverse Primer	0.5 µM
Template DNA	10 ng/µL
dNTPs	200 µM
Phusion DNA Polymerase	0.4 U

Table 3.10: *Reagents for the production of protein truncation DNA using HF phusion polymerase in a 20* µ*L reaction.*

Step	Cycles	Temperature (°C)	Time (secs)
1: Initial Denaturation	1	98	30
2a: Denaturing		98	10
2b: Annealing	25	Ta	30
2c: Extension		72	30
3: Final Extension	1	72	600
Hold	1	8	-

Table 3.11: *Thermocycler settings for the HF phusion polymerase site directed mutagenesis PCR reactions.*

3.3.9.2 Gel Extraction

DNA gel extraction was performed on the PCR products in order to remove the PCR primers and phusion polymerase after the PCR reaction and to ensure no secondary products were used in the restriction digest. The PCR mixture was first run on an agarose gel following the protocol described in section 3.3.7. The DNA band corresponding to the PCR product was excised and purified using a Sigma Aldrich GenElute Gel Extraction Kit following the instructions provided by Sigma Aldrich.

3.3.9.3 Double Restriction Digest

Double digests were performed on both the PCR DNA insert and plasmid in order to produce complementary "sticky ends" which would be used for ligating the insert and plasmid [217]. Each reaction was incubated at 37 °C for 1 hour, and The cut products were gel extracted as described in section 3.3.9.2 to ensure the removal of restriction enzymes and excised plasmid DNA.

Insert	Vector	Restriction Enzymes
D5-1	pGEX 4T-1	*BamHI, XhoI*
D5-2	pGEX 4T-1	*BamHI, XhoI*
gC1q-R ΔN	gC1q-R T7	*NdeI, BamHI*

Table 3.12: *Plasmid vectors and restriction digest enzymes used for the cloning of phusion PCR products.*

Reagent	Final Concentration
CutSmart Buffer	1 X
Restriction Enzymes	20 U each
DNA Insert	5 ng/µL
Vector	50 ng/µL

Table 3.13: *Reagents used for double restriction digest reactions in a 20 µL reaction.*

3.3.9.4 Ligation Using T4 Ligase

Prior to ligation, Antarctic Phosphatase was incubated with the cut vector for 15 mins at 37 °C. This enzyme removes the 5′ phosphate group from the vector, preventing it from ligating back together during the ligase reaction [218]. The enzyme was heat inactivated by incubating for 20 mins at 65 °C.

T4 DNA ligase was used to ligate the DNA insert and vector, producing a circular recombinant plasmid. The ligase reaction was left at R.T for 20 mins before being placed on ice and transformed into XL1-Blue cells as described in section 3.3.6.

Reagents	**Final Concentrations**
Antarctic Phosphatase Reaction Buffer	1 X
Cut Vector	2.5 ng/µL
Antarctic Phosphatase	5 U

Table 3.14: *Reagents used for the treatment of cut vector with Antarctic Phosphatase in a 20* µL *reaction.*

Reagents	**Final Concentrations**
T4 DNA ligase buffer	1 X
Cut Vector	2.5 ng/µL
Cut Insert	7.5 ng/µL
T4 DNA Ligase	20 U

Table 3.15: *Reagents used for the ligation of digested insert and vector in a 20* µL *reaction.*

3.3.10 DNA Sequencing

DNA sequencing was carried out on recombinant plasmids to confirm the DNA sequence following mutagenesis experiments. Sequencing was performed by the DNA Sequencing Facility at the University of Nottingham Medical School, QMC. 5 µL of freshly mini prepped plasmid DNA was used per reaction along with pGEX-R or T7-F sequencing primers supplied by the DNA Sequencing Facility.

3.4 Protein Expression

The T7 bacterial expression system was used for the over expression of recombinant proteins. This system is the most widely used expression system for the production of proteins in bacteria due to its ability to produce large amounts of the target protein whilst maintaining a significant amount of control over expression rates [219]. The T7 expression system involves the transcription of a gene of interest, which is inserted just after a T7 promoter and *lac* operator, through IPTG dependent inhibition of the lac repressor [220].

3.4.1 Non Labelled Protein

All non isotopically labeled growths were performed in 1 L of LB Broth which had been autoclaved in 2 L baffled flasks. Overnight cultures of the correct BL21 (DE3) strain were inoculated, 1 in 50, into the main growth flask with the appropriate antibiotic (Table 3.4). The cells were left to grow at 37 °C with 180 rpm agitation to an OD^{595nm} of 0.6 before cells were induced with 1 mM IPTG and left for 5 hours at 30 °C. Cells were harvested by centrifuging at 3,000 g,

and cell pellets were collected and stored at - 80 °C.

3.4.2 Labelled Protein

^{15}N labelled proteins were produced in order to be studied by 2D NMR. Protein labelling can be achieved by growing cells in a minimal media in the presence of a source of ^{15}N [221]. 1 L minimal media was prepared, using the reagents shown in Table 3.16, and autoclaved. Prior to inoculation, 4 g of glucose and 1 g of ^{15}N ammonium chloride was dissolved in 20 mL autoclaved water and added to the media using a 0.2 µm syringe filter. In addition, 1 mL of vitamins, 10 mg of thiamine and biotin, and the appropriate antibiotic were added. The final addition was 1 mL trace elements, as prepared according to Table 3.17.

Reagent	Concentration (mg/mL)
Disodium Hydrogen Orthophosphate	6
Potassium Dihydrogen Orthophosphate	3
Sodium Chloride	0.5
Magnesium Sulphate	0.3
Calcium Chloride Dihydrate	0.015

Table 3.16: *Individual components used to make minimal media.*

Reagent	Concentration (ng/mL)
Iron(III) Chloride Hexahydrate	400
Zinc(II) Chloride	80
Ammonium Molybdate Tetrahydrate	20
Copper(II) Chloride dihydrate	20
Manganese(II) Chloride Tetrahydrate	20
Sodium Tetraborate Decahydrate	20

Table 3.17: *Individual components used to make trace element stock solution.*

Overnight cultures, enough for a 1 in 25 mL inoculation, were centrifuged at

3,000 g, and pellets were resuspended in 25 mL of fully prepared minimal media. These mini cultures were grown at 37 °C for 30 mins before being inoculated into the main 1 L flask. Cells were left at 37 °C with 180 rpm agitation for 2 hours before the temperature was dropped to 30 °C. Cells were induced at an OD^{595nm} of 0.7 with 1 mM IPTG and left for roughly 15 hours. Cells were cultured by centrifuging at 3,000 g, and pellets were collected and stored at - 80 °C.

3.5 Preparation of Protein Lysates

Cell pellets were defrosted at RT and resuspended in 10 mL Lysis Buffer per L of pellet. Prior to sonicating, 1 mg of DNase I and 100 microL of 10 x protease inhibitor cocktail, from Sigma-Aldrich, was added. The cells were sonicated at 10 microns for 10 X 30 secs, with 30 secs rests. Cell lysates were then centrifuged at 35,000 g, filtered using a 0.2 µm, 25 mm syringe filter and kept on ice.

3.6 Protein Purification by FPLC

All purifications were performed using an AKTA Prime and AKTA Start FPLC system from GE Healthcare. FPLC columns were also purchased from GE Healthcare. All buffers were filtered and degassed using 47 mm, 0.2 µm pore sized Whatman filters before usage. The UV absorption was monitored at 280 nm, and protein purity was analysed by SDS-PAGE as described in section 3.7.1. Protein concentrations were calculated by nanodrop as described in section 3.8.1 using the individual extinction coefficients of each protein. Protein extinction coefficients, which are dependent on the number of Trp, Tyr and Cys residues within the protein sequence [222], were calculated using equation

3.6.1

$$\epsilon^{280} = nW \times 5,500 + nY \times 1,490 + nC \times 125 = A^{0.1\,\%} \times MW \qquad (3.6.1)$$

Where ϵ^{280} is the molar extinction coefficient at 280 nm, nW is the number of Trp residues, nY is the number of Tyr residues and nC is the number of reduced Cys residues. A $^{0.1\,\%}$ is the absorbance of a 1 g/L sample of the protein and MW is the molecular weight of the protein in Da. The units of ϵ^{280} are M^{-1} cm^{-1}

All the calculated extinction coefficients for the proteins used with an absorption at 280 nm are shown in Table 3.18.

Protein	**ϵ^{280} (M^{-1} cm^{-1})**	**Absorbance 0.1 %**
D5	11000	0.671
gC1q-R	22460	0.944
D5-1	5500	0.654
D5-2	5500	0.83
gC1q-R ΔN	22460	1.053
gC1q-R Δ β3-β4	22460	0.955
gC1q-R Δ β5-β6	22460	0.967
gC1q-R Δ β7-β8	15470	0.688

Table 3.18: *Extinction coefficients calculated for all the protein constructs that had an absorption at 280 nm. These values were used to calculate protein concentration.*

The HMWK peptides did not contain any Trp, Tyr or Cys residues so concentrations were calculated by measuring peptide backbone absorbance at 205 nm [223]. The extinction coefficient of A 0.1 % = 31 was used to calculate peptide concentration from absorbance.

3.6.1 Purification of D5

3.6.1.1 Cationic Exchange Chromatography

Cationic exchange chromatography was the first purification step used for His-tagged D5. This method involves the trapping of positively charged proteins using a column pre-packed with sulfopropyl (SP) sepharose, followed by protein elution by introducing a salt gradient [224]. The D5 lysate was diluted 5 fold with Ion Exchange Low Salt Buffer and loaded onto a 5 mL HiTrap SP cationic exchange column equilibrated in 5 CV of Ion Exchange Low Salt Buffer. His-tagged D5 was eluted by introducing 40 % Ion Exchange High Salt Buffer. The operating parameters used are reported in Table 3.19.

Pressure Limit (MPa)	Flow Rate (mL / min)	Step Elution (% B)
0.5	4	40

Table 3.19: *Operating parameters used for the cationic exchange purification step of His-tagged D5 using a 5 mL HiTrap SP column.*

3.6.1.2 Immobilised Metal-Affinity Chromatography

His-tagged D5 was further purified using Immobilised Metal-Affinity Chromatography (IMAC). This method involves the use of a column pre packed with sepaharose modified with a metal chelating group. The modified sepharose can be charged with metal ions which will effectively bind poly His tags present within a protein [225]. Bound proteins are then eluted by washing free imidazole through the column [226]. A 5 mL HisTrap HP column was used for the IMAC purification step of His-tagged D5. The column was washed with 1 CV of EDTA (500 mM) and charged by loading 1 CV of $NiSO_4$ (100 mM). The column was then equilibrated in 5 CV of HisTrap Binding Buffer.

Pressure Limit (MPa)	Flow Rate (mL / min)	Elution Gradient (% B)	Gradient Length (mL)	Fraction Size (mL)
0.5	3	10 - 55	170	10

Table 3.20: *Operating parameters used for the IMAC purification step of His-tagged D5 using a 5 mL HisTrap HP column.*

3.6.1.3 Thrombin Cleavage

His-tagged D5 was concentrated to roughly 5 mL using a Vivaspin 3 kDa MWCO spin concentrator from GE Healthcare and topped up with 10 mL Thrombin Cleavage Buffer. The sample was concentrated back to 5 mL, and this process was repeated three times to ensure the sufficient removal of imidazole. 2.5 U of human alpha thrombin was added, and the sample was left for roughly 16 hour with gentle rotation in order to fully cleave the N-terminal His tag. The protein solution was diluted 5 fold with Ion Exchange Low Salt Buffer and loaded onto a 5 mL HiTrap SP. Cleaved D5 was eluted using a 20-40 % Ion Exchange High Salt Buffer gradient. Cleaved D5 fractions were pooled and spin concentrated to 3 mL using a Vivaspin 3 kDa MWCO spin concentrator from GE Healthcare. The protein sample was dialysed against 1 L Protein Storage Buffer for 12 hours at 4 °C using a Mini Dialysis Kit, 1 kDa cut-off from GE Healthcare. The dialysed protein was snap frozen and stored at - 80 °C.

3.6.2 Purification of gC1q-R

3.6.2.1 Anionic Exchange Chromatograph

Anionic exchange chromatography was used as the first purification step for gC1q-R and is similar in principle to cationic exchange (described in section

3.6.1.1) but involves the binding of negatively charged proteins due to the presence of quaternary ammonium (Q) sepharose [227]. The gC1q-R cell lysate was diluted 5 fold and loaded onto a 5 mL HiTrap Q column equilibrated in 5 CV of Ion Exchange Low Salt Buffer. gC1q-R was eluted by introducing a 0-40 % Ion Exchange Hight Salt Buffer gradient.

Pressure Limit (MPa)	Flow Rate (mL / min)	Elution Gradient (% B)	Gradient Length (mL)	Fraction Size (mL)
0.5	3	0 - 40	200	10

Table 3.21: *Operating parameters used for the anionic exchange purification step of gC1q-R using a 5 mL HiTrap Q column.*

3.6.2.2 Size Exclusion Chromatography

Size exclusion chromatography, also commonly referred to as gel filtration, is a method of separating proteins by size [228]. The gel filtration media, consisting of dextran covalently bound to a highly cross linked agarose, behaves as a molecular sieve that large proteins will travel a more direct route through than smaller proteins. gC1q-R fractions from the anionic exchange step were pooled and spin concentrated to a final volume of 10 mL using a Vivaspin 10 kDa MWCO spin concentrator from GE Healthcare. This was loaded onto a superdex 200 gel filtration column equilibrated in 1.5 CV of Gel Filtration Buffer. Fractions corresponding to pure gC1q-R were pooled, spin concentrated to a final volume of 3 mL and dialysed against 1 L of Protein Storage Buffer for 12 hours at 4 °C using a 1 kDa cut-off Mini Dialysis Kit from GE Healthcare. The dialysed protein was snap frozen and stored at - 80 °C.

Pressure Limit (MPa)	Flow Rate (mL / min)	Fraction Size (mL)
0.5	2	10

Table 3.22: *Operating parameters used for the size exclusion purification step of gC1q-R using a superdex 200 column.*

3.6.3 Purification of D5-1, D5-2 and HMWK 401-443

3.6.3.1 Glutathione Affinity Chromatography and Thrombin Cleavage

The presence of an N-terminal GST tag on the three D5 truncations allowed for their purification using glutathione affinity chromatography. Glutathione sepharose is used to bind the GST-tagged protein and consists glutathione-linked sepharose beads that bind directly to the GST tag [229]. This purification method allows for binding, washing and cleavage using one simple purification method. A 20 mL drip column was prepared by adding 1 mL of Glutathione Sepharose 4B beads, supplied by GE healthcare, followed by washing with 3 CV of Lysis Buffer. The cell lysates were added to the column and left tumbling for 5 hours at RT. The cell lysate were then eluted from the column, and the beads were washed with 2 CV of Lysis Buffer and 2 CV of Cleavage Buffer. 1 mL of Cleavage Buffer plus 5 U of thrombin was added to the column and left tumbling for 15 hours at 4 °C. The cleaved protein was eluted from the column, and beads were further washed with 1 CV of lysis buffer to ensure all cleaved protein had been removed form the beads.

3.6.3.2 Cationic Exchange Chromatography

Cleaved D5-1 and D5-2 were diluted 5 fold in Ion Exchange Low Salt Buffer and loaded on to a 5 mL HiTrap SP column equilibrated in 5 CV of Ion Exchange Low Salt Buffer. D5-1 was eluted using a 0-30 % Ion Exchange High Salt Buffer gradient, whereas D5-2 was eluted using a 20-40 % Ion Exchange High Salt Buffer gradient.

HMWK 401-443 was diluted 5 fold in Ion Exchange Low Salt Buffer B and loaded onto a 5 mL HiTrap SP column equilibrated in Ion Exchange Low Salt Buffer B. The protein was eluted using a 0-20 % Ion Exchange High Salt Buffer B gradient.

Protein	Pressure Limit (MPa)	Flow Rate (mL / min)	Elution Gradient (% B)	Gradient Length (mL)	Fraction Size (mL)
D5-1	0.5	3	0 - 30	200	10
D5-2	0.5	3	20 - 40	100	10
401 - 443	0.5	3	0 - 20	100	10

Table 3.23: *Operating parameters used for the cationic exchange purification steps of D5-1, D5-2 and HMWK 401-463 using a 5 mL HiTrap SP column.*

3.7 Protein Analytical Techniques

3.7.1 SDS-PAGE

SDS-PAGE was used to assess protein identity and purity throughout protein purifications and analytical gel filtration experiments. This technique involves the coating of proteins with negatively charged sodium dodecyl sulphate (SDS) and their subsequent size separation through the application of an electrical current [230].

3.7.1.1 Gel Preparation

All gels consisted of a stacking gel set on top of a resolving gel with the dimensions 10 x 10 cm with 1 mm width. The individual components of both the stacking and resolving gels are shown in Table 3.3. For the resolving gel, 25 µL of 10 % APS (w/v) and 5 µL of TEMED was added to 3 mL of 15 % Acrylamide Buffer and poured into the gel caster. 200 µL of iso-propanol was added to the gel bed to ensure the gel set evenly. The stacking gel was made by adding 25 µL of 10 % APS (w/v) and 5 µL of TEMED to 1 mL of Stacking Acrylamide Buffer which was then poured on top of the resolving gel. 15 well combs were added to the stacking solution prior to setting. Once set, the gel was assembled in the gel electrophoresis tank and filled with SDS Running Buffer. The well combs were carefully removed prior to sample loading.

3.7.1.2 Sample Preparation and Gel Electrophoresis

5 µL of sample was combined with 5 µL 2 X SDS Loading Dye and heated to 90 °C for 5 mins before being loaded into the well. The first lane of each gel was loaded with 5 µL of Novex Sharp Pre-Stained Protein Standard, 3.5 - 260 kDa, from ThermoFisher. Gels were run for 70 mins at 180 V and 40 mA per gel using a BioRad PowerPack Basic.

3.7.1.3 Gel Staining

Gels were carefully transferred to a plastic container and rinsed in water before covering in GelCode Blue Safe Protein Stain from ThermoFisher and microwaving on a low setting for 10 mins. Gels were left to destain in water until protein bands became clearly visible. Digital pictures for each gel were recorded on a

SynGene G:Box.

3.7.2 Analytical gel filtration

Analytical gel filtration was used to analyse changes in protein conformation, stability and detecting protein-protein interactions between gC1q-R and D5 ligands. Gel filtration elution profiles are dependent on both protein size and conformation; this allows for the detection of folding/unfolding events and the formation of large protein complexes. A Superdex 75 10/60 HR column from GE Healthcare was used for the analytical gel filtration experiments, and proteins were loaded using a 500 µL loading loop from GE Healthcare. 600 µL of sample was loaded for each experiment to ensure the loop had been filled and all experiments had the same volumes loaded.

Pressure Limit (MPa)	Flow Rate (mL / min)	Fraction Size (mL)
1.6	0.3	1.0

Table 3.24: *Operating parameters used for the Superdex 75 10/60 HR gel filtration column.*

3.7.2.1 gC1q-R Chemical Stability Experiments

The native fold of gC1q-R was examined by diluting a concentrated stock of gC1q-R to 10 µM with AGF Buffer. This sample was incubated at 37 °C for 10 mins and loaded onto a Superdex 75 column equilibrated in 2 CV of AGF Buffer. Attempts to study the chemical stability of gC1q-R were made by repeating the experiment whilst the column was equilibrated in 2 CV of Denaturing AGF Buffer. The elution profiles for both experiments were compared.

3.7.2.2 Zinc Binding Experiments

Zinc binding experiments were performed using the superdex 75 column equilibrated in 2 CV AGF Buffer. D5 protein stocks were diluted to 10 µM with AGF Buffer and increasing concentrations of zinc chloride. Protein solutions were incubated at 37 °C for 10 mins. A no zinc control was also run where 5 mM EDTA was added to the protein solution instead of zinc chloride.

3.7.2.3 gC1q-R : D5 Ligand Complexes

For the interaction studies of gC1q-R and D5, protein stock solutions were diluted to 10 µM using AGF Buffer. Proteins were initially run separately and acted as non-bound controls. In order to obtain protein complexes of varying molar ratios, the D5 concentration was kept at 10 µM and mixed with reducing concentrations of gC1q-R at 10, 5, and 3.33 µM trimer concentration. Protein samples were incubated at 37 °C for 10 mins prior to loading onto the superdex 75 column equilibrated in 2 CV of AGF Buffer. 1 mL fractions were collected and analysed using SDS-PAGE analysis as described in section 3.7.1. The methods used for observing the interactions of gC1q-R with D5-1 and D5-2 were similar to that described for full length D5. The only differences were that the column was equilibrated in either 2 CV of EDTA AGF Buffer, for D5-2; or Zinc AGF Buffer, for D5-1. Furthermore, D5-2 concentrations were increased to 20 µM to increase protein detection by SDS-PAGE gels.

3.8 Biophysical Techniques

3.8.1 UV Spectroscopy

Concentrations of both plasmid DNA and protein samples were calculated by measuring the absorbance at specific wavelengths using an ND-1000 Spectrophotometer NanoDrop. DNA was measured at 260 and 280 nm, whereas proteins were measured at 280 nm or 205 nm for the HMWK D5 peptides. A 2 µL buffer blank was initially run followed by 2 µL of DNA or protein sample.

3.8.2 Circular Dichroism

3.8.2.1 Sample Preparation

Protein stocks were diluted with CD Buffer to a final concentration of 0.1-0.5 mg/mL in a final volume of 500 µL. A dilution factor of at least ten was required in order to sufficiently reduce the concentration of Cl^- ions to produce minimal buffer signal at 180-200 nM.

3.8.2.2 Secondary Structure Prediction

400 µL of sample, 15µM D5, D5-1 and D5-1, and 5µM of gC1q-R trimer, was loaded into a 1 mm quartz cuvette from Hellma Analytics, and scans were collected at 190-260 nm at 25 °C using using a Chirascan-plus spectrometer from Applied Photophysics. An average of three scans were taken for each sample, and a buffer blank sample, consisting of CD Buffer alone, was subtracted from the averaged spectra. Protein secondary structure was then predicted using

CDNN software.

3.8.2.3 Zinc Binding Experiments

The effects of zinc binding on D5 and D5 truncations was analysed by comparing CD spectra run in the presence or absence of zinc. Protein samples were incubated at 37 °C for 10 mins in the presence of 1 mM EDTA or 75 µM zinc chloride. Spectra were recorded at 190-260 nm at 25 °C.

3.8.2.4 Thermal Melting Curves

Thermal melting curves were generated by running CD spectra between 10 and 90 °C at 5 °C intervals, and melting curves were generated by plotting CD at 222 nm against temperature. Accurate sample temperatures were recorded using the ESHU temperature probe.

3.8.3 Units

Units were converted from ellipticity to mean residue ellipticity (MRE) using the following equation:

$$M.R.E = \theta/(C \times N \times l) \tag{3.8.1}$$

Where θ is ellipticity in degrees, C is the protein concentration in M, N is the number of amino acids in the protein and l is the cell path length in m. M.R.E has the units deg.cm^2.dmol^{-1}. M.R.E gives the total ellipticity per mole of individual residue, and was used to compare CD spectra of proteins with a range

of different molecular weights.

3.8.4 Isothermal Titration Calorimetry

3.8.4.1 Sample Preparation

Proteins were dialysed against 1 L of either; ITC Buffer, ITC Zinc Buffer or ITC EDTA Buffer for 15 hours at 4 °C using a 1kDa MWCO mini dialysis from GE Healthcare. gC1q-R or gC1q-R mutants were diluted to 3-5 µM in a final volume of 2 mL using the same dialysis buffer, and D5 ligands were diluted to 80-100 µM: roughly 25 fold higher than the protein in the cell.

For the zinc titrations with D5 or the D5 truncations, potassium phosphate was replaced with 50 mM HEPES buffer to prevent precipitation of the large concentrations of zinc present. Protein samples were dialysed against 1 L of 50 mM HEPES, 100 mM NaCl pH 7.4, and the same dialysis buffer was used to make a 1 mM zinc chloride stock solution. Protein concentrations were diluted to 10 µM and zinc chloride was diluted to between 200 and 700 µM with the HEPES dialysis buffer.

For the competition experiments of D5-1 and D5-2 with gC1q-R, a 20 µM gC1q-R sample was saturated with 5 fold excess of either D5-1 or D5-2, and samples were loaded onto a superdex 200 gel filtration column, equilibrated in 1.5 CV of AGF Zinc Buffer, in order to remove any unbound ligand. The protein complex was concentrated to approximately 5 µM using a Vivaspin 10 kDa MWCO spin concentrator from GE Healthcare, followed by dialysis into ITC Zinc Buffer. This was then loaded into the cell and titrated with the apposing D5 ligand.

3.8.4.2 ITC Titrations

A MicroCal VP ITC system from Malvern was used for all ITC experiments. The cell was filled with 1.4 mL of gC1q-R or gC1q-R mutant whilst 300 µL of D5 ligand was loaded into the syringe. For the zinc binding experiments, the cell was filled with D5 ligand and the syringe with zinc chloride. The reference power was set at 5 µcal/sec, and the syringe stirring speed was set at 300 rpm. All ITC experiments were performed at 25 °C. An initial pre equilibration step of 1 h was set followed by 30 x 10 µL injections. Ligand dilution effects were tested by running a ligand to buffer control using the same titration parameters.

3.8.4.3 Binding Curve Fitting

ITC data were analysed using MicroCal Analysis software. The ligand to buffer control was subtracted from the experimental data, and any anomalous titration points were removed. Curves for D5-2 and HMWK 493-516 were fit to a one binding site model, whereas curves for D5 and D5-1 were fit to a three site sequential binding model.

3.8.5 NMR Experiments

3.8.5.1 Sample Preparation

D5 was concentrated to 200 µM using a Vivaspin 3 kDa MWCO spin concentrator from GE Healthcare. The concentrated protein sample was dialysed against 500 mL NMR Buffer for 15 hours at 4 °C using a 1kDA MWCO Mini Dialysis Kit from GE Healthcare. Prior to obtaining NMR spectra, the sample was centrifuged at 18,000 g for 10 mins to remove any insoluble aggregates.

3.8.5.2 Data Acquisition

All NMR experiments were performed on a Bruker 800 MHz Advance III Spectrometer using a QCI cryoprobe. Initial 1D experiments were collected in order to evaluate sample purity and confirm that protein aggregates had not formed. 1D proton spectra were acquired with 64 scans, acquiring 32,768 points in T1 over a spectral width of 15.02 ppm. All solvent suppression was performed using excitation sculpting.

^{15}N labelled samples were used to acquire 2D heteronuclear single quantum coherence (HSQC) spectra over a spectral width of 12 ppm in the 1H dimension and 38 ppm in the ^{15}N dimension. Spectra were acquired using 64 scans and 64 complex points in F1, and the nitrogen transmitter frequency was centred at 119 ppm. Data processing was carried out using Topspin 3.1.b.53 and CcpNmr Analysis 2.1.2 software.

3.8.6 Mass Spectrometry

3.8.6.1 Sample Preparation

Prior to analysis by mass spectrometry, proteins were desalted using a stacked 5 x 5 mL desalting column equilibrated in 3 CV of Desalt Buffer.

Pressure Limit (MPa)	Flow Rate (mL / min)	Fraction Size (mL)
0.5	0.3	10

Table 3.25: *Operating parameters used for the 5 x 5 mL desalting column.*

3.8.6.2 Mass Measurements

A SYNPAT High Definition Mass spectrometry system using electrospray ionisation and a quadrupole time of flight analyser was used to acquire the mass spectrometry data, and analysis was performed on Mass Lynx software. Protein samples were diluted to 20 µM using Desalt Buffer and, 100 µL was injected onto the mass spectrometer at 5 µL / min.

3.8.7 X-Ray Crystallography

3.8.7.1 Sample Preparation

A variety of different complexes were made in attempts to crystallise gC1q-R with various D5 ligands. Two different methods were used in order to produce the protein complex. The first involved mixing gC1q-R with 5 fold excess of ligand, followed by gel filtration (described in section 3.6.2.2) to separate the complex from the unbound ligand. The fractions corresponding to the protein complex were concentrated to a final concentration of 3-15 mg/mL using a Vivaspin 3 kDa MWCO spin concentrator from GE Healthcare. The second method involved directly mixing gC1q-R with varying ratios of ligand without the following gel filtration step. To ensure the non-gel filtrated samples were in comparable conditions to the gel filtrated samples, the complex was dialysed against 500 mL AGF Buffer using a Mini Dialysis Kit, 1 kDa cut-off from GE Healthcare. Samples made from both methods were snap frozen and stored at - 80 °C.

3.8.7.2 96 Well Screens

96 well crystallisation plates were set up using MRC Crystallisation Plates from Molecular Dimensions in conjunction with crystal screens purchased from Hampton Research. 100 µL of precipitant solution was added to each of the reservoirs, and 200 nL of protein solution was combined with 200 nL of each precipitant solution using a Mosquito Crystal liquid handler from ttplabtech. Trays were sealed using HD Clear Duck packaging tape, and crystals were left to grow at 20 °C.

3.8.7.3 24 Well Optimisation Screens

Optimisation screens were set up using 24 well VDX plates from Hampton Research. 1 mL of precipitant was added to each reservoir, and 1 µL of protein and precipitant were mixed in each well. The plates were sealed using HD Clear Duck packaging tape and crystals were left to grow at 20 °C.

3.8.7.4 Crystal Fishing and Freezing

5 µL of cryo protectant buffer was carefully added to the crystal. The cryo protectant buffer was made up of the specific crystallisation buffer plus 30 % glycerol (v/v) or, if any PEG was present in the crystallisation conditions, the PEG concentration was increased to 30-50 %. Crystals were then fished using MicroLoops from MiTeGen and immediately snap frozen and stored under liquid nitrogen ready for shipping.

3.8.7.5 X-Ray Diffraction and Modelling

Crystal X-ray diffraction was performed at the Diamond light source. Structures were created with ccp4 software by using molecular replacement against the wild type crystal structure (PDB:1P32). This solved structure was used as, due to the relatively small ligands being tested, the crystal structure for the ligand bound receptor would likely be very similar.

CHAPTER 4

The Interaction of D5 with the Major Endothelial Cell Receptor gC1q-R: Recruitment of HMWK to Endothelial Cells

4.1 Introduction

gC1q-R is the major endothelial cell for HMWK [118], but little detail is known about the domain responsible for gC1q-R binding. D5 has been shown to bind endothelial cell surfaces [78], and seeing as HWMK binds gC1q-R on the surface of endothelial cells [118] [231] D5 is likely the domain responsible. D3 has also been shown to bind endothelial cells [75], but due to the highly basic and acidic natures of D5 and gC1q-R respectively, it was proposed that D5 is more likely to comprise the major binding site for gC1q-R. Furthermore, peptides derived from D5 but not D3 were shown to inhibit the interaction of D5 with gC1q-R [107] providing evidence that gC1q-R interacts with D5 and not D3. The inter-

action of isolated D5 with gC1q-R was tested in order to further understand the binding of HMWK with endothelial cells and the potential mechanisms for cell surface receptor-dependent contact activation.

4.1.1 Aims

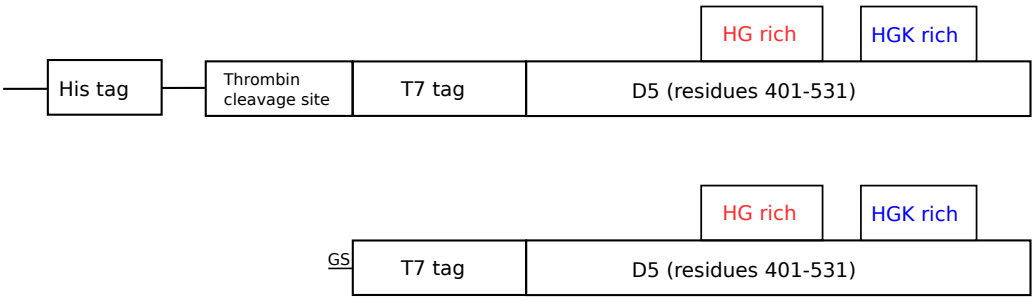
The aims of this chapter were to express and purify protein constructs of both human D5 of HMWK and gC1q-R. Once purified, D5 was characterised for its secondary structure and zinc binding ability in order to determine how metal binding affects the secondary structure. This was used to help understand the role zinc plays on the surface binding properties of HWWK.

gC1q-R was characterised for trimer formation and stability as this is key for the binding of HMWK. Once confident the receptor was forming a trimer in solution, attempts to detect and quantify the interaction with both proteins were carried out. An interaction of D5 and gC1q-R has been suggested in the literature but never has the direct interaction of isolated D5 been detected. Furthermore, the trimeric structure of gC1q-R could potentially allow multiple ligands to bind simultaneously: a theory that has never been fully explored. The data within this chapter provide the first direct evidence that D5 binds directly to this receptor, and the trimeric nature of gC1q-R and its significance for contact activation is explored. This has helped in further understanding the mechanism of contact activation.

4.1.2 Protein Constructs

4.1.2.1 Human D5

The domain boundaries for D5 are not well defined in the literature. Many different constructs have been used but they all span the His-rich region of HMWK differing only in the N and C terminal boundaries. The human D5 construct used in these experiments was supplied by collaborators and spanned residues 401-531 of full length HMWK. The D5 construct also contained a thrombin cleavable N-terminal His tag followed by a non cleavable T7 tag. The aa and DNA sequences for this construct can be found in section 9.1.1.1.



GS**H M A S M T G G Q Q M G R G S** T V S P P H T S M A P A Q D E E R D S G K E Q G H T
R R H D W G H E K Q R K H N L G H G H K H E R D Q G H G H Q R G H G L G H G H E Q Q
H G L G H G H K F K L D D D L E H Q G G H V L D H G H K H K H G H G H G K H K N K G K
K N G K H N G W K T E H L A S S S E D S

Figure 4.1: *Cartoon representation of the human D5 construct including the D5 aa sequence. The top panel represents the D5 construct before thrombin cleavage, whereas the bottom panel represents the D5 construct after His tag cleavage. The amino acid sequence of cleaved D5 is shown below the cartoon. The His-Gly-rich and His-Gly-Lys-rich regions are highlighted in red and blue respectively.*

4.1.2.2 Human gC1q-R

The gC1q-R construct consisted of residues 74-282 which was the same construct used to solve the crystal structure [154]. The N terminal region of gC1q-R is responsible for targeting the receptor to the mitochondrial membrane [145] and was left out of the construct to avoid any unwanted membrane binding properties and potential protein aggregation or solubility issues. The aa and DNA sequences for this construct can be found in section 9.1.1.2.

4.2 Results

4.2.1 Protein Purification

4.2.1.1 Purification of D5

D5 is a highly basic protein with a calculated pI of 9.4. The basic nature of this domain combined with the presence of an N-terminal His tag allowed for the purification of D5 using ionic exchange and affinity chromatography. Initial purification attempts consisted of Ni affinity chromatography as the first step, but it was observed that running an initial cationic exchange column greatly increased the binding affinity of His-tagged D5 to the nickel column. This is likely due to the removal of a large number of impurities which were initially blocking the His tag from effectively binding the nickel. Due to the large number of His residues present through out the protein, His-tagged D5 eluted at a high imidazole concentration of approximately 500 mM (Figure 4.2). The protein did not elute as one clean peak but as a series of peaks. It is possible that the disordered nature of this domain combined with the high His content resulted in a

distribution of different binding conformations to the nickel, and these different conformations had slight differences in their binding affinities.

Optimal thrombin cleavage of the His tag was achieved over 15 hours at RT in Thrombin Cleavage Buffer. The final purification step involved one final cationic exchange column using a shallow gradient spanning approximately 20 mS/cm. This step is important for the removal of the His tag, thrombin and any uncleaved D5 that may still be present. Similarly to the Ni affinity column, D5 did not elute as a single peak but a collection of peaks (Figure 4.3) which, once again, may be due to multiple binding conformations with the sulfopropyl groups. The yield of pure D5 produced using this method was approximately 3 mg/mL.

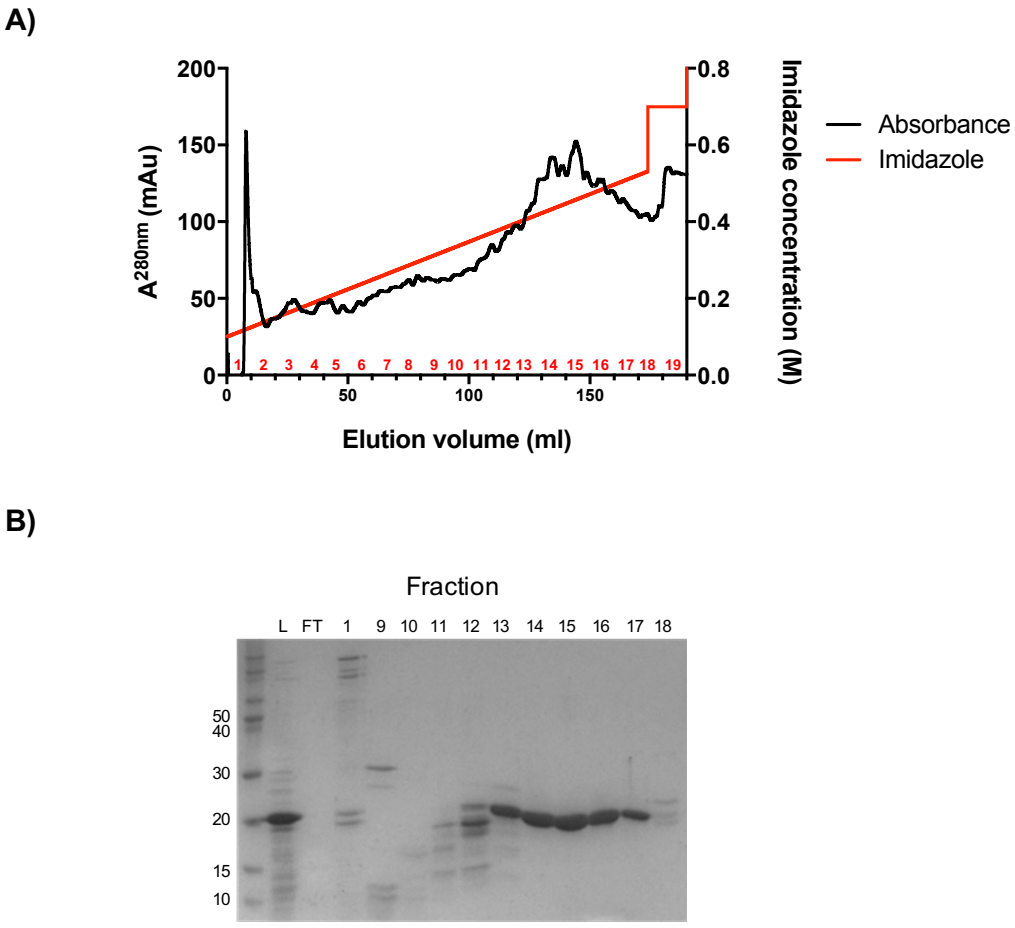


Figure 4.2: *Nickel affinity purification step of His-tagged D5.* ***A)*** *Absorbance at 280 nm (black line) and concentration of imidazole (red line) plotted against elution volume. 10 ml fractions were collected and fraction numbers are shown in red.* ***B)*** *SDS-PAGE gel showing the protein content of the load material (L), flow through (FT) and the various fractions from panel A collected for analysis. The first lane is the protein marker where useful sizes in kDa have been marked on the side. Pure His-tagged D5 can be seen in lanes 14-17 running at approximately 20 kDa.*

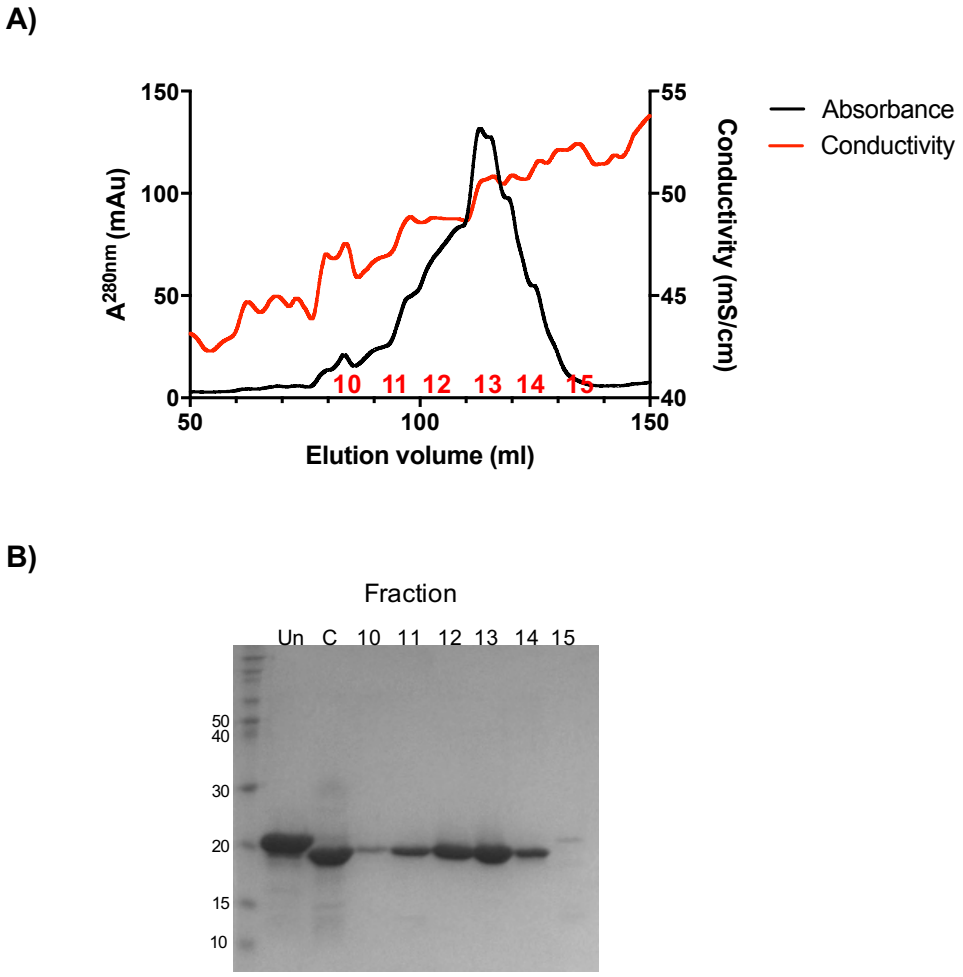


Figure 4.3: *Cationic exchange purification step of cleaved D5.* ***A)*** *Absorbance at 280 nm (black line) and salt concentration (red line) plotted against elution volume. 10 ml fractions were collected and fraction numbers are shown in red.* ***B)*** *SDS-PAGE gel showing the protein content of the uncleaved D5 (Un), D5 after thrombin cleavage (C) and fractions 10-15 from the cationic exchange column. The first lane is the protein marker where useful sizes in kDa have been marked on the side. Pure, cleaved D5 can be seen in lanes 10-14 running just below 20 kDa.*

4.2.1.2 Purification of gC1q-R

The first purification step for gC1q-R involved the binding of gC1q-R to an anionic exchange column. gC1q-R binds strongly to this column due to its acidic nature and elutes at approximately 45 mS/cm (Figure 4.4). A series of larger peaks elute shortly after the major gC1q-R peak and corresponds to DNA or other non-protein based impurities.

After anionic exchange, the protein was loaded onto a superdex 200 gel filtration column. The majority of gC1q-R eluted as a single peak (Figure 4.5); however, some larger molecular weight species of gC1q-R also elute from the column at an earlier volume. These high molecular weight species are spread over a larger volume and are likely due to large oligomers or aggregates of gC1q-R. To ensure only the single trimer species was collected, only the major peak, fractions 25-28 , was collected. The yield of pure gC1q-R produced using this method was approximately 15 mg/ L.

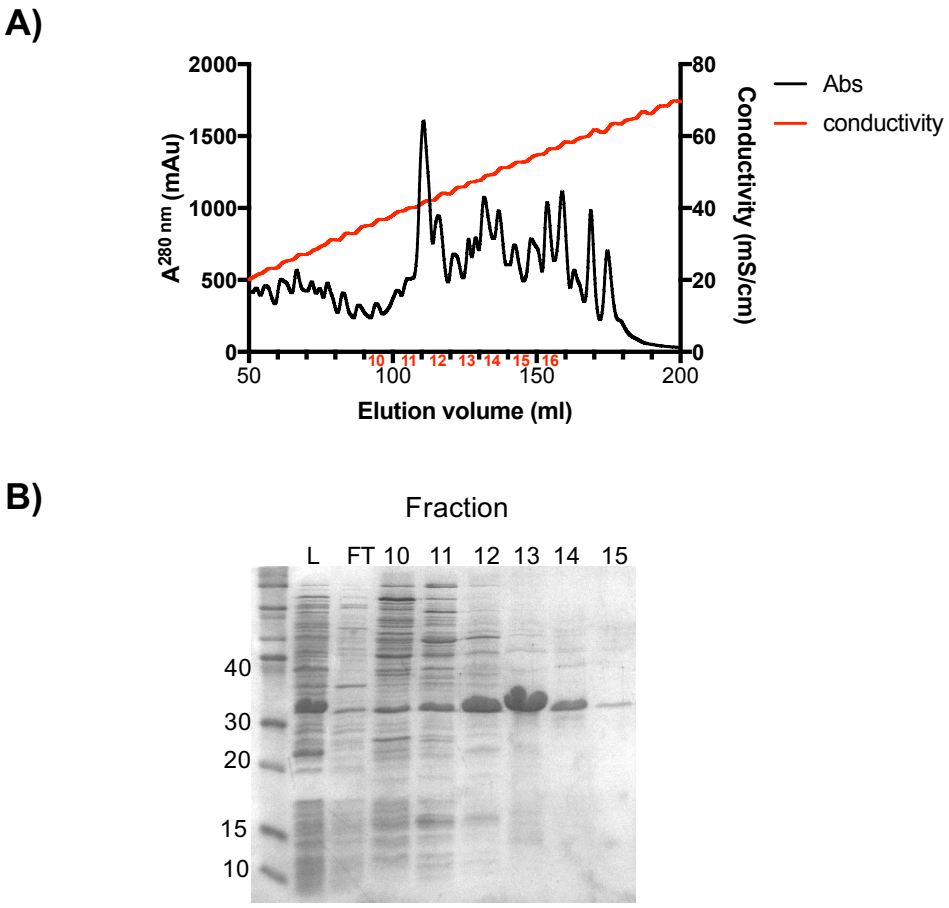


Figure 4.4: *Anionic exchange purification step of gC1q-R.* ***A)*** *Absorbance at 280 nm (black line) and salt concentration(red line) plotted against elution volume. 10 ml fractions were collected and fraction numbers are shown in red.* ***B)*** *SDS-PAGE gel showing the protein content of the load material (L), flow through (FT) and fractions 10-15 from the anionic column. The first lane is the protein marker where useful sizes in kDa have been marked on the side. gC1q-R can be seen running at 32 kDa in all lanes but is concentrated in lanes 12-14*

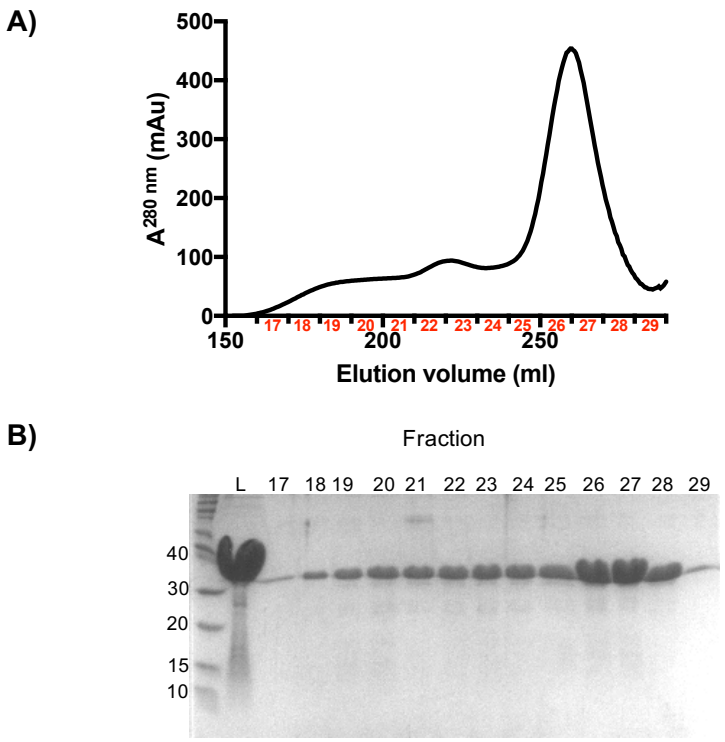


Figure 4.5: *Gel filtration purification step of gC1q-R.* ***A)*** *Absorbance at 280 nm plotted against elution volume. 10 ml fractions were collected and fraction numbers are shown in red.* ***B)*** *SDS-PAGE gel showing the protein content of the load material (L) and fractions 17-29 from the gel filtration column. The first lane is the protein marker where useful sizes in kDa have been marked on the side. Pure gC1q-R can seen at 32 kDa, and the major species is found in fractions 26-28.*

4.2.2 Characterisation of D5

4.2.2.1 Secondary Structure

Circular dichroism was used to determine the secondary structure of D5. Previous studies have shown D5 to be an intrinsically disordered protein [232] so it was important to confirm that the structure of the T7-tagged D5 construct was consistent with previous reports. The CD spectra collected was consistent with a largely random coiled domain indicated by a large minimum at approximately 195-200 nm.

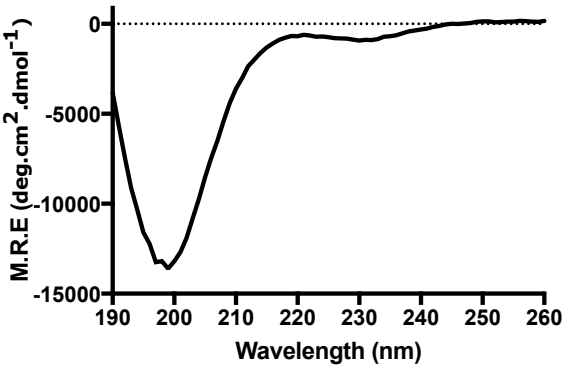


Figure 4.6: *Far UV CD spectra of D5 (15 μM) at 298 K. This spectrum is consistent with a largely disordered protein.*

4.2.2.2 Zinc Binding Propterties

The His-rich sequence of D5 has been linked with the ability to directly bind Zn^{2+}; however, there is conflicting evidence in the literature as to how zinc binding may affect the structure and folding of D5. Due to the large concentration of His residues, there is potential for multiple metals to bind throughout the entire length of D5. The domain readily precipitates in the presence of high zinc concentrations which is a likely a sign of non specific metal binding, result-

ing in the misfolding or aggregation of this domain.

Native mass spectrometry confirmed the binding of multiple zinc atoms to D5, with species bound to one, two and three Zn^{2+} ions detected (Figure 4.7). There were no peaks corresponding to unbound D5 suggesting this domain readily binds metal ions. provided the first direct evidence that D5 is able to bind multiple Zn^{2+} ions.

	Species A	**Species B**	**Species C**
MW (Da)	16517.96 ($\pm$1.62)	16579.50 ($\pm$4.30)	16455.33 ($\pm$1.85)
Number of bound Zn^{2+} ions	Two	Three	One

Table 4.1: *Calculated masses from the ESI-FT mass spectrum of D5 in the presence of zinc chloride.*

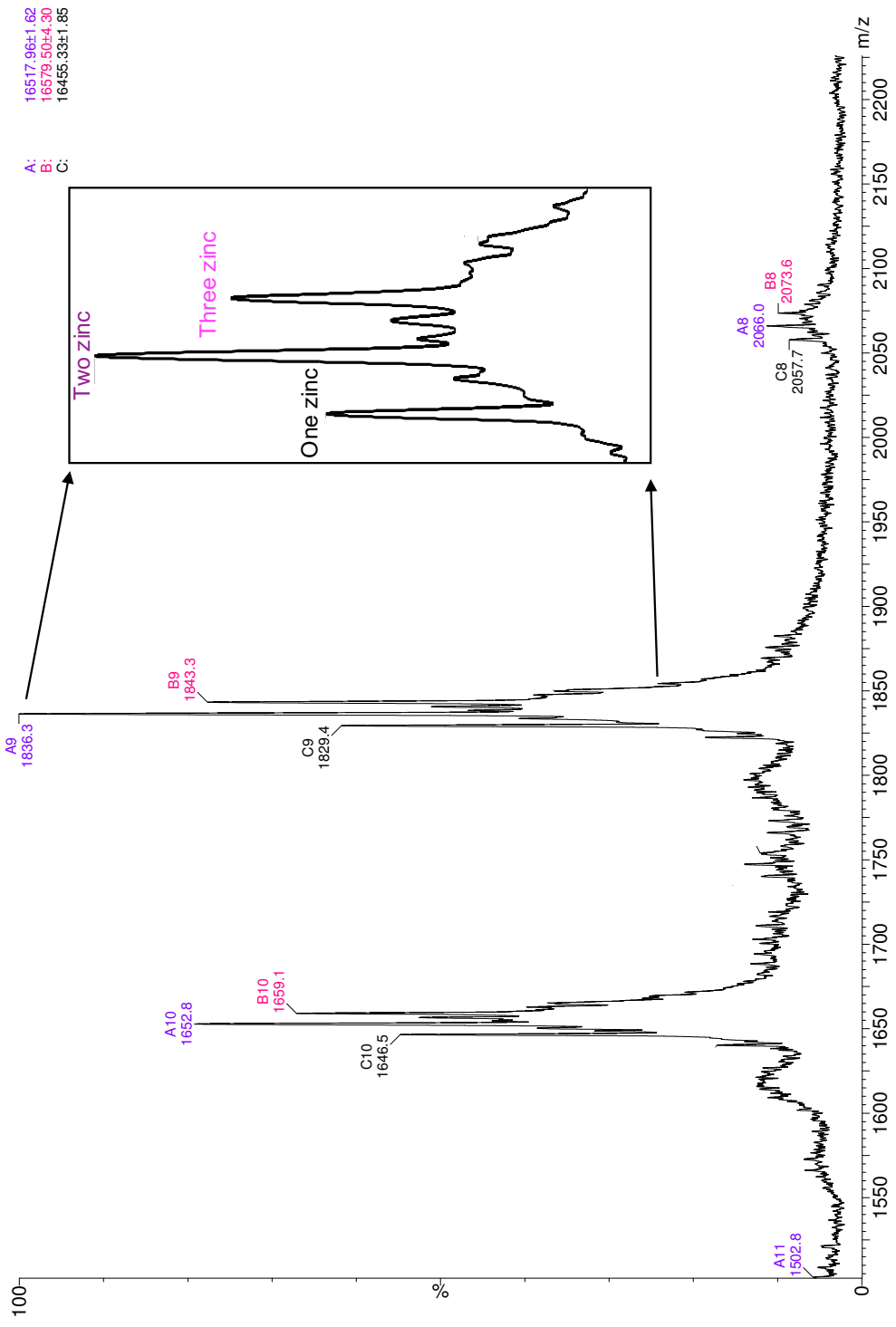


Figure 4.7: *Native ESI-FT mass spectrum of D5 (20 μM) in the presence of 100 μM zinc chloride. Three different species can been seen in this spectrum corresponding to D5 bound to one (species C), two (species A) and three (species B) zinc ions. The charge states and M/Z ratios are indicated above each peak, and the calculated mass for each species is shown on the right. A zoomed in image showing one set of peaks corresponding to the three species is also shown.*

ITC was used to further confirm the multiple zinc binding properties of D5. The ITC binding curve was fit to a one site model (Figure 4.8), and the thermodynamic parameters are shown in Table 4.2. These titrations were performed in 50 mM HEPES, 100 mM NaCl pH 7.4 instead of the phosphate based ITC Buffers documented in Table 3.3. This was to prevent large amounts of precipitation due to the high concentration of zinc present and the insolubility issues associated with zinc phosphate. Furthermore, these experiments relied heavily on knowing the exact zinc concentration at each injection; therefore, zinc phosphate precipitation would have resulted in a large discrepancy in the total zinc concentration. D5 was dialysed against 500 mL of 50 mM HEPES, 100 mM NaCl pH 7.4 and the $ZnCl_2$ was dissolved in the same buffer from the crystalline form.

These data further confirm the multiple zinc binding properties of D5, but unlike the three Zn^{2+} detected by mass spectrometry, ITC detected the binding of four Zn^{2+} ions. Discrepancies in stoichiometry could be a result of concentration errors during the ITC experiments or simply due to differences when working with D5 in the gaseous phase or the more native solution phase. Despite these differences in stoichiometry, both these techniques revealed that multiple Zn^{2+} ions bind directly to D5.

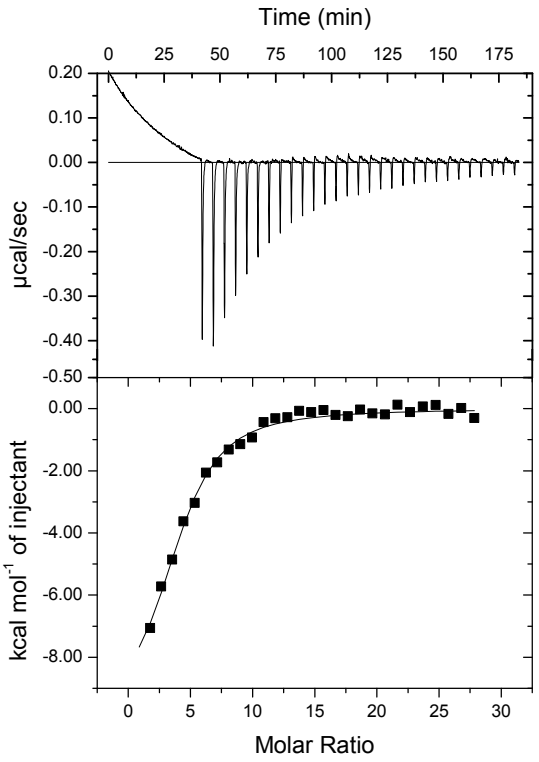


Figure 4.8: *ITC data for the titration of zinc chloride (700 µM) into D5 (10 µM). The top panel is the raw ITC data showing the exothermic release of heat upon zinc binding, and the bottom panel is the isotherm produced from the integration of the peaks.*

N	**K_D (µM)**	**ΔH (kJ mol^{-1})**	**ΔS (J K^{-1} M^{-1})**	**ΔG (kJ mol^{-1})**
4.0 (± 0.4)	1.5 (± 0.2)	-41.8 (± 5.9)	-28.5	-33.3

Table 4.2: *ITC-derived thermodynamic properties for the binding of zinc with D5.*

Despite mass spectrometry and ITC confirming the binding of Zn^{2+}, data collected by analytical gel filtration and CD suggest that metal binding does not induce any detectable folding within D5. No observable changes in D5 secondary structure was detected by CD in the presence or absence of zinc (Figure 4.9). Furthermore, analytical gel filtration elution profiles remained unchanged upon the titration of zinc (Figure 4.10). Metal-induced folding within

D5 would have resulted in an altered elution profile, and the lack of any observable change provided further evidence that metal binding does not affect secondary structure.

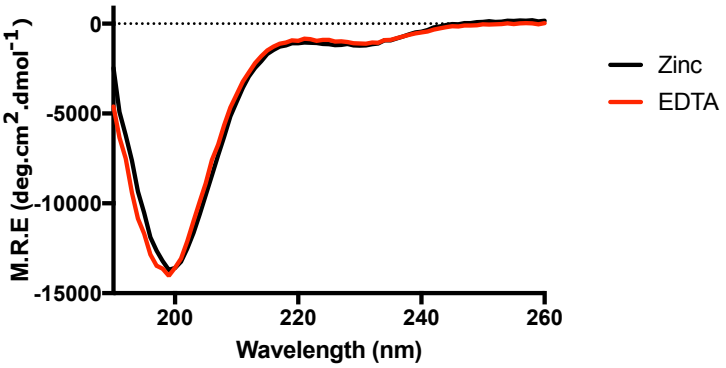


Figure 4.9: *Far UV CD spectra of D5 (15 μM) in the presence of 1 mM EDTA or 75 μM zinc chloride.*

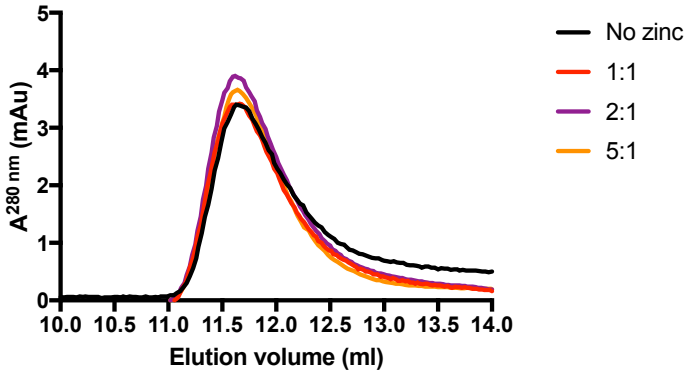


Figure 4.10: *Elution profiles of D5 (10 μM) in the presence of increasing concentrations of zinc chloride.*

4.2.3 Characterisation of gC1q-R

4.2.3.1 Chemical Stability of the Trimer

The molecular weight of gC1q-R in solution was calculated using analytical gel filtration. gC1q-R ran as a single species with a calculated molecular weight of

90 kDa. This is similar to the predicted trimeric molecular weight of 71.4 kDa, and the 20 kDa difference in calculated molecular weight is likely due to the less compact nature of the donut like structure, resulting in a slightly decreased elution volume when compared to the more globular protein standards. To obtain a more accurate molecular weight of the gC1q-R sample, analytical ultracentrifugation could be used. This method provides quantitative analysis of macromolecules in solution and calculates an accurate protein mass which is independent on the protein shape [233].

Attempts to break the trimer into the individual monomers proved difficult, and it was only under extreme denaturing conditions that any low molecular weight species could be observed (Figure 4.11). This low molecular weight species had a calculated molecular weight of 34 kDa which is approximately a third of the calculated weight of the trimer and, therefore, is consistent with the breakdown of the trimer to the monomer. Despite the presence of this smaller peak, the major species present was still the high molecular weight trimer suggesting that gC1q-R forms a stable trimer.

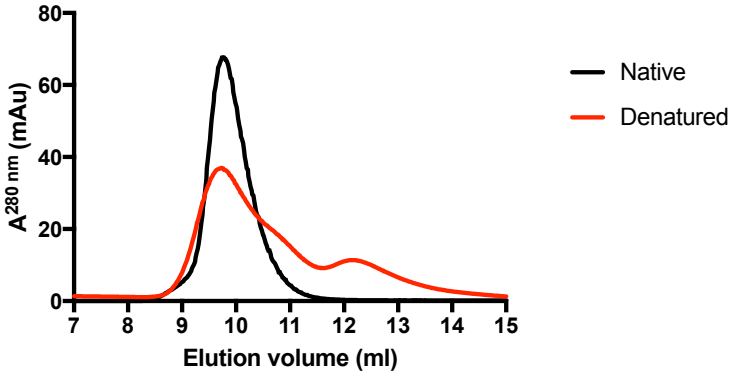


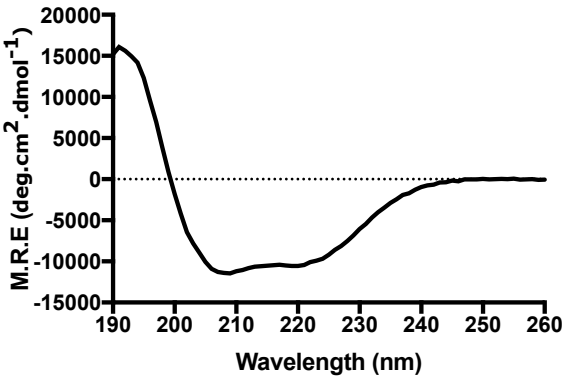
Figure 4.11: *Elution profiles of gC1q-R (10 µM trimer) in its native state (black) or in the presence of 0.1 M 2-mercaptoethanol and 0.1% w/v SDS (red). The peak corresponding to the trimer eluted at 9.9 ml, whereas the denatured monomeric peak eluted at 12.2 ml. The calculated molecular weights of the high and low molecular weight species are 90 and 34 kDa respectively.*

4.2.3.2 Secondary Structure and Thermal Stability by Circular Dichroism

CD was used to confirm that the secondary structure of purified gC1q-R was consistent with the crystal structure and assess the thermal stability of the protein. A key structural feature of this protein is the assembly of three monomers forming a symmetrical homo trimeric ring. The stability of this trimer can be examined by looking at the thermal unfolding profile of gC1q-R. The trimeric structure was predicted to have two distinct unfolding transitions: the first being the breakdown of the trimer into its monomeric components, and the second being the subsequent unfolding of the monomers. The stability of the trimer is concentration dependent, and once the concentration is reduced below the K_D of trimer formation the monomeric form will start to be the major species, resulting in a change in melting curve. Therefore, trimer stability can be assessed by comparing thermal denaturation spectra at reducing protein concentrations.

The CD spectra and thermal melting curves for gC1q-R are presented in Figure 4.12 and Figure 4.13 respectively. The calculated secondary structure from the CD spectrum is consistent with the crystal structure, with significant contributions from α helix, β sheet and random coil elements. gC1q-R undergoes a minor unfolding transition at approximately 26 °C, but no observable secondary unfolding event was detected at higher temperatures. There were minor changes in calculated secondary structure between gC1q-R at 10 °C and 30 °C with the major differences being a 10 % reduction in α helical content. This shift cannot be due to the breakdown of the trimer as no monomeric peaks were detected by analytical gel filtration experiments run at a similar temperature. This transition is likely due to slight unfolding of the N-terminal αA or C-terminal αC helices. Furthermore, this transition is relatively small and no further unfolding events occurred at higher temperatures suggesting the trimeric structure of gC1q-R is highly thermally stable.

Furthermore, reducing the gC1q-R concentration from 25 to 2.5 μM had little to no effect on the melting curve, showing that gC1q-R trimer formation has a low K_D and is stable at low concentrations. Concentrations could not be reduced past 2.5 μM as attempts to lower the concentration any further resulted in too weak a signal for obtaining accurate melting curves.



α Helix	Antiparallel *β*-Sheet	Parallel *β*-Sheet	*β*-turn	Random Coil
32.0 %	12.2 %	8.0 %	17.7 %	28.4 %

Figure 4.12: *Far UV CD spectra of gC1q-R (5 μM trimer) at 298 K. Predicted secondary structure elements are provided underneath the spectrum.*

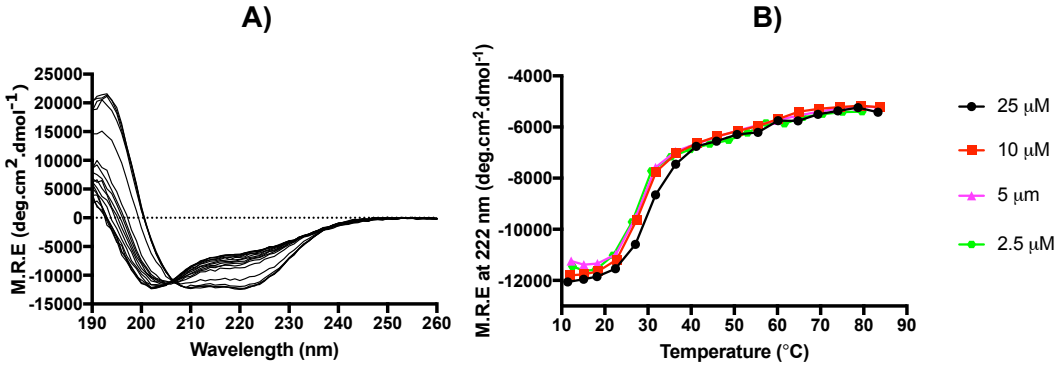


Figure 4.13: *Thermal denaturation CD spectra of gC1q-R.* ***A)*** *CD spectra at increasing temperatures from 10 to 90 °C.* ***B)*** *M.R.E at 222 nm plotted as a function of temperature at trimer concentrations of 25, 10, 5 and 2.5 μM.*

4.2.4 Detecting the interaction between D5 and gC1q-R

4.2.4.1 Analytical Gel Filtration

Analytical gel filtration was first used to identify the interaction between gC1q-R and D5 (Figure 4.14). There is an offset in elution of approximately 3 mL between both proteins; therefore, the presence of a binding interaction would result in an obvious shift of D5 into the higher molecular weight peak, whereas unbound D5 would remain in the low molecular weight peak. This observable separation of the bound and unbound protein species was also used to determine an approximate binding stoichiometry.

For these binding experiments, D5 concentrations were kept constant at 10 μM, whereas gC1q-R concentrations started at 10 μM (trimer) and were reduced in order to obtain different binding stoichiometries. In the presence of gC1q-R, D5 eluted in the high molecular weight peak and the unbound D5 peak disappeared. This provided the first direct evidence that isolated D5 binds to gC1q-R. Unbound D5 was only detected at a D5:gC1q-R ratio of 3:1; this suggested that multiple D5 ligands are able to bind one trimer. It was predicted that three D5 ligands would bind to the trimer, one per monomer, but the fact unbound D5 started to appear at a ratio of 3:1 could be due to the third binding event being too weak to detect by gel filtration.

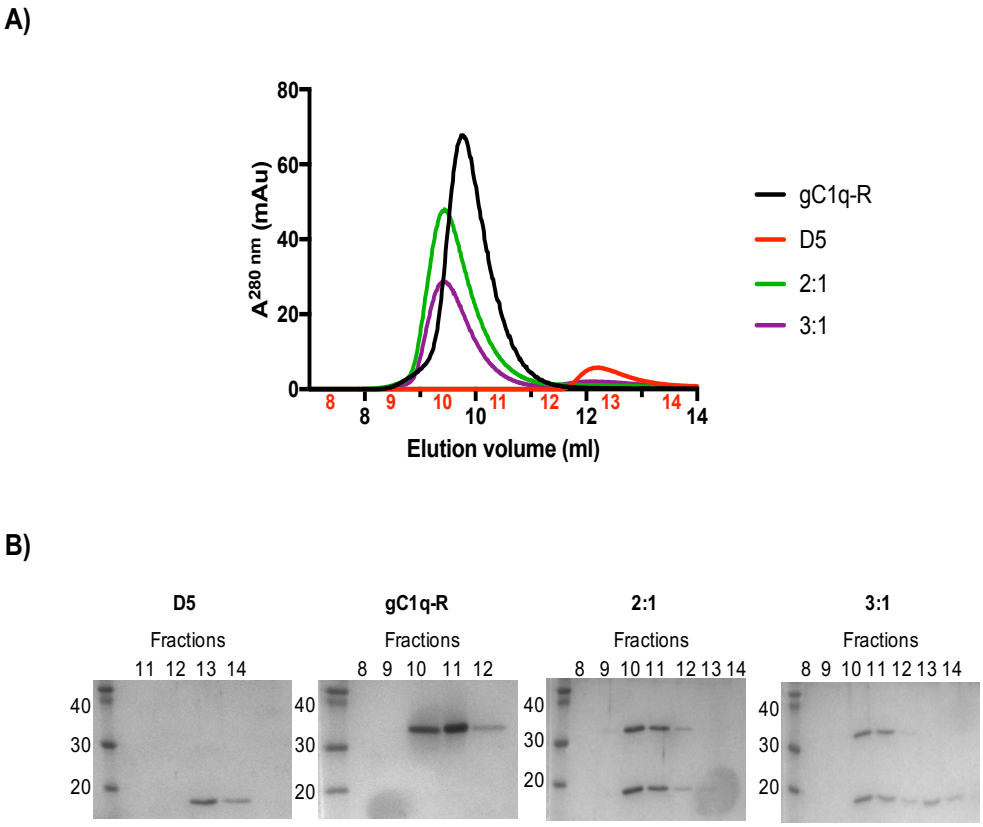


Figure 4.14: *Analytical gel filtration of D5 (10 μM) combined with 10, 5 and 3.3 μM of gC1q-R trimer to get varying molar ratios.* ***A)*** *Elution profiles of the various protein solutions with fraction numbers indicated in red.* ***B)*** *SDS-PAGE gels of the 1 ml fractions collected from the column showing the protein content of each peak. The first lane is the protein marker where useful sizes in kDa have been marked on the side. gC1q-R can be seen at approximately 32 kDa and D5 at just below 20 kDa.*

4.2.4.2 Isothermal Titration Calorimetry

Whilst analytical gel filtration was able to detect an interaction between D5 and gC1q-R, it was unable to provide quantitative information for the interaction of D5 with gC1q-R. ITC was used to calculate the binding affinity and thermodynamic profile for this interaction as well as to obtain a more accurate stoichiometry. The ITC data for the binding of D5 with gC1q-R is presented in Figure 4.15.

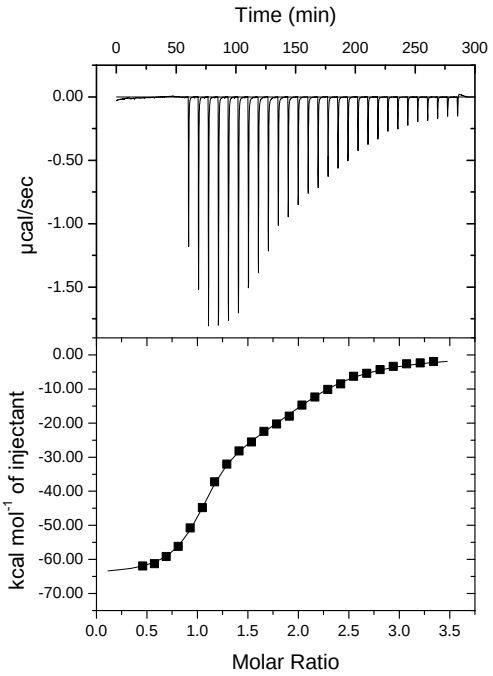


Figure 4.15: *ITC results when D5 (100 µM) is titrated into gC1q-R (5 µM trimer). The top panel is the raw ITC data showing the exothermic release of heat upon the injection of D5, whilst the bottom panel is the isotherm produced from the integration of the peaks.*

The titration of D5 with gC1q-R resulted in a strong response with a large exothermic release of heat. From the binding curve, it was clear that this interaction does not occur through a single site model, and there are clearly multiple, distinct binding events. Figure 4.16 shows the binding curve fit to a variety

of different binding models. A one site binding model fits these data poorly, whereas a two site sequential model provides a much better fit than the one site model but fits poorly to the well defined second binding event. The best fit for these data was the three site sequential model where three separate K_D values can be extrapolated from the one binding curve. The binding affinities decrease from the first to the third binding event due to negatively cooperative sequential binding [234]. A summary for the sequential mechanism is shown in equation 4.2.1.

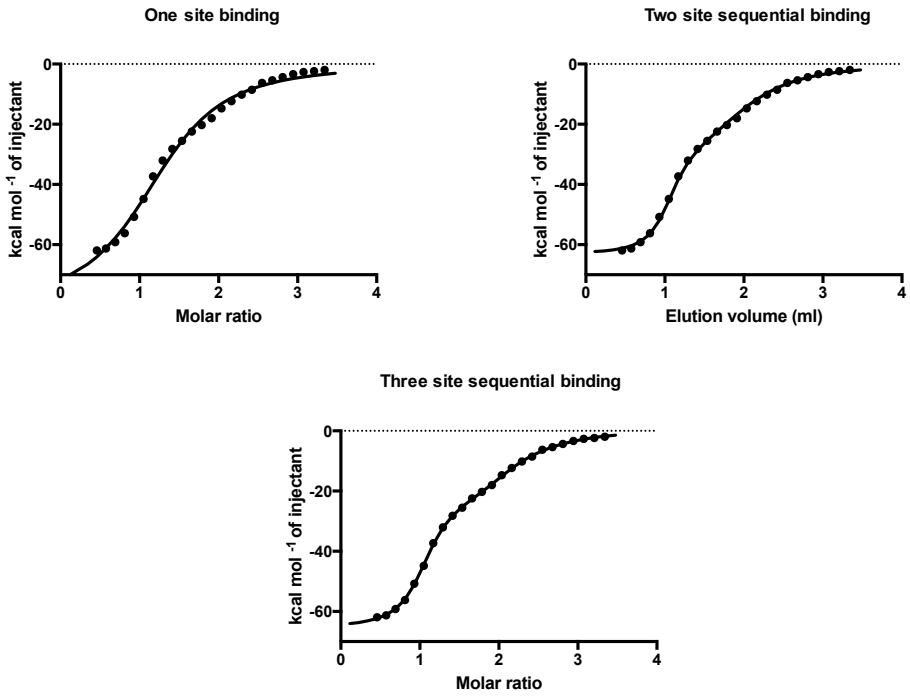


Figure 4.16: *The ITC binding curve produced by the titration of D5 with gC1q-R fit to a one site binding model (top left), two site sequential binding model (top right) and a three site sequential binding model (bottom). The three site sequential binding model, where three separate K_D values are generated, gave the best fit to these data.*

$$K_{\mathrm{D}}1 = \frac{[gC1q - R][D5]}{[gC1q - R : D5]} \quad < K_{\mathrm{D}}2 = \frac{[gC1q - R : D5][D5]}{[gC1q - R : D5_2]} \quad < K_{\mathrm{D}}3 = \frac{[gC1q - R : D5_2][D5]}{[gC1q - R : D5_3]} \tag{4.2.1}$$

	K_D (nM)	ΔH (kJ mol^{-1})	ΔS (J K^{-1} M^{-1})	ΔG (kJ mol^{-1})
First binding event	1.9 ($\pm$ 0.1)	-268.8 ($\pm$ 1.1)	-732.7	-50.5
Second binding event	64.9 ($\pm$ 1.9)	-105.2 ($\pm$ 1.3)	-214.8	-41.2
Third binding event	1011.1 ($\pm$ 75.8)	-23.5 ($\pm$ 1.4)	36.0	-34.3

Table 4.3: *ITC-derived thermodynamic properties for the three site sequential binding of D5 with gC1q-R.*

The thermodynamic properties of each binding event are presented in Table 4.3. The first two binding events have an unfavourable decrease in entropy, but a large decrease in enthalpy is enough to compensate for this. Typically, large negative enthalpy changes are suggestive of favourable electrostatic or hydrogen bonding [235] which is consistent with both D5 and gC1q-R being highly charged proteins, and it's likely that there are multiple electrostatic interaction patches involved in binding. ΔS is less negative for the subsequent binding events and becomes positive during the third binding event. Negative ΔS values are caused by the loss of conformational flexibility [236]; as both D5 and gC1q-R are either highly disordered or have regions of disorder, this large decrease in entropy is likely due to increased ordering within particular flexible regions involved in binding. This effect is reduced as the binding affinity becomes weaker and these regions are no longer as tightly bound and restricted. The final binding event has a positive ΔS which may reflect the release of solvent perhaps due to some hydrophobic effect [237]. There is only a small increase in entropy, and the first two binding events may have similar hydrophobic effects but the increased ordering of D5 upon binding results in an overall negative ΔS. The relationship between ΔH and ΔS for the different binding events is summarised in Figure 4.17.

The negative cooperativity observed could be caused by increased steric hinderance at the binding site, or alternatively allosteric modification of the receptor upon D5 binding could reduce the affinity of D5 to the neighbouring

binding sites. These would both act as regulatory mechanisms for the binding of HMWK and FXII with gC1q-R, preventing one protein from binding all available sites within gC1q-R.

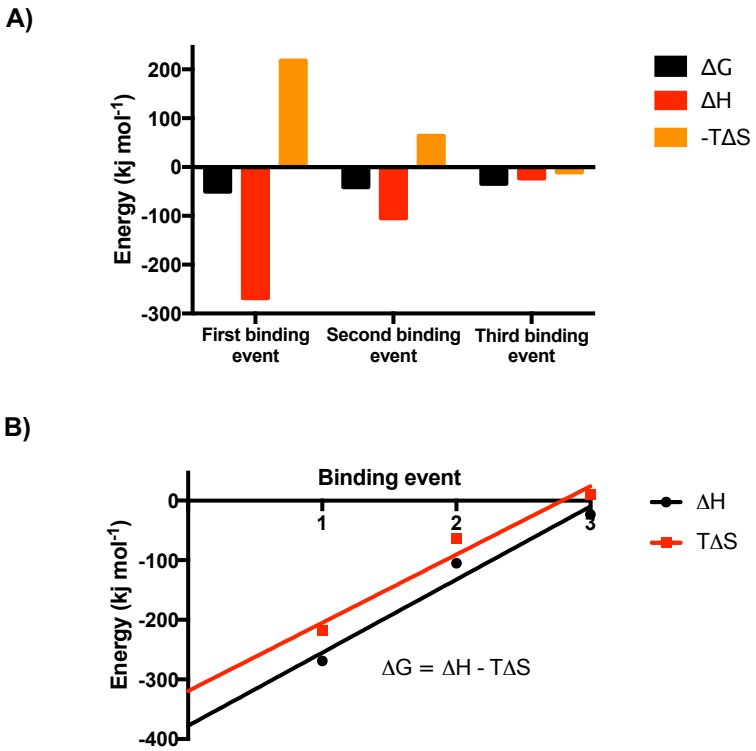


Figure 4.17: *The relationship between ΔH and ΔS for the three binding events of D5 with gC1q-R.* ***A)*** *Chart showing ΔG, ΔH and -TΔS for each binding event. The first two binding events consist of a favourable negative ΔH and unfavourable positive -TΔS. This suggests that these binding events are enthalpy driven which is characteristic of electrostatic interactions and favourable hydrogen bond formation. The third binding event is associated with a small increase in ΔS and is both entropy and enthalpy driven.* ***B)*** *Ethalpy-entropy compensation plot where ΔH and TΔS values are plotted for each binding event. This plot highlights the proportional relationship between enthalpy and entropy as more D5 ligands bind gC1q-R. ΔG is consistent between binding events and is equal to ΔH-TΔS.*

Thus far, all the interaction experiments were performed in the absence of any zinc. In order to investigate the effects of zinc on the interaction of D5 with gC1q-R, the ITC experiment was repeated in the presence of zinc or EDTA. It is important to note that these ITC experiments were performed in phosphate buffer and many divalent cations, including zinc, precipitate readily in the pres-

ence of phosphate. Unlike the zinc titrations documented in section 4.2.2.2 however, the zinc buffer could not be switched to HEPES, or alternate buffers such as Tris, due to the much higher enthalpy of protonation associated with these buffers [238] resulting in large buffer effects when combined with the highly electrostatic interaction of gC1q-R and D5. These effects were not an issue when detecting the much weaker interactions of zinc with D5. Due to this, the ITC Zinc Buffer contained both 20 mM KP and 50 μM $ZnCl_2$; therefore, some of the zinc would have precipitated as zinc phosphate. The anion responsible for precipitating the zinc is the PO_4^{3-} which would account for less than 0.01 % of the total phosphate species in the solution [239]. This is still is enough to cause significant precipitation within the buffer but does not mean all the free Zn^{2+} precipitated out of solution. Both D5 and gC1q-R were dialysed against 500 mL of buffer and, therefore, overtime they would become saturated with Zn^{2+} irrespective of the reduced concentration of zinc in the main buffer solution. This hypothesis was confirmed by detecting the zinc-dependent interaction of gC1q-R and D5-1 (documented in section 5.2.3) after extensive dialysis against ITC Zinc Buffer using analytical gel filtration. This suggests that, despite the issue of zinc phosphate precipitation, enough free zinc was present in the buffer for the interaction to occur.

The same dialysis was performed with a large excess of EDTA to remove any zinc that may have been present. No significant changes between the gC1q-R and D5 isotherms were detected between the zinc and EDTA buffers providing evidence that the presence of Zn^{2+} does have an effect on the interaction of isolated D5 with gC1q-R.

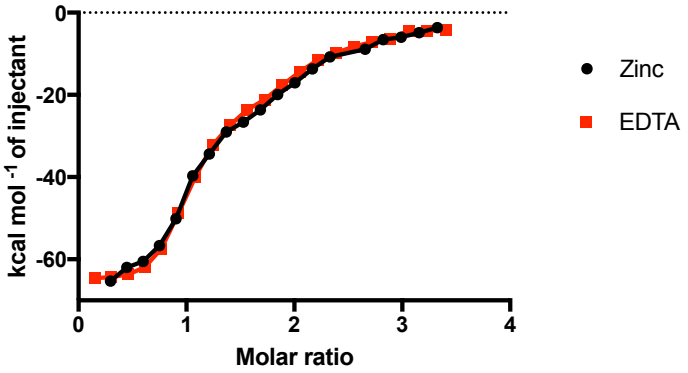


Figure 4.18: *ITC isotherms produced when D5 was titrated into gC1q-R after saturation with zinc (black) or in the presence of 1 mM EDTA (red). These data confirmed the zinc-independent nature of this interaction.*

4.2.4.3 gC1q-R Induced Folding of D5 by 2D NMR

An important question regarding this interaction is if binding would induce a structural change within D5 and lock the domain into a more structured and well folded orientation. It is common for disordered proteins to fold upon binding to other proteins, either within distinct regions or throughout the entire protein [240].

2D NMR spectra of ^{15}N labelled D5 with or without unlabelled gC1q-R were obtained in order to try and observe any structural changes or additional folding upon binding (Figure 4.19). The unbound D5 spectrum has a large degree of peak overlap, partially due to the large number of His, Lys and Gly residues but also due to the intrinsically disordered nature of this domain. A total of 61 peaks are observed in the spectrum; this covers approximately 40 % of the expected total peak number for D5. This poor coverage is common when studying disordered proteins as whilst the protein is switching between multiple conformations, some of the nuclei will broaden due to intermediate chemical exchange [241]. This broadening can become so severe that the peak is not visible in the NMR spectrum.

Upon the addition of gC1q-R, a large number of peaks move but there is no obvious folding of D5 as the anticipated change in peak dispersion, characteristic of formation of a folded, globular domain, was not observed. There are even fewer peaks present in the bound form, with only 30 % total coverage. The further reduction in the number of peaks is potentially due to intermediate exchange between the bound and unbound form causing more peaks to broaden and disappear. Efforts to assign the HSQC spectrum by performing additional 2D and 3D NMR experiments were not made due to the extremely poor coverage and high levels of peak overlap present.

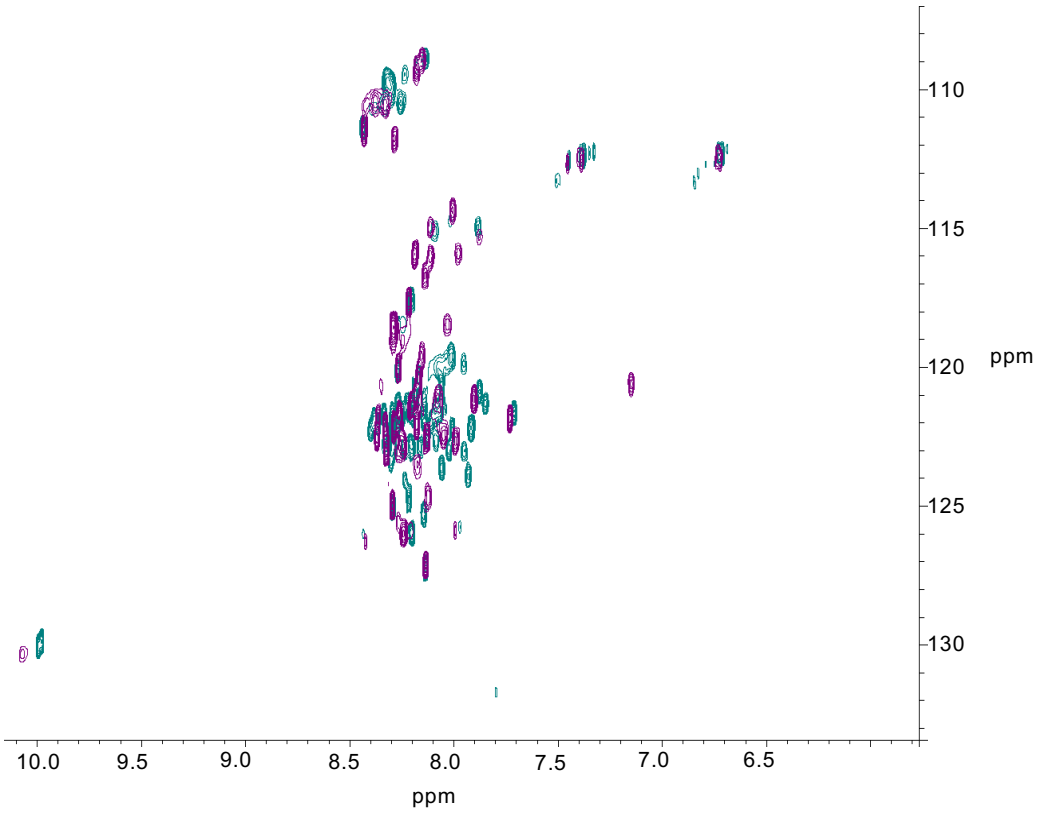


Figure 4.19: *HSQC NMR spectra of ^{15}N labelled D5 (200 µM) in the absence (green) or presence (purple) of unlabelled gC1q-R (200 µM). The high degree of peak overlap, both in the presence and absence of gC1q-R, is suggestive of a disordered protein.*

4.3 Discussion

D5 is an intrinsically disordered protein and does not undergo any detectable structural change upon binding to zinc. This is in contrast with data suggesting that zinc induces a conformational change within D5 [242]. A past experiment, looking at the shift in Trp fluorescence within D5, detected an emission shift upon the addition of zinc [243] which was proposed to correspond to a conformational change within the protein. However, looking closer at the D5 sequence, the two Trp residues are next to His-rich sequences and, due to the multiple zinc binding regions within the sequence, the shift in fluorescence is likely due to direct binding of the metal as opposed to a metal induced conformational change. The binding of zinc to D5 can be described as like having 'beads on a string' where the binding of the metal doesn't have a significant effect on the secondary structure. Zinc binding could serve a more significant structural function within full length HMWK as the N and C-terminus of D5 are no longer exposed but bound to the rest of the protein, resulting in a more rigid and confined structure compared to isolated D5. Although the exact metal binding residues remain unknown, the large number of His residues found throughout D5 are expected to be responsible for the binding of the multiple Zn^{2+} ions. Histidine is the most common binding residue in zinc catalytic sites and the second most common, after cysteine, in structural zinc sites [244]. There are no cysteine residues in the D5 sequence so zinc binding cannot occur through a classic Cys_2His_2 zinc finger motif [245]. The binding of zinc with feredoxin is important for stabilising the proteins structure, and binding occurs through three His and one Asp residue [246]. D5, which has a total of eight Asp residues, could be binding through a similar mechanism with His and Asp residues coordinating zinc in a tetrahedral arrangement. There are also cases of three His residues and a water molecule co ordinating to one zinc, including carbonic anhydrase II [247] and DD carboxypeptidase [248]; however, unlike D5 both these zinc atoms are bound within catalytic sites. The coordination

of four His residues to one metal atom is not well documented but could be possible due to the high frequency of His residues throughout the domain. A common motif which is repeated throughout D5 is the HG_XHG_XH motif where X is either one residue, normally a Gly, or three alternate residues. It is therefore possible for metal binding to be localised to one of these short binding motifs as opposed to the more wide spread binding found in class II zinc binding motifs [249]. Zinc binding is likely coordinated by either three His residues or three His and one Asp residue. To determine the key zinc coordinating residues, mutation studies will have to be performed in attempts to knock out specific metal binding sites.

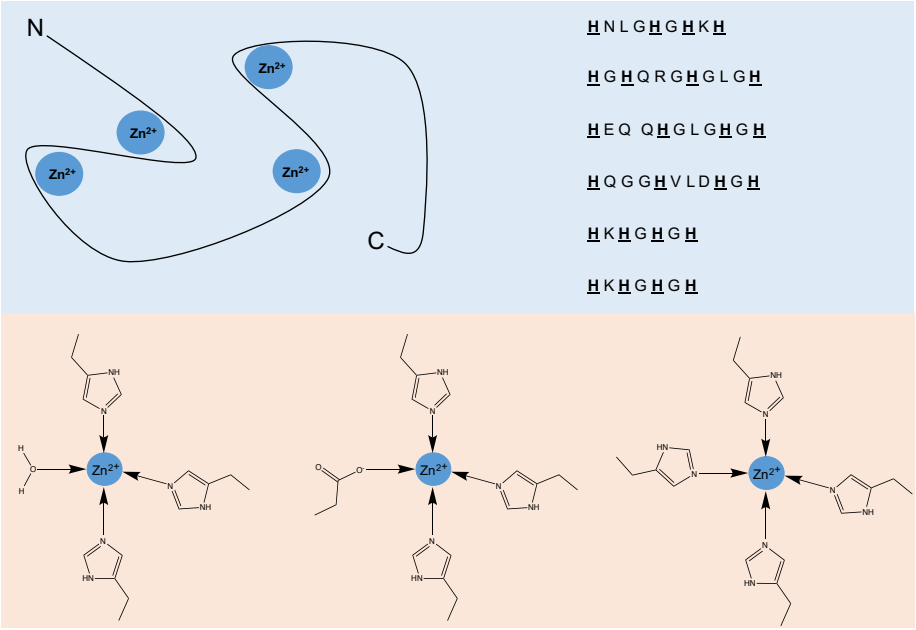


Figure 4.20: *The 'beads on a chain' model for zinc binding to D5 with HG_XHG_XH motifs found within the D5 sequence shown to the side. Potential zinc bound structures are presented in the bottom panel.*

The sequential binding mechanism of D5 and gC1q-R is likely a key mechanistic feature of contact activation. Considering that gC1q-R has a very stable symmetrical trimeric structure, it is not surprising that multiple D5 domains are binding to one receptor as the presence of one binding site on one monomer means there are two identical sites on the other two monomers. However, the

significant observation from this work is the demonstration of negative cooperativity for the multiple binding events of D5 to gC1q-R. This negative cooperativity observed is due to a large increase in ΔH which outweighs the favourable increase in ΔS, and this mechanism could be important for the assembly of FXII at the same surface. Both these proteins compete for a similar binding site on endothelial cells, and the reduced binding of the second and third D5 to the endothelial receptor would allow FXII, with its weaker binding affinity, to bind to one of the monomers. With HWMK and FXII bound to one receptor surface it is now possible to present all the initiator proteins at one endothelial cell surface and trigger contact activation.

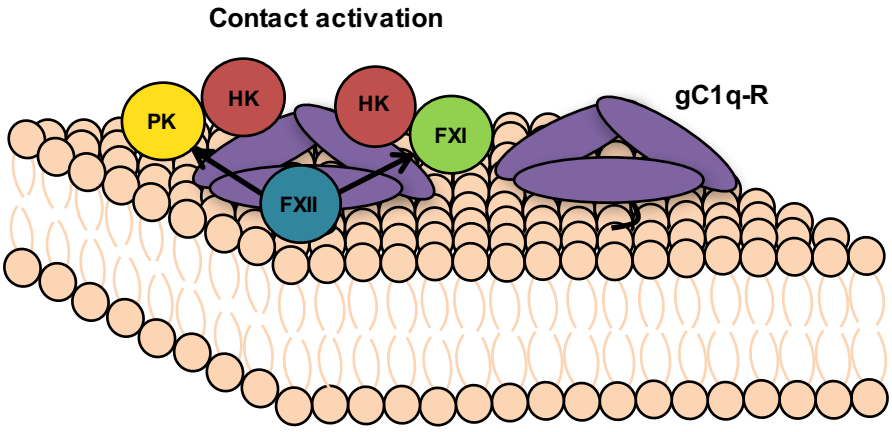


Figure 4.21: *Cartoon diagram of the proposed model for the assembly of FXII, FXI and prekalikrein at the endothelial cell surface through the interaction with HMWK (HK) and gC1q-R.*

Additionally to the observation of negative cooperativity, the interaction of D5 with gC1q-R was shown to be zinc-independent. This is in contest with previous work on full length HMWK that showed that the interaction with gC1q-R required the presence of zinc [106] [117]. Zinc binding may be more critical for the binding of full length HMWK than for isolated D5 as it might help expose key binding sites within D5, something which is not necessary for isolated D5 with its vastly increased flexibility and less confined structure. To test this hypothesis the interaction of gC1q-R with full length HMWK will need to be

performed, and these experiments were not performed due to the high production costs of full length HMWK.

It may seem strange that the strong interaction between D5 and gC1q-R did not induce folding within D5 when studied by NMR, but structural disorder within protein complexes is now more widely appreciated [250] and serves an important biological function. Intrinsically disordered proteins are particularly common in cell signalling and regulation [251] where their flexible structure allows them to bind a series of different ligands and surfaces. Disordered proteins have also been linked with optimal allosteric modulation of protein interactions [252] which is consistent with the negative cooperativity detected between D5 and gC1q-R. The binding cooperativity shown by the disordered D5 domain is important for the 'molecular hub' like properties of HMWK [253] where its major role is to assemble multiple proteins at one surface. Taking all this into account, it is clear that the disordered nature of D5 plays a pivotal role in contact activation.

CHAPTER 5

D5 Truncations: Two Major gC1q-R Binding Regions

5.1 Introduction

To understand the binding interaction of D5 and gC1q-R in more detail it was important to pinpoint the key regions and residues within D5 responsible for binding. D5 has two major regions: an N-terminal His-Gly-rich region and a C-terminal His-Gly-Lys-rich region. The N-terminal region has a neutral charge with a calculated pI of 7.5, whereas the C-terminal region is highly basic due to the large number of Lysine residues and has a calculated pI of 9.4. Due to the acidic nature of gC1q-R, it was predicted that it would be this basic C-terminal region of D5 that would hold the key binding properties for the receptor, and previous experiments showed that a Lys-rich peptide from this region inhibited the interaction of HMWK with gC1q-R [107] as well as the binding of FXII to HUVEC cells [46].

There is also evidence suggesting that the N-terminal region of D5 is partially

responsible for the anionic surface binding properties of HMWK. An N-terminal His-rich peptide of D5 was shown to inhibit the procoagulant activity of HWMK and the zinc-dependent binding of HMWK to kaolin [84]. Further deletion studies within D5 identified at least two anionic surface binding sites: one within the His-Gly-rich region and a second within the His-Gly-Lys-rich region [70]. It is possible that, similarly to past surface binding experiments, both the N and C-termini are able to bind to gC1q-R.

5.1.1 Aims

The binding of D5 to cell surfaces does not appear to occur through one binding site but a number of different sites present throughout the sequence. In order to understand the roles of both the N and C-terminal regions of D5 in gC1q-R binding, a major truncation was made which split D5 into two fragments, separating the N-terminal region from the C-terminal region. Using similar binding experiments to those used for full length D5, the roles of both the His-Gly-rich region and the basic His-Gly-Lys-rich region in receptor and zinc binding were explored. The domain was further segmented through the purchase of three His-Gly-rich peptides which were used in conjunction with the larger truncations to pin point the key binding regions for gC1q-R.

5.1.2 D5-1 and D5-2 Protein Constructs

D5 was truncated into two smaller fragments, resulting in an N terminal truncation rich in His and Gly residues, named D5-1; and a basic C terminal truncation rich in Lys residues, named D5-2. Due to the relatively small size of both these truncations, they were both cloned as thrombin cleavable GST fusion proteins in order to make initial visualisation easier by SDS-PAGE; furthermore,

it helped with the purification of the neutral D5-1 which would not bind as strongly to an ionic exchange column compared with D5 or D5-2. Having the GST tag added an extra purification step which was used to improve the purification procedure. Cartoon representations of both constructs are shown in Figure 5.1, and the aa and DNA sequences for D5-1 and D5-2 can be found in sections 9.1.1.3 and 9.1.1.4 respectively.

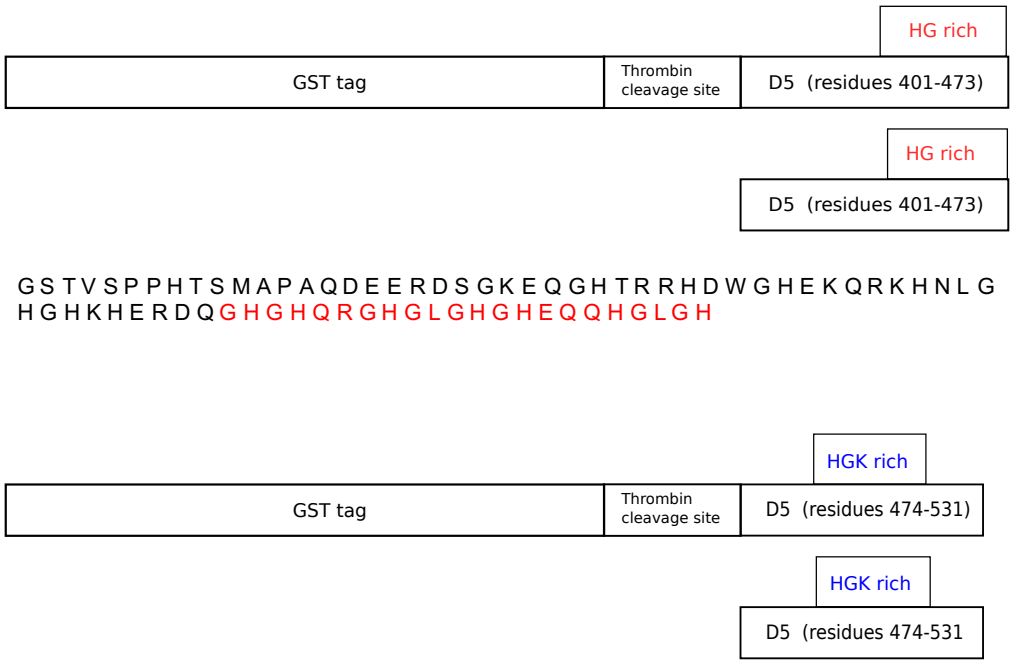


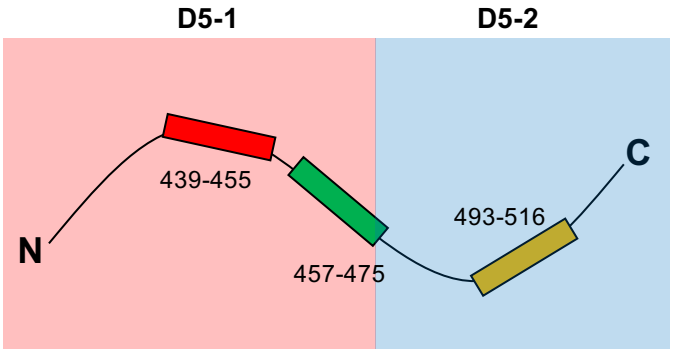
Figure 5.1: *Cartoon representation of the D5-1 and D5-2 constructs outlining the position of the GST tag, thrombin cleavage site and HMWK residues. The top and bottom images represent the D5-1 and D5-2 constructs before and after thrombin cleavage respectively, and the amino acid sequence for the cleaved products are shown underneath each cartoon with the His-Gly and His-Gly-Lys-rich regions highlighted in red or blue respectively.*

	GST-tagged Molecular Weight (kDa)	Cleaved MW (kDa)	pI
D5-1	34.6	8.4	7.5
D5-2	32.8	6.6	9.4

Table 5.1: *Theoretical molecular weights of tagged and cleaved D5-1 and D5-2 along with the calculated pI of each cleaved protein.*

5.1.3 D5 Peptide Constructs

The three initial peptides used corresponded to His-Gly-rich repeats found throughout D5. Two of these peptides were located within the N terminal D5-1 region, named HMWK 439-455 and HMWK 457-475 respectively, whilst the final peptide comprised a His-Gly-Lys-rich sequence found within the D5-2 region, named HMWK 493-516. HMWK 457-475 was previously been shown to be involved in zinc-dependent surface binding. These peptides were used in attempts to narrow down the gC1q-R binding site to shorter regions of D5.



HMWK 439-455 - H N L G H G H K H E R D Q G H G H

HMWK 457-475 - R G H G L G H G H E Q Q H G L G H G H

HMWK 493-516 - H G H K H K H G H G H G K H K N K G K K N G K H

Figure 5.2: *Cartoon representation of the three His-Gly-rich peptides from D5. The red and blue squares represent the regions which lie within the D5-1 and D5-2 boundaries respectively. The aa sequences for each peptide are shown below the cartoon.*

5.2 Results

5.2.1 D5-1 and D5-2 Purification

GST-tagged D5-1 and D5-2 were first captured using glutathione sepharose beads, and any unbound protein impurities were washed from the column. The GST tag was cleaved with thrombin followed by the elution of the cleaved D5-1 and D5-2 products. Elution profiles and SDS-PAGE gels for the glutathione purification step are shown in Figure 5.3. The binding of both GST-tagged fusion proteins with the glutathione beads was extremely poor with most of the material running straight through the column into the flow through. This poor binding efficiency is likely due to the flexible D5-1 and D5-2 inserts blocking the glutathione binding site on the GST tag, preventing efficient binding. Despite this, the overall yield of cleaved protein collected from this step was sufficient to carry on to the next purification step so no further attempts to improve binding efficiency were performed.

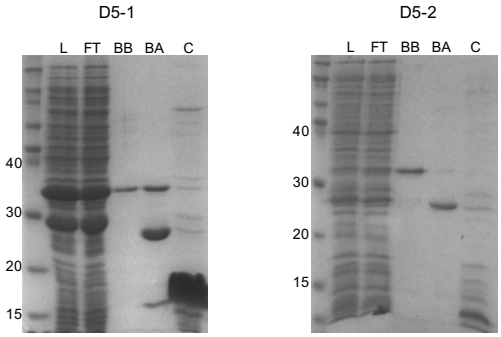


Figure 5.3: *SDS-PAGE gel of fractions taken from the glutathione sepharose purification of D5-1 (left) and D5-2 (right). Both gels show the load material (L), flow through (FT), beads before cleavage (BB), beads after cleavage (BA) and the cleaved product (C). The first lane is the protein marker where useful sizes in kDa have been marked on the side. GST-tagged D5-1 and D5-2 can be seen running at 35 and 33 kDa respectively, and the cleaved products can be seen running at approximately 17 and 10 kDa in lanes C. The cleaved GST tag is present in lanes BA, running at 28 kDa.*

After the GST column, a cationic exchange column was used to fully purify both D5-1 and D5-2 (Figures 5.4 and 5.5 respectively). Due to having a pI of 9.4, D5-2 bound strongly to the column and eluted at 55-60 mS/cm. D5-1 has a pI of 7.5 and, therefore, did not bind as strongly as D5-2 and eluted at 30-35 mS/cm. Both proteins ran as clean single bands on the gel with no impurities.

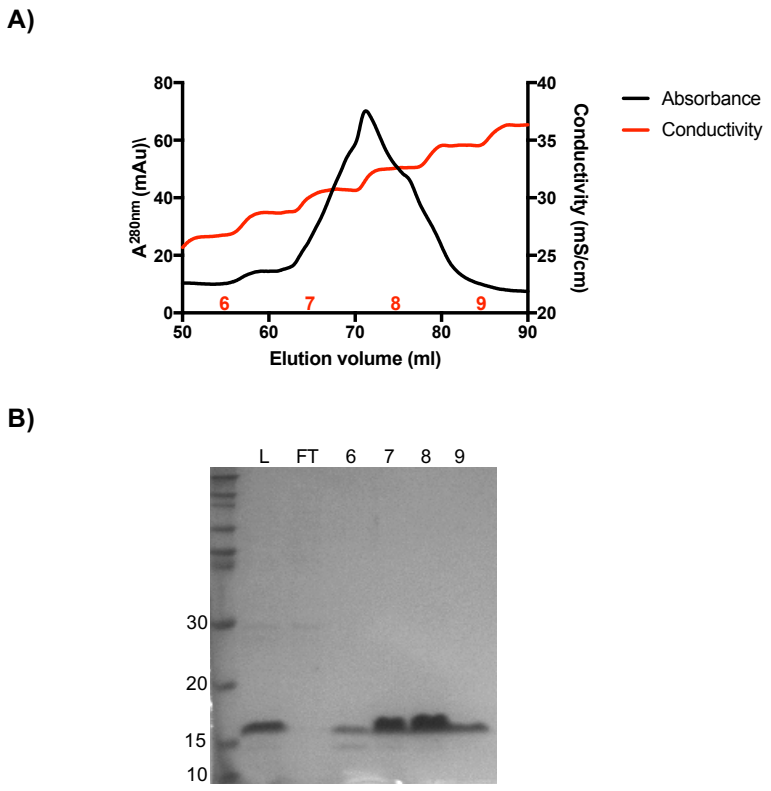


Figure 5.4: *Cationic exchange column of cleaved D5-1.* ***A)*** *Absorbance at 280 nm (black line) and salt concentration (red line) plotted against elution volume. 10 ml fractions were collected and fraction numbers are shown in red.* ***B)*** *SDS-PAGE gel showing the protein content of the load material (L), flow through (FT) and fractions 6-9 from the cationic column. The first lane is the protein marker where useful sizes in kDa have been marked on the side. Pure D5-1 can be seen in lane 9 running at approximately 17 kDa.*

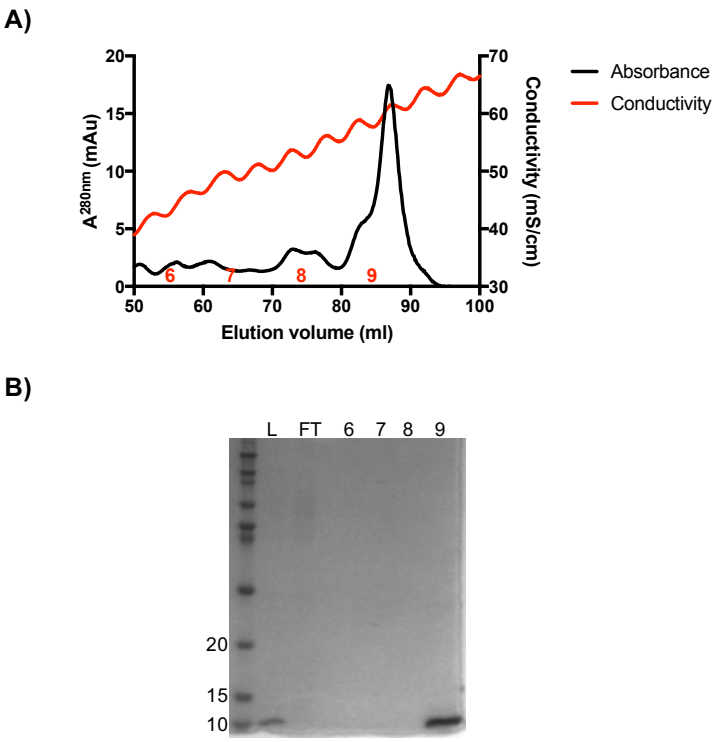


Figure 5.5: *Cationic exchange column of cleaved D5-2.* ***A)*** *Absorbance at 280 nm (black line) and salt concentration (red line) plotted against elution volume. 10 ml fractions were collected and fraction numbers are shown in red.* ***B)*** *SDS-PAGE gel showing the protein content of the load material (L), flow through (FT) and fractions 6-9 from the cationic column. The first lane is the protein marker where useful sizes in kDa have been marked on the side. Pure D5-2 can be seen in lane 9 running at 10 kDa.*

5.2.2 Characterisation of D5-1 and D5-2

5.2.2.1 Secondary Structure

Both D5-1 and D5-2 produced comparable CD spectra to full length D5 representative of largely disordered structures (Figure 5.6). This confirmed that splitting D5 into two fragments had not resulted in protein aggregation; therefore, all further results would be due to the absence of either the N or C terminus and not due to aggregation of the truncations.

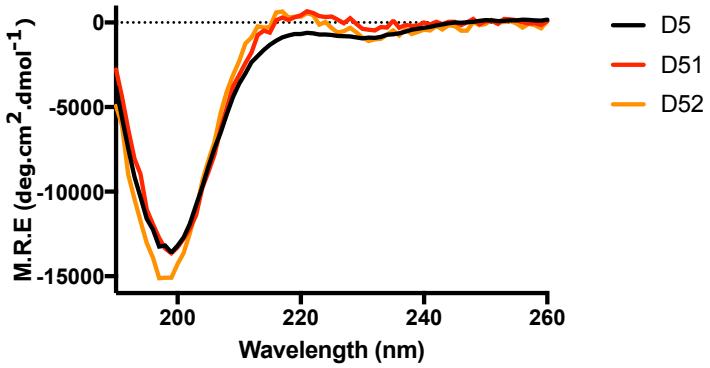


Figure 5.6: *Overlaid far UV CD spectra of D5, D5-1 and D5-2 (15 µM each) at 298 K. Both D5 truncations are consistent with full length D5 and have a largely disordered structure.*

5.2.2.2 Zinc Binding Properties

Similar zinc binding experiments were performed on D5-1 and D5-2 as for full length D5. Mass spectrometry data for both D5 truncations in the presence of zinc is shown in Figures 5.7 and 5.8 respectively. The mass spectrum of D5-1 in the presence of zinc identified three species: unbound D5-1, D5-1 bound to one Zn^{2+} and D5-1 bound to two Zn^{2+}. The same experiment performed on D5-2 identified two species: D5-2 bound to one Zn^{2+} and D5-2 bound to two Zn^{2+}. Therefore, two metal ions bound through the N-terminal region and two through the C-terminal region, and zinc binding is not confined to a particular region of D5 but is spread throughout the sequence.

	D5-1 species A	**D5-1 species B**	**D5-1 species C**	**D5-2 species A**	**D5-1 species B**
Molecular mass (Da)	8409.65 (±0.31)	8472.55 (±0.43)	8535.16 (±2.20)	6689.71 (±0.44)	6752.60 (±0.75)
Bound Zn^{2+} ions	None	One	Two	One	Two

Table 5.2: *Calculated masses from the ESI-FT mass spectrum of D5-1 and D5-2 in the presence of zinc.*

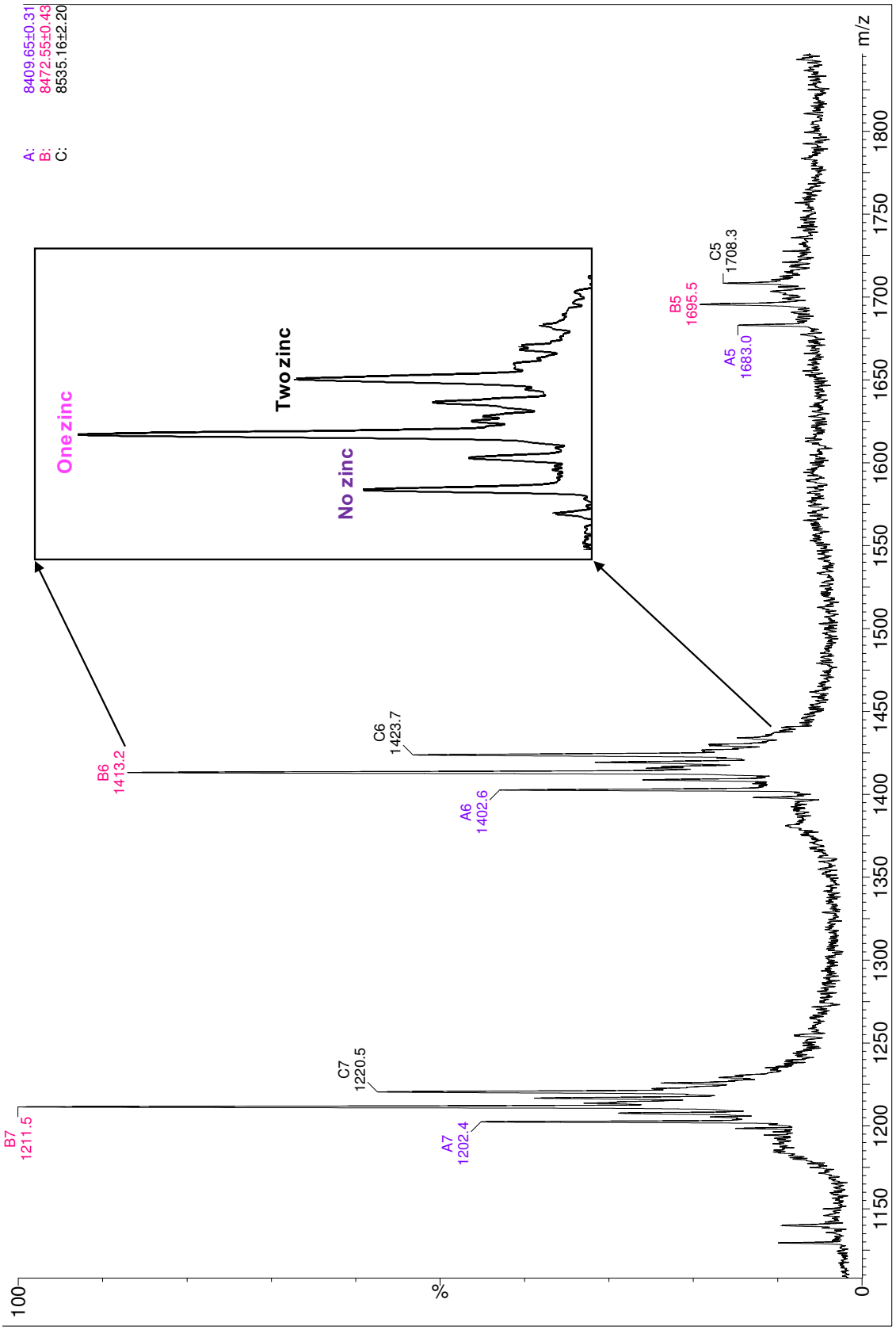


Figure 5.7: *Native ESI-FT mass spectrum of D5-1 (20 µM) in the presence of 100 µM zinc chloride. Three different species can been seen in this spectrum corresponding to unbound D5-1 (species A), D5-1 bound to one (species B)and two zinc ions (species C). The charge state and M/Z ratios are indicated above the peak, and the calculated mass for each species is shown on the right. A zoomed in image showing one set of peaks corresponding to the three species is also shown.*

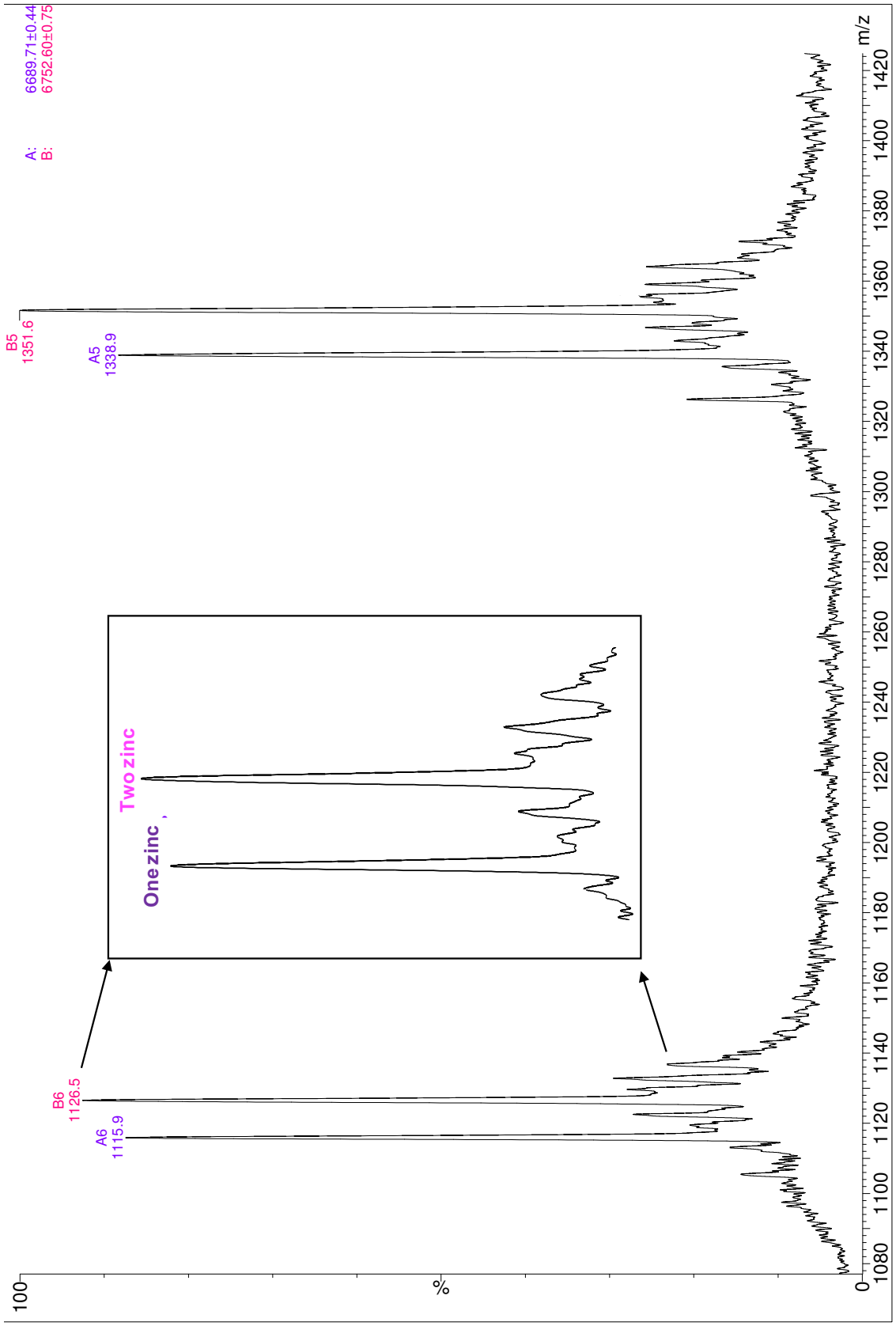


Figure 5.8: *Native ESI-FT mass spectrum of D5-2 (20 µM) in the presence of 100 µM zinc chloride. Two different species can been seen in this spectrum corresponding to D5-2 bound to one (species A) and two (species B) zinc ions. The charge state and M/Z ratios are indicated above the peak and the calculated mass for each species is shown on the right. A zoomed in image showing one set of peaks corresponding to the two species is also shown.*

ITC further confirmed the binding of two zinc atoms to both D5-1 and D5-2; this is consistent with the ITC data for full length D5 that identified the binding of four zinc atoms. Once again, these zinc binding ITC experiments were performed in 50 mM HEPES buffer instead of potassium phosphate to prevent zinc precipitation. The metal binding affinities for D5-1 and D5-2 were comparable to that of full length D5 (Table 5.3) suggesting that both these metal binding regions are independent of one another and do not show any cooperative binding properties.

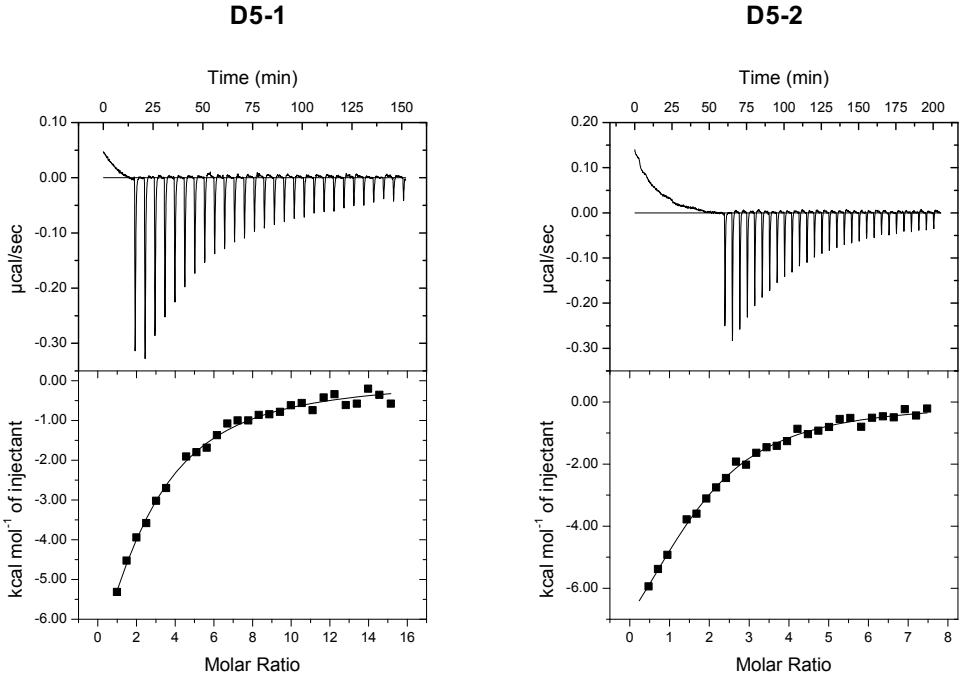


Figure 5.9: *ITC data for the titration of zinc chloride into D5-1 or D5-2 (10 µM). The top panel is the raw ITC data showing the exothermic release of heat upon zinc binding, and the bottom panel is the isotherm produced when you integrate the peaks.*

	N	KD (µM)	ΔH (kJ mol^{-1})	ΔS (J K^{-1} M^{-1})	ΔG (kJ mol^{-1})
D5	4.0 (± 0.4)	1.5 (± 0.2)	-41.8 (± 5.9)	-28.5	-33.3
D5-1	2.0 (± 0.39)	8.5 (± 1.1)	-68.0 (± 15.4)	-130.6	-29.1
D5-2	1.7 (± 0.1)	4.7 (± 0.5)	-44.9 (± 3.4)	-48.6	-34.1

Table 5.3: *ITC-derived thermodynamic properties for the binding of zinc with D5-1 and D5-2 compared with full length D5.*

Each D5 fragment contains 15 and and 12 His residues respectively; however, specific Zn^{2+} binding stoichiometries were observed with both D5-1 and D5-2 by mass spectrometry and ITC which suggests specific co-ordination of Zn^{2+}. Similar to full length D5, CD did not detect any change in secondary structure upon the binding of zinc to both the D5-1 and D5-2 fragments (Figure 5.10), and subsequent addition of EDTA had no effect on the CD spectra

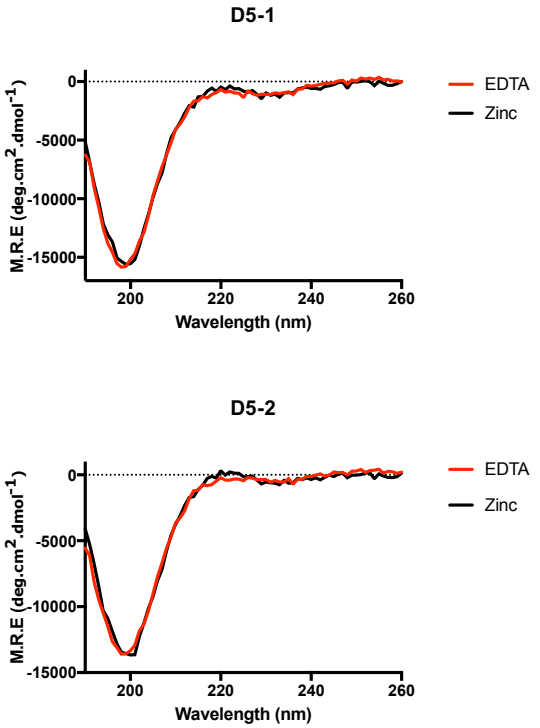


Figure 5.10: *Far UV CD spectra of D5-1 and D5-2 (15 µM each) in the presence of 1 mM EDTA or 75 µM $ZnCl_2$.*

5.2.3 D5 Truncations and Their Binding to gC1q-R

5.2.3.1 Analytical Gel Filtration

Similarly to full length D5, the binding of both D5-1 and D5-2 with gC1q-R was initially tested using analytical gel filtration (Figure 5.11). Comigration of D5-1

with gC1q-R was only observed in the presence of zinc; this was the first sign of any metal dependent binding whilst working on D5 in these studies. In the presence of EDTA, D5-1 eluted in the low molecular weight peak and no gC1q-R bound ligand was detected; however, in the presence of zinc, D5-1 comigrated with gC1-R in the high molecular weight peak. In contrast, D5-2 comigrated with gC1q-R in both the presence and absence of zinc. D5, therefore, binds to gC1q-R through two distinct regions: the N-terminus in a zinc-dependent interaction and the C-terminus in a zinc-independent interaction.

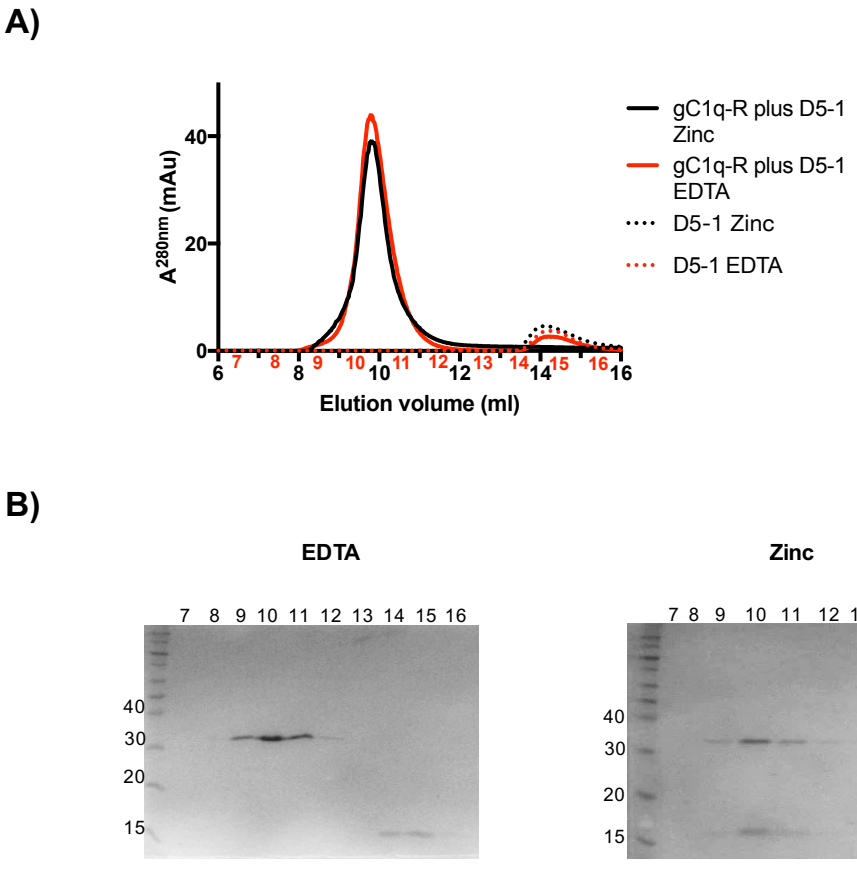


Figure 5.11: *Analytical gel filtration of D5-1 (10 µM) and gC1q-R (10 µM trimer).* ***A)*** *Elution profiles in the presence of 50 µM zinc chloride (black) or 5 mM EDTA (red). Fraction numbers are indicated in red. 10 µM D5-1 only controls in the presence of zinc and EDTA are also shown as dotted lines.* ***B)*** *SDS-PAGE gels of the 1 ml fractions collected showing the protein content of each peak. The first lane is the protein marker where useful sizes in kDa have been marked on the side. gC1q-R can be seen at approximately 32 kDa and D5-1 at approximately 17 kDa. D5-1 comigrated with gC1q-R only in the presence of zinc outlining the zinc-dependent nature of the interaction.*

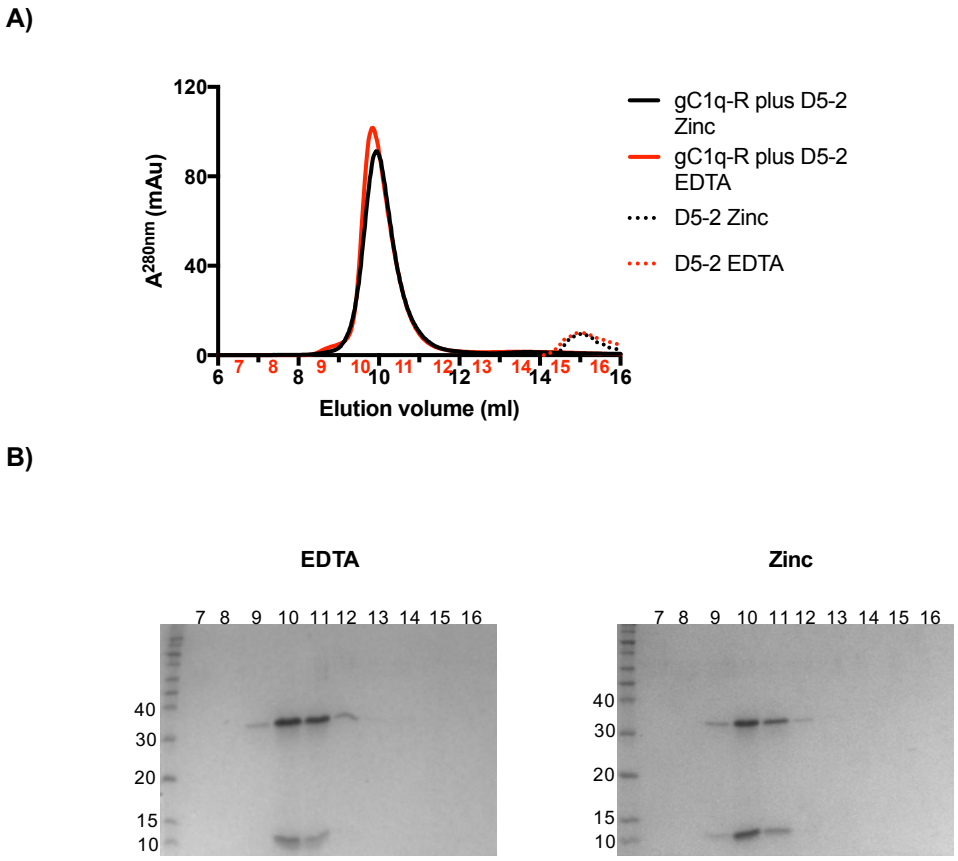


Figure 5.12: *Analytical gel filtration of D5-2 (20 µM) and gC1q-R (20 µM) trimer).* ***A)*** *Elution profiles in the presence of 50 µM zinc chloride (black) or 5 mM EDTA (red). Fraction numbers are indicated in red. 20 µM D5-2 only controls in the presence of zinc and EDTA are also shown as dotted lines.* ***B)*** *SDS-PAGE gels of the 1 ml fractions collected showing the protein content of each peak. The first lane is the protein marker where useful sizes in kDa have been marked on the side. gC1q-R can be seen at approximately 32 kDa and D5-2 at 10 kDa. D5-2 comigrated with gC1q-R in both the presence and absence of zinc outlining the zinc-independent nature of the interaction.*

5.2.3.2 Two Distinct Binding Mechanisms Detected by ITC

ITC was used to calculate the thermodynamic properties for the interactions of the D5 truncations with gC1q-R. The ITC binding curves generated from the addition of either D5-1 or D5-2 with gC1q-R are not comparable (Figure 5.13) and clearly represent two different binding mechanisms.

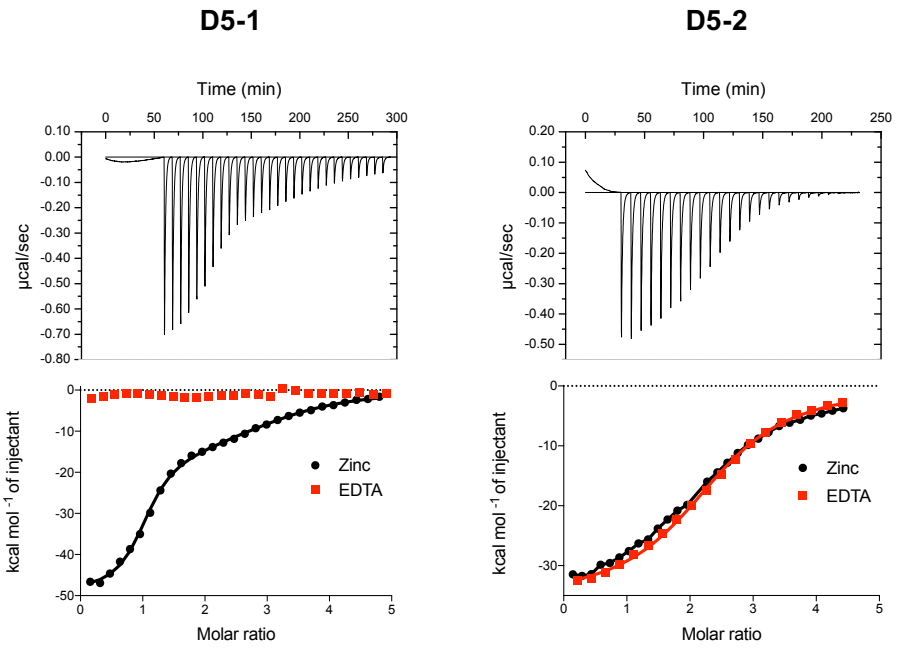


Figure 5.13: *ITC data when D5-1 or D5-2 (100 µM each) were titrated into gC1q-R (3 µM trimer). The top panels show the raw ITC data in the presence of zinc, and the bottom panels show the binding curves produced after saturation with zinc (black) or in the presence of 1mM EDTA (red). D5-1 binding is zinc-dependent and shows negative cooperativity, whereas D5-2 binding is zinc-independent and shows no sign of cooperativity.*

The curve generated for the binding of D5-1 with gC1q-R in the presence of zinc was fit to the same three site sequential binding model as full length D5. The derived thermodynamic properties for each binding event are shown in Table 5.4. The first and second binding events are approximately 40 and 20 fold weaker than full length D5 respectively, whereas the binding affinity of the third binding event is more comparable with the full length domain. The relationship between ΔH and ΔS is shown in Figure 5.14 and, similarly to full length D5, the interaction is enthalpy driven. Unlike full length D5 however, the third binding event is not associated with a favourable increase in entropy, and this could be due to the decreased size of the ligand, resulting in reduced hydrophobic interactions.

	K_D (nM)	ΔH (kJ mol^{-1})	ΔS (J K^{-1} M^{-1})	ΔG (kJ mol^{-1})
First binding event	73.0 (± 4.0)	-200.9 (± 1.4)	-535.9	-41.2
Second binding event	1579.8 (± 145.3)	-92.2 (± 4.9)	-198.5	-33.0
Third binding event	2673.8 (±257.7)	-62.2 (± 5.9)	-101.7	-31.9

Table 5.4: *ITC-derived thermodynamic properties for the sequential binding of D5-1 with gC1q-R.*

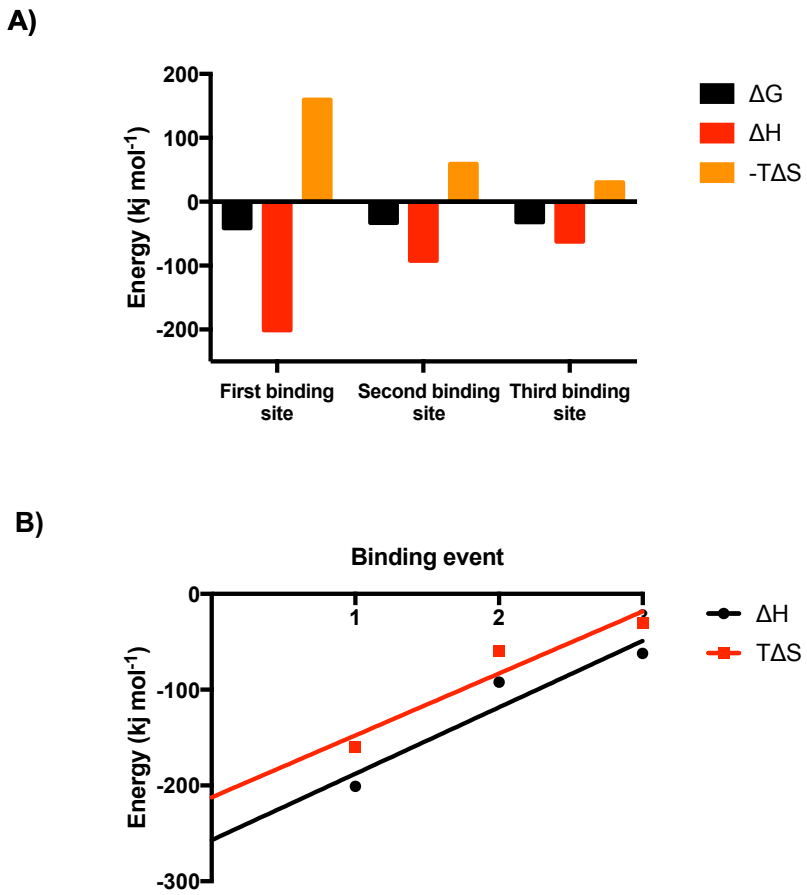


Figure 5.14: *Relationship between ΔH and ΔS for the three binding events of D5-1 with gC1q-R.* ***A)*** *Chart showing the ΔG, ΔH and -TΔS values for each binding event. Each binding events consist of a favourable negative ΔH and unfavourable positive -TΔS showing that these binding events are enthalpy driven, characteristic of electrostatic interactions and favourable hydrogen bond formation.* ***B)*** *Ethalpy-entropy compensation plot where ΔH and TΔS values are plotted for each binding event. This plot highlights the proportional relationship between enthalpy and entropy as more D5 ligands bind gC1q-R. ΔG is consistent between binding events and is equal to ΔH-TΔS.*

N	K_D (nM)	ΔH (kJ mol^{-1})	ΔS (J K^{-1} M^{-1})	ΔG (kJ mol^{-1})
2.3 (± 0.01)	763.4 (± 27.5)	-149.1 (± 1.1)	-382.7	-35.1

Table 5.5: *ITC-derived thermodynamic properties for the binding of D5-2 with gC1q-R.*

The binding curve produced upon the binding of D5-2 with gC1q-R was fit to a more classical single site binding model. The calculated thermodynamic parameters are shown in Table 5.5. The N value was calculated to be 2.3 indicating that there are multiple, equivalent binding sites as apposed to the inequivalent binding sites of D5 and D5-1. Each binding site had a calculated K_D of 763 nM. An N value closer to 3 would be more consistent with the model of one D5-2 binding to each monomer. An N value of 2.3 corresponds to 0.8 D5-2 ligands per monomer of gC1q-R, and a 20 % concentration error at each site would result in the lower than predicted N value of 2.3. On the other hand, having only two D5-2 domains binding to the receptor could explain why D5 has a tighter binding affinity during the first two binding events when compared with D5-1, whereas the third binding events are more comparable; the first two D5 binding events are a combination of interactions through both D5-1 and D5-2, whereas the third binding event only involves D5-1. A third D5-2 might be prevented from effectively binding the remaining available site due to steric clashes with the other D5-2 ligands or allosteric modification of the receptor. D5-2 binding is associated with a large, negative ΔH and a negative ΔS (Figure 5.15) which is consistent with both D5 and D5-1, and suggests a largely electrostatic interaction. This is not surprising as D5-2 is a highly basic region and can make extensive electrostatic contact with the acidic patches of gC1q-R.

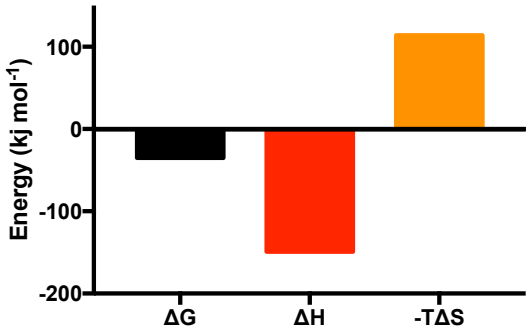


Figure 5.15: A) *Chart showing the ΔG, ΔH and -TΔS values for the binding of D5-2 with gC1q-R. The binding consist of a favourable negative ΔH and unfavourable positive -TΔS showing that these binding events are enthalpy driven, characteristic of electrostatic interactions and favourable hydrogen bond formation.*

5.2.3.3 D5 Truncation Competition Experiments

Both the gel filtration and ITC data clearly show that both the N and C-termini of D5 hold key binding sites for gC1q-R. One important question that needed to be addressed was if these two binding regions bound to separate or overlapping sites on the receptor. One way to determine this is by performing competition experiments where both ligands will either compete with one another or bind simultaneously. Competitive binding between D5-1 and D5-2 would be suggestive of a single binding site, whereas simultaneous binding would suggest that they have different binding sites. The fact that both these ligands show different binding mechanisms by ITC is more suggestive of there being two different sites

Analytical gel filtration was the first method used to confirm the presence of one or two binding sites on gC1q-R. If both proteins bound to separate sites then all three proteins would elute from the column as one bound species, whereas if there was only one site gC1q-R would only bind one ligand and the other would elute separately as an unbound species. In the presence of zinc, all three

proteins coeluted in the high molecular weight peak suggesting the formation of a ternary complex (Figure 5.16). D5-1 was added in excess so each D5-1 binding site on gC1q-R would have been saturated; therefore, the additional binding of D5-2 must be due to binding within a different region of gC1q-R. In the absence of zinc, only D5-2 co-migrated with the receptor, and unbound D5-1 could be seen eluting as a low molecular weight peak. This result was expected and further confirmed that, even when D5-2 is bound to the receptor, the D5-1 interaction is still strictly zinc-dependent.

ITC was used to further confirm the presence of two separate binding sites for D5-1 and D5-2. After full saturation of gC1q-R with either D5-1 or D5-2 the opposing ligand was titrated into the complex, and the binding curves were compared to those produced with unbound gC1q-R. If both regions competed for the same site there would be a significant reduction in binding affinity as both ligands competed with one another for the gC1q-R binding site. On the other hand, no change in binding affinity would represent two independent binding sites. The binding curve produced upon the titration of D5-1 with gC1q-R saturated with D5-2 remained virtually unchanged when compared with binding to gC1q-R alone, and the same applied for the titration of D5-2 with gC1q-R saturated with D5-1 (Figure 5.17). This provided further evidence that both D5-1 and D5-2 bind separate sites, and the binding of one of these regions to gC1q-R does not have any significant affect on the binding of the other.

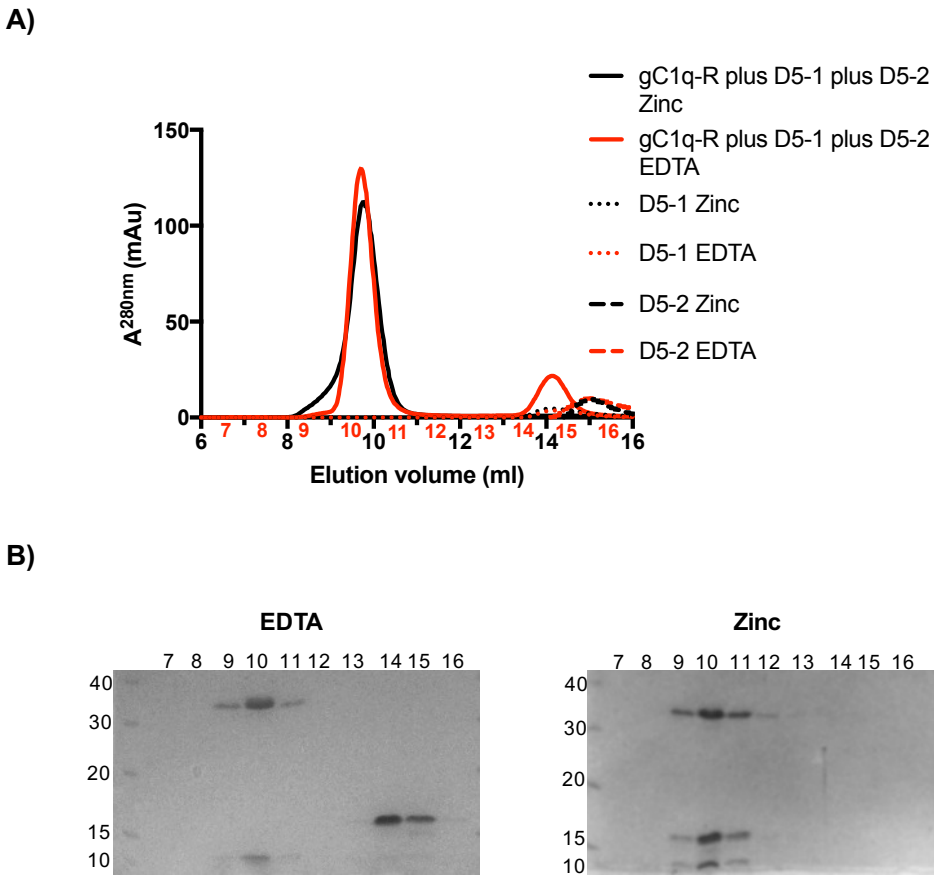


Figure 5.16: *Analytical gel filtration competition experiments of D5-1 (30 μM), D5-2 (10 μM) and gC1q-R (10 μM trimer).* ***A)*** *Elution profiles in the presence of 50 μM zinc chloride (black) or 5 mM EDTA (red) with fraction numbers indicated in red. 10 μM D5-1 only and 20 μM D5-2 only controls in the presence of zinc and EDTA are also shown.* ***B)*** *SDS-PAGE gels of the 1 ml fractions collected showing the protein content of each peak. The first lane is the protein marker where useful sizes in kDa have been marked on the side. gC1q-R can be seen at approximately 32 kDa, D5-1 at 17 kDa and D5-2 at 10 kDa. A ternary complex is formed in the presence of zinc signified by the co-migration of all three proteins.*

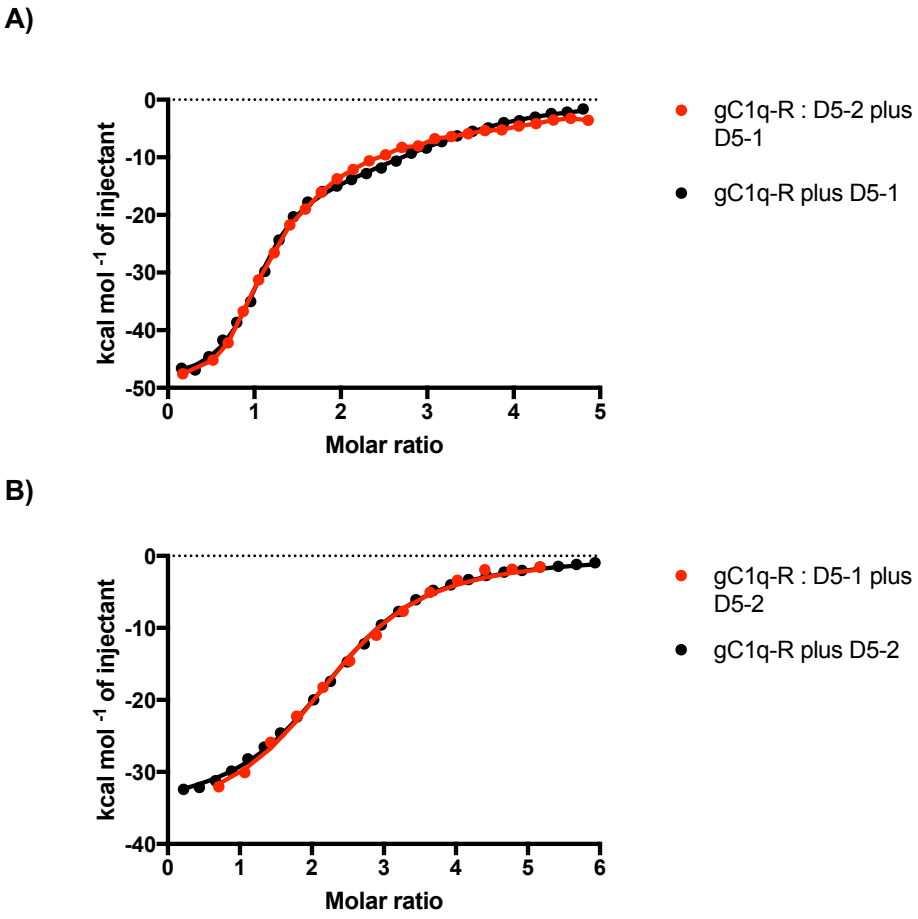


Figure 5.17: *ITC experiment showcasing the dual binding sites within gC1q-R for D5-1 and D5-2. These titrations were performed after saturation of both ligand and receptor with Zn^{2+}.* ***A)*** *ITC binding curve produced when D5-1 (100 μM) was titrated into the complex of gC1q-R and D5-2 (3 μM) compared with the binding curve of D5-1 with unbound gC1q-R.* ***B)*** *ITC binding curve produced when D5-2 (100 μM) was titrated into the complex of gC1q-R and D5-1 (3 μM) compared with the binding curve of D5-2 with unbound gC1q-R.*

	K_D1 (nM)	K_D2 (nM)	K_D3 (nM)
gC1q-R + D5-1	73.0 (± 4.0)	1579.8 (± 145.3)	2673.8 (± 2673.8)
gC1q-R : D5-2 + D5-1	81.3 (± 6.0)	254.5 (± 26.5)	862.1 (± 104.3)
gC1q-R + D5-2	763.4 (± 27.5)	-	-
gC1q-R : D5-1 + D5-2	471.7 (± 58.5)	-	-

Table 5.6: *ITC-derived binding affinities for the interaction of D5-1 or D5-2 with the gC1q-R : D5-2 or gC1q-R : D5-1 complexes respectively. The calculated binding affinities for these ligands with unbound gC1q-R is also presented as a comparison. The K_D values are comparable between bound and unbound gC1q-R which is suggestive of two distinct binding sites for D5-1 and D5-2 on gC1q-R.*

5.2.4 Peptide Binding Detection by ITC

The binding of the three His-Gly-rich D5 peptides with gC1q-R was tested using ITC in order to narrow down the D5-1 and D5-2 binding sites and further pinpoint the key gC1q-R binding regions. The peptide sequences and their positions within D5 can be found in Figure 5.2. Both N-terminal His-Gly-rich peptides showed no binding to the receptor after being dialysed against ITC Zinc Buffer, the same buffer that was used to detect the interaction of D5-1 with gC1q-R; whereas, the Lys-rich HMWK 493-516 peptide produced a comparable binding curve to full length D5-2 (Figure 5.18) in the absence of zinc, yielding similar thermodynamic parameters (Table 5.7). This provided evidence that the gC1q-R binding site of D5-2 is located within this Lys-rich region.

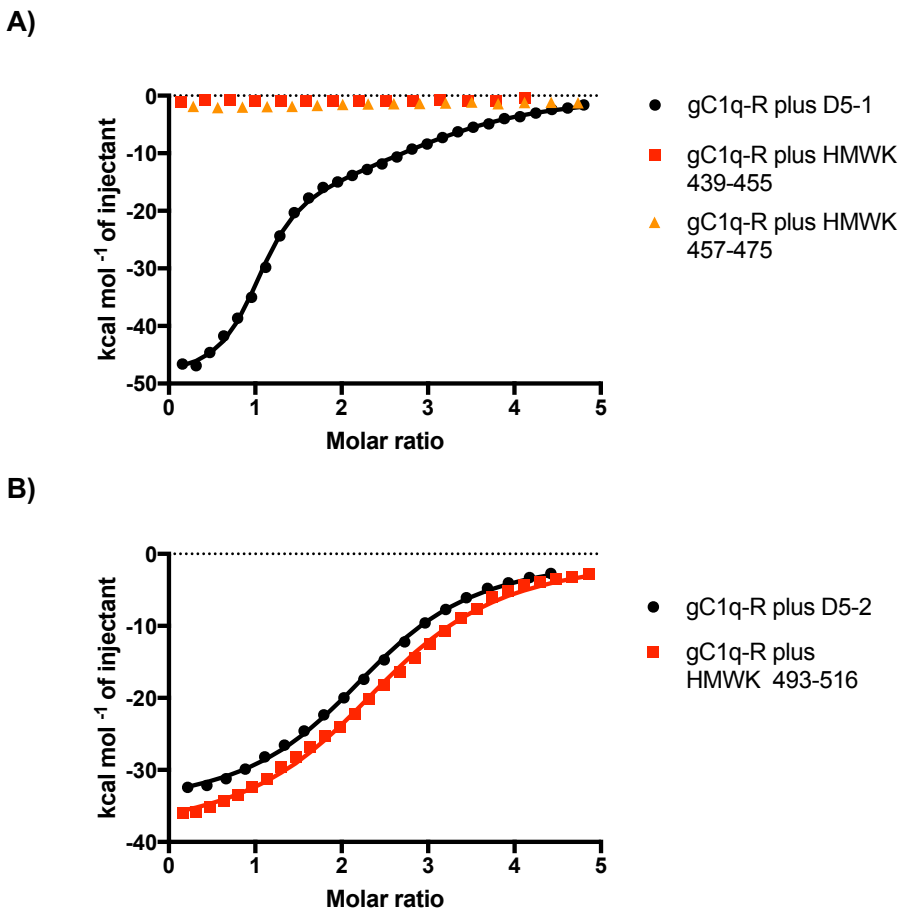


Figure 5.18: *ITC binding curves for the three His-Gly-rich D5 peptides (200 μM) were titrated into gC1q-R (10 μM).* ***A)*** *Binding curve for the titration of HMWK 439-455 or HMWK 457-475 with gC1q-R compared with the D5-1 binding curve. ITC titrations were performed after saturation of both ligand and receptor with Zn^{2+}. No binding curves were generated by the N-terminal peptides.* ***B)*** *Binding curve for the titration of HMWK 493-516 with gC1q-R compared with the D5-2 binding curve. Both binding curves are comparable suggesting that HMWK 493-516 is the major binding region of D5-2.*

N	K_D (nM)	ΔH (kJ mol^{-1})	ΔS (J K^{-1} M^{-1})	ΔG (kJ mol^{-1})
2.56 (±0.02)	1612.9 (± 90.3)	-164.0 (± 1.8)	-439.6	-33.2

Table 5.7: *ITC-derived thermodynamic properties for the binding of HMWK 493-516 with gC1q-R.*

5.2.5 HMWK 401-438 N-terminal Truncation

Both HMWK 439-455 and HMWK 457-475 showed no detectable signs of binding by ITC. These two peptides span the C-terminal portion of D5-1 which suggested that it is potentially the N-terminal region that is responsible for gC1q-R binding. To test this hypothesis, one final D5 truncation was produced, HMWK 401-438, consisting of residues 401-438 at the very N-terminus of D5. This mutant was produced by cloning a premature stop codon within the D5-1 pGEX plasmid.

5.2.5.1 Purification of HMWK 401-438

The purification of HMWK 401-438 was similar to that of D5-1 and D5-2. The only difference was, due to the lower theoretical pI of 6.84, the pH of the cationic exchange buffers was reduced to 6.0 in order to increase the binding efficiency of this truncation to the column. The GST and cationic exchange purification steps of HMWK 401-438 are shown in Figures 5.19 and 5.20 respectively.

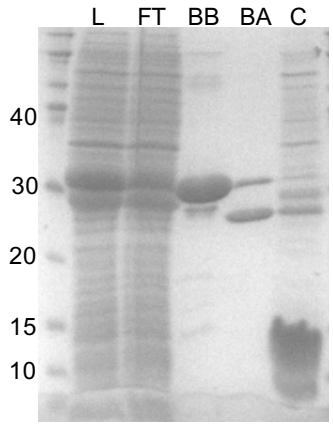
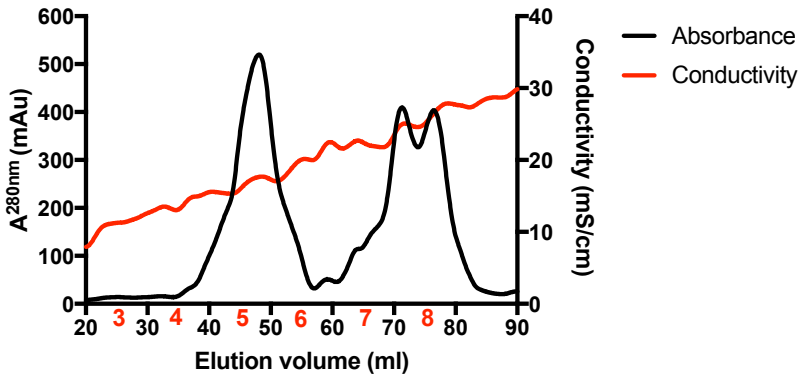


Figure 5.19: *SDS-PAGE gel of fractions taken from the glutathione sepharose purification of HWMK 401-438. The gel shows the load material (L), flow through (FT), beads before cleavage (BB), beads after cleavage (BA) and the cleaved product (C). The first lane is the protein marker where useful sizes in kDa have been marked on the side. GST-tagged HMWK 401-438 can be seen running at 32 kDa, and the cleaved product can be seen running at approximately 12 kDa in lane C. The cleaved GST tag is present in lane BA running at 28 kDa.*

A)



B)

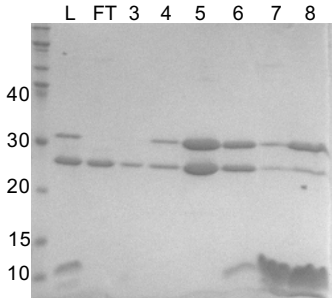


Figure 5.20: *Cationic exchange column of cleaved HMWK 401-438.* ***A)*** *Absorbance at 280 nm (black line) and salt concentration (red line) plotted against elution volume. 10 ml fractions were collected and fraction numbers are shown in red.* ***B)*** *SDS-PAGE gel showing the protein content of the load material (L), flow through (FT) and fractions 3-8 from the cationic column. The first lane is the protein marker where useful sizes in kDa have been marked on the side. Pure cleaved HWMK 401-438 can be seen in lanes 6-8 running at 10 kDa. The two contaminants running at approximately 32 and 27 kDa are un-cleaved HMWK 401-438 and cleaved GST respectively. Protein from fraction 7 was the only material used for experiments as it was the purest sample.*

5.2.5.2 HMWK 401-438 Binding to gC1q-R by ITC.

HMWK 401-438 showed no detectable binding to gC1q-R by ITC after being dialysed against ITC Zinc Buffer. Therefore, after testing three peptides from N terminal region of D5, the key gC1q-R binding regions within D5-1 remain unclear. The binding site for D5-1 is likely spread over a larger region that requires multiple binding motifs to effectively bind.

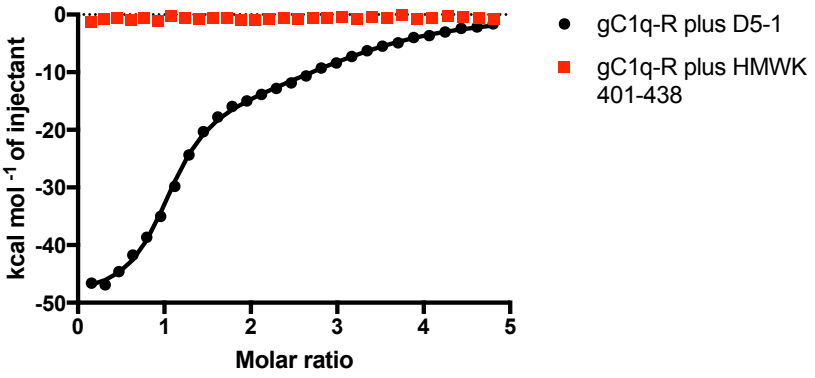


Figure 5.21: *ITC binding curve produced when HMWK 401-438 (200 μM) was titrated into gC1q-R (10 μM) compared with the D5-1 binding curve. ITC titrations were performed after saturation of both ligand and receptor with Zn^{2+}. No binding curve was generated by HMWK 401-438.*

5.3 Discussion

Both the N and C-termini of D5 bind to different regions of gC1q-R and are able to bind independently of one another. One of the major differences between the two regions of D5 is that, unlike the C terminus, the N terminus binds in a zinc-dependent manner. This is consistent with previous studies that showed peptides from the N-terminal region of D5 also bind heparin in a zinc dependent manner, whereas peptides from the C-terminal Lys-rich region bound in the absence of zinc [254]. This study linked the increased Lys content of the C-

terminal peptide with its ability to bind surfaces in a zinc-independent manner; therefore, the Lys residues within D5-2 are likely making key contacts with the receptor in the absence or presence of zinc. It remains unclear as to why full length D5 showed no zinc dependence when binding gC1q-R seeing as isolated D5-1 binds zinc-dependently. One explanation could be that the binding of the Lys-rich C-terminal region of D5 orientates the N-terminal region in such a way that zinc is no longer required for binding; therefore, full length D5 is still able to bind gC1q-R through both binding regions in the absence of zinc. This does not appear to be the case for full length HMWK, which requires zinc for binding, and could suggest that the Lys-rich region within HMWK does not play as big a role in binding compared with isolated D5.

The gC1q-R binding site within D5-2 was pinpointed to the HMWK 493-516 peptide. This binding region is made up of 33 % Lys content and contains 57 % of the total Lys content for the entire D5 sequence. This provided further evidence which suggests the Lys residues play a pivotal role in gC1q-R binding, and that this binding region is localised to a 24 aa sequence near the C-terminal end of D5. What remains unknown is if all the Lys residues are interacting with the receptor or if there are only a few key binding residues. There are a total of 49 acidic residues within one gC1q-R monomer; this number is sufficient to allow all 8 Lys residues within HMWK 493-516 to make key electrostatic interactions with the receptor. However, in order to probe the key binding residues even further, specific residues within this region would have to be mutated in order to locate individual binding residues.

The binding region within D5-1, on the other hand, could not be narrowed down through the use of shorter peptides suggesting the N-terminal binding site is more diffuse. The peptides used spanned the entire D5-1 sequence, excluding one Gln residue which lies between HMWK 439-455 and HMWK 457-475, but no binding was detected for any of the peptides. The most likely expla-

nation for the lack of binding is that more than one of these peptide regions is required for gC1q-R binding. Larger peptide fragments from D5-1 will need to be tested in order to test this hypothesis and locate key binding regions.

The sequential binding mechanism observed within full length D5 is largely due to the binding of the N-terminal D5-1 fragment which also binds in a sequential manner as shown by ITC. It is possible that the negative cooperativity shown by D5-1 is linked with its more diffuse binding region compared with the shorter binding motif found in D5-2. The more interactions being made between the ligand and receptor the higher the protein density within the binding site, resulting in increased steric hinderance for subsequent ligands. Furthermore, if D5-1 is binding through the central cavity of gC1q-R, where the distance between equivalent sites on the other monomers is at its shortest, then steric clashes will be amplified. The fact D5-2 does not show the same sequential binding mechanism could suggest it is binding towards the outside of the donut structure where steric clashes won't have as much of an effect on the binding of other D5-2 ligands. Combined with a smaller binding patch, this would remove any sequential binding that is present with full length D5 or D5-1.

A model for the duel binding of D5-1 and D5-2 is shown in Figure 5.22. In this figure both D5-1 and D5-2 are shown to be binding different sites on the same monomer of gC1q-R. However, it is also possible that one D5 could simultaneously bind two different monomers. A gC1q-R mutant which prevents the trimer formation of gC1q-R could be used to determine if both D5 regions are able to simultaneously bind to the one monomer. Future experiments could be performed where D5-1 and D5-2 are used in competition experiments for full length HMWK binding with gC1q-R; this would help confirm if both these regions are equally critical for the binding of full length HWMK.

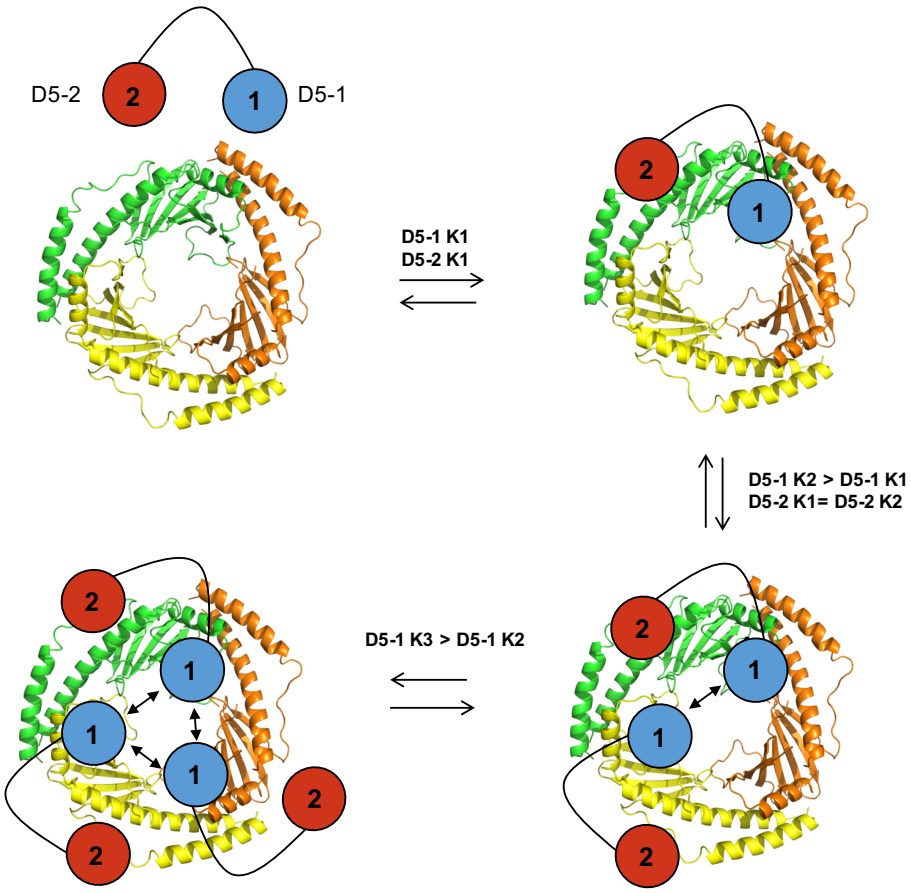


Figure 5.22: *A cartoon representation for the proposed model of D5 binding to gC1q-R. D5 binds through both the N-terminal D5-1 and C-terminal D5-2 regions through different sites on gC1q-R. The binding of D5-1 exhibits negative cooperativity, and the binding of the first D5-1 is followed by the binding of a second and third, each with reduced binding affinities. The binding of D5-2 shows no sign of cooperativity, and all the binding events have equivalent binding affinities.*

It has long been believed that D5 is able to bind to surfaces through multiple regions [84] [70], and the results in this chapter not only support this but also provide evidence that both regions bind two independent sites on gC1q-R. The ITC data suggests full length D5 binds approximately 40 fold and 400 fold stronger than D5-1 and D5-2 respectively, indicating avidity effects in the binding of both these fragments when part of the same peptide chain. The significance of this dual binding, although not fully understood, could be crucial for the localisation of HMWK, FXI and PK to the receptor surface. D5 acts as a flexible arm that is able to extend outwards and bind gC1q-R through both the N-terminal or C-terminal ends. The simultaneous binding of the N and C-termini would lock this flexible domain, and therefore the entire HMWK, into a specific orientation which allows FXII to activate PK and FXI.

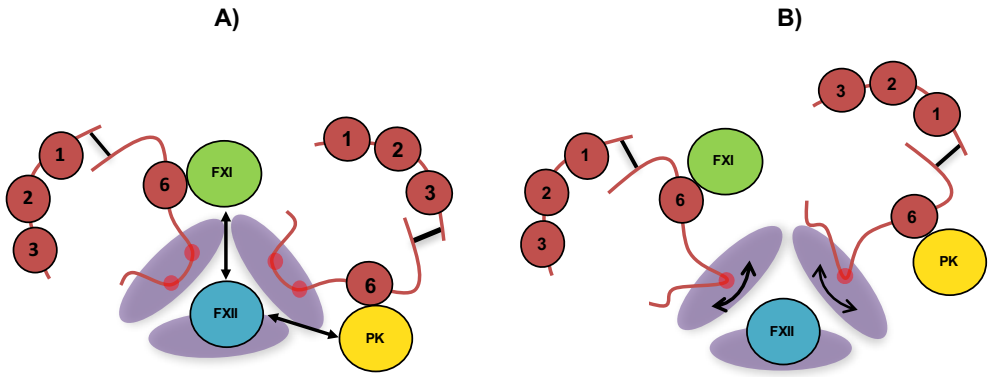


Figure 5.23: *Cartoon representation for the activation of contact initiation and how the multiple D5 binding sites have an important role in presenting the other contact initiators at the cell surface. Activated HMWK is used in this cartoon due to its increased coagulant activity over HMWK, but this model would also apply to un-cleaved HMWK.* ***A)*** *The binding of activated HMWK (red) to gC1q-R (purple) through two different regions of D5. This locks the complex into a particular conformation which allows FXII, which is bound to another monomer of gC1q-R, to activated HMWK bound PK and FXI.* ***B)*** *The alternative binding of D5 through one region allows more rotation and flexibility within HMWK and would not fix PK and FXI in close proximity with FXII, reducing contact activation.*

CHAPTER 6

Pinpointing the D5 Binding Sites on gC1q-R Using Site Directed Mutagenesis

6.1 Introduction

The production of D5 truncations proved extremely useful in pinpointing key gC1q-R binding regions within the D5 sequence and further understanding the interaction of D5 with gC1q-R. This chapter focusses on point mutations made within gC1q-R that were used to narrow down the D5 binding patch on the receptor. Testing the binding of gC1q-R mutants with full length D5 and the D5 truncations has resulted in an increased understanding regarding the nature of this interaction and has as helped narrow down the key D5 binding regions on the receptor. These data will potentially aid in the future development of small molecule inhibitors targeting contact activation through the interaction of HMWK and gC1q-R.

6.1.0.3 HMWK Interaction Surface on gC1q-R

A past study, looking at a variety of gC1q-R mutants and their binding to HMWK using ELISA, narrowed the interaction patch to residues 190-218 of gC1q-R [155]. This region spans the β5-β6 flexible loop at the top of the trimer, along the β6 strand and to the β6-β7 loop situated at the bottom of the receptor. Furthermore, the deletion of a small acidic patch of residues within the β3-β4 loop, consisting of residues 144-148, also reduced binding with HMWK. There are regions within this proposed binding site that are within flexible loop regions not resolved in the original crystal structure, and in order to easily visualise these missing loops a model of gC1q-R was built using SWISS-MODEL. This model was built using the crystal structure for wild type gC1q-R as a template (PDB ID:1P32), where the missing loop regions were simply modelled into the pre-existing structure. This allowed for easier representation of the HMWK binding patch without modifying the solved crystal structure (Figure 6.1).

In addition to the binding of HMWK, residues 144-148 and 196-202 were reported to also be involved in FXII binding [155]. The deletion of residues 204-218, however, had no effect on FXII binding suggesting the FXII and HMWK binding sites overlap but are not identical. The proposed FXII binding site is present on the upper face of gC1q-R, whereas the HMWK binding site spans both the upper and lower face. The observation of a larger HMWK binding site is in agreement with data presented in chapter 5 reporting the presence of multiple D5 binding sites for both the N and C-termini. However, the current proposed HMWK binding is broad and the positions of each of the individual D5 binding sites are not known. Attempts to narrow down the binding site to a more specific region must be made, and the D5 truncations produced will allow for the mapping of the individual D5 binding sites on gC1q-R creating a more specific and detailed map of the HMWK binding region on gC1q-R. Due to the very distinct ITC curves produced upon the binding of the various D5

ligands with gC1q-R, subtle changes in binding can be detected between wild type and mutant gC1q-R allowing for key HWMK binding sites to be identified with higher sensitivity than previous ELISA experiments.

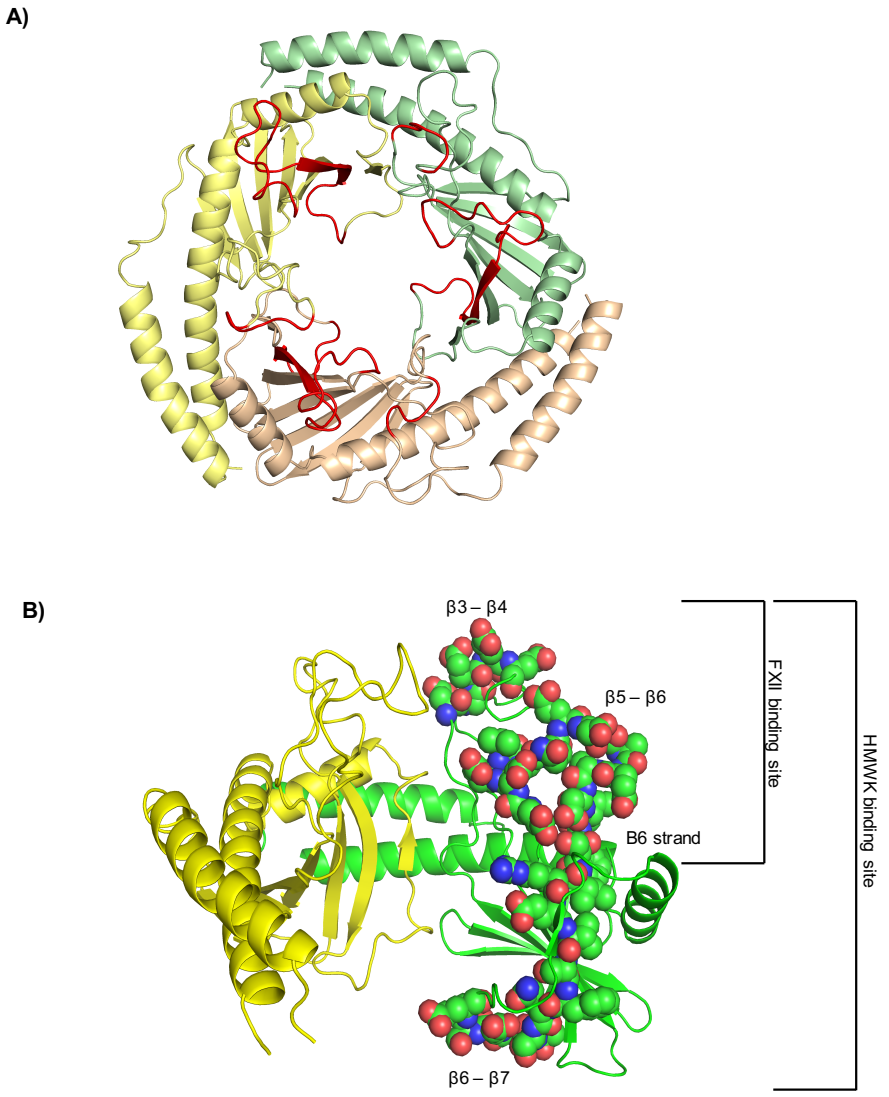


Figure 6.1: *The HWMK binding region, as proposed by Ghebrehiwet et al [155], highlighted on gC1q-R. This model was built based on the 1P32 crystal structure and includes the unresolved flexible loop regions.* ***A)*** *A birds eye view of the gC1q-R trimer with the HMWK binding site highlighted in red. The proposed binding patch spans from the β3-β4 and β5-β6 loops, and down the β6 strand to the β6-β7 loop.* ***B)*** *A side on view of gC1q-R with the HMWK binding site highlighted as spheres on a single monomer. One monomer has been removed from the image in order to make visualisation of the binding site easier. Within this binding region, carbon, oxygen and nitrogen atoms are coloured green, red and blue respectively, and the overlapping HMWK and FXII binding site regions are shown on the side.*

6.1.1 Aims

The aims of this chapter were to produce a range of gC1q-R mutants which would be used in attempts to localise the key D5 binding sites. The proposed binding region, consisting of amino acids 190-218, was the prime target for these mutations along with other acidic rich patches present in nearby disordered loops. Furthermore, a gC1q-R truncation lacking the N-terminal αA helix was produced in an attempt to break the trimer into the monomeric subunits and test the binding of D5 to the monomeric form of gC1q-R. The binding of these mutants to full length D5 and D5 truncations was quantified using ITC, and binding affinities were compared to gC1q-R WT.

6.1.2 gC1q-R Mutant Constructs

The first mutant made was a gC1q-R N-terminal helix truncation where the αA helix, corresponding to the amino acid sequence HTDGDKAFVDFLSDEIKEERK, was deleted (Figure 6.2). This mutant has been suggested to prevent trimer formation due to the fact this helix makes key contacts with the αB helix of the adjacent monomer [155]. This mutant was made in order to disrupt trimer formation and test D5 binding to the monomeric form of gC1q-R.

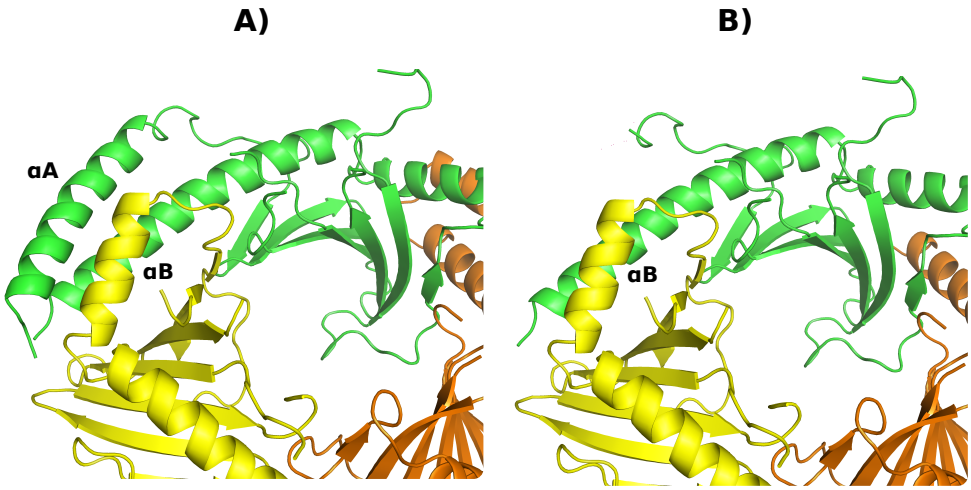


Figure 6.2: ***A)*** *The crystal structure of gC1q-R showing the packing of the αA helix from one monomer with the αB helix of the neighbouring monomer, resulting in stabilisation of the trimer.* ***B)*** *A representation of the αA helix deletion mutant which has been shown to destabilise the trimer into the monomeric form.*

Further gC1q-R mutants were produced where regions from the proposed HWMK binding site were deleted in attempts to narrow down the key D5 binding sites. The entire HMKW binding site was not deleted as this would have removed the the entire β6 strand, and previous analytical gel filtration data on a β6 deletion mutant revealed that it was unable to from a stable trimer [155]. In this study, the gC1q-R mutations were designed to cause minimal disruption to the overall structure of the receptor; therefore, the key areas targeted were acidic rich flexible loop regions. Acidic rich patches were targeted because the basic and acidic natures of D5 and gC1q-R respectively suggests electrostatic interactions play an important role in the interaction of these two proteins.

The first mutations made targeted the flexible loops situated above and below the β6 strand. gC1q-R $\Delta\beta$5-β6 had the acidic sequence Glu196-Glu200, found in the β5-β6 loop region above the β6 strand, deleted from the protein. gC1q-R $\Delta\beta$6-β7 consisted of the deletion of the β6-β7 loop that points into the central cavity. Due to the flexible nature of these regions, it was anticipated that the deletion of these residues should not have caused any significant changes in

protein secondary or tertiary structure; therefore, any changes in D5 binding would be due to the deletion of key binding residues and not due to misfolding of the receptor. The final mutant made focused on short acidic clusters found within the β3-β4 loop. gC1q-R $\Delta\beta$3-β4 had a short cluster of Glu residues, Glu146-148, substituted for Ala. These residues had been previously targeted by Ghebrehiwet [155] and were shown to reduce HMWK binding. Additionally, gC1q-R $\Delta\beta$3-β4 targeted a second Glu rich patch, Glu156-157, also present within the β3-β4 loop. In contrast to the deletion mutants, gC1q-R $\Delta\beta$3-β4 consisted Ala substitutions. Substitutions were not used for the other mutants as the targeted patches were significantly larger, and introducing a large chain of hydrophobic residues would have had a larger affect the structure. However, seeing as gC1q-R $\Delta\beta$3-β4 consisted of two very short acidic patches, the Ala substitutions had reduced risks of introducing unwanted hydrophobic effects and protein misfolding. These substitutions, unlike the deletions, had the advantage of keeping the targeted loop region at a similar length to the wild type protein.

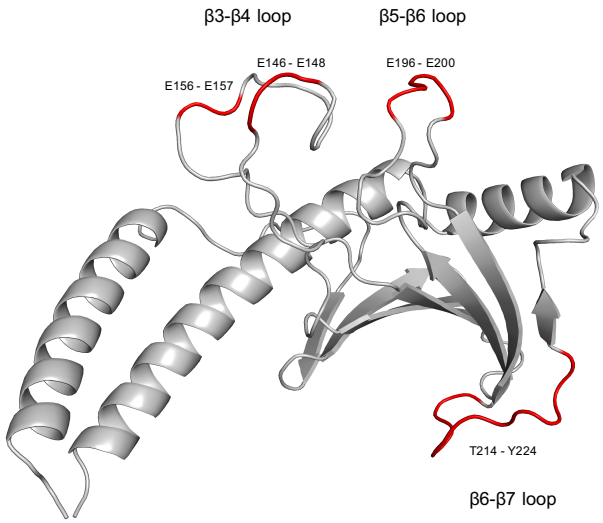


Figure 6.3: *The structure of a single monomer of gC1q-R with the three loop regions that were mutated, in attempts to narrow down the D5 binding sites, highlighted in red. Residues Glu146, Glu147, Glu148, Glu156 and Glu157 from the β3-β4 loop were substituted to Ala residues, and a second mutant had residues Glu196-Glu200 within the β5-β6 loop deleted. A final mutant was made where the β6-β7 loop was deleted.*

Mutant name	Region targeted	Residues mutated	Anticipated effect
gC1q-R ΔN	αA helix	Deletion of His75-Lys95	Trimer breakdown
gC1q-R Δβ3-β4	β3-β4 loop	Glu146, 147, 148Ala Glu156, 157Ala	Disruption of D5 binding
gC1q-R Δβ5-β6	β5-β6 loop	Deletion of Glu196-Glu200	Disruption of D5 binding
gC1q-R Δβ6 - β7	β6-β7 loop	Deletion of Thr214-Tyr224	Disruption of D5 binding

Table 6.1: *Summary of the gC1q-R mutations made highlighting the regions targeted and their desired effect.*

6.2 Results

6.2.1 gC1q-R Mutation Expression and Purification.

All gC1q-R mutants were expressed and purified using the same method as gC1q-R WT. The mutations did not have any significant effect on the theoretical pI's of the protein and, therefore, they all bound with similar binding affinities to the anionic exchange column. Additionally, protein expression levels and solubility were not significantly affected and protein yeilds were comparable to gC1q-R WT.

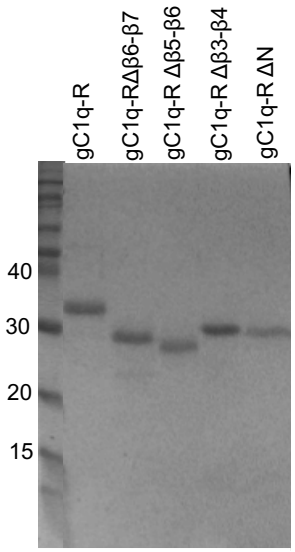


Figure 6.4: *SDS-PAGE gel of the four purified gC1q-R mutants compared with gC1q-R WT. The first lane is the protein marker where useful sizes in kDa have been marked on the side. gC1q-R WT runs at 32 kDa, both gC1q-R Δβ6-β7 and gC1q-R Δβ5-β6 run slightly lower at roughly 29kDa and 28kDa respectively and both gC1q-R Δβ3-β4 and gC1q-R ΔN run at 30kDa.*

6.2.2 Characterisation of gC1q-R Mutants.

6.2.2.1 Secondary Structure

CD was used to assess the effects that the various gC1q-R mutations had on the secondary structure (Figure 6.5). All four mutants gave comparable spectra to gC1q-R WT with only subtle changes in CD profile. Both gC1q-R ΔN and Δ*β6-β7* had slightly increased minima at 207 nm and decreased minima at 222 nm. These spectra were more comparable with the gC1q-R WT spectrum at slightly elevated temperatures of approximately 35 °C. Furthermore, both these mutants did not undergo the same low temperature melting transition seen in the wild type protein, proposed to be due to the unfolding of the *α*A or *α*C helices. gC1q-R ΔN does not have the N-terminal *α*A helix and explains the change in spectrum and reduced melting curve as there is no N-terminal helix to unwind.

The deletion of the β6-β7 loop could have slightly distorted the β sheet and affected the interaction and stabilisation of the αC helix of the adjacent monomer. It's therefore possible that the slight changes in spectrum and melting curve for $\Delta\beta$6-β7 were due to the destabilisation of the αC helix. The effects of both these mutations, however, were small and did not produce major distortions of the secondary structure of gC1q-R.

Both the spectra of gC1q-R $\Delta\beta$5-β6 and gC1q-R $\Delta\beta$3-β4, on the other hand, were more comparable with gC1q-R WT at lower temperatures of approximately 10 °C, and they both had decreased minima at 207 nm and increased minima at 222 nm. Furthermore, the melting curve is shifted to the right by roughly 10 °C which could be due to increased hydrophobic stabilisation of the αC helix upon the removal of acidic loop regions. Once again, the changes in secondary structure were only minimal and did not represent major misfolding of the receptor.

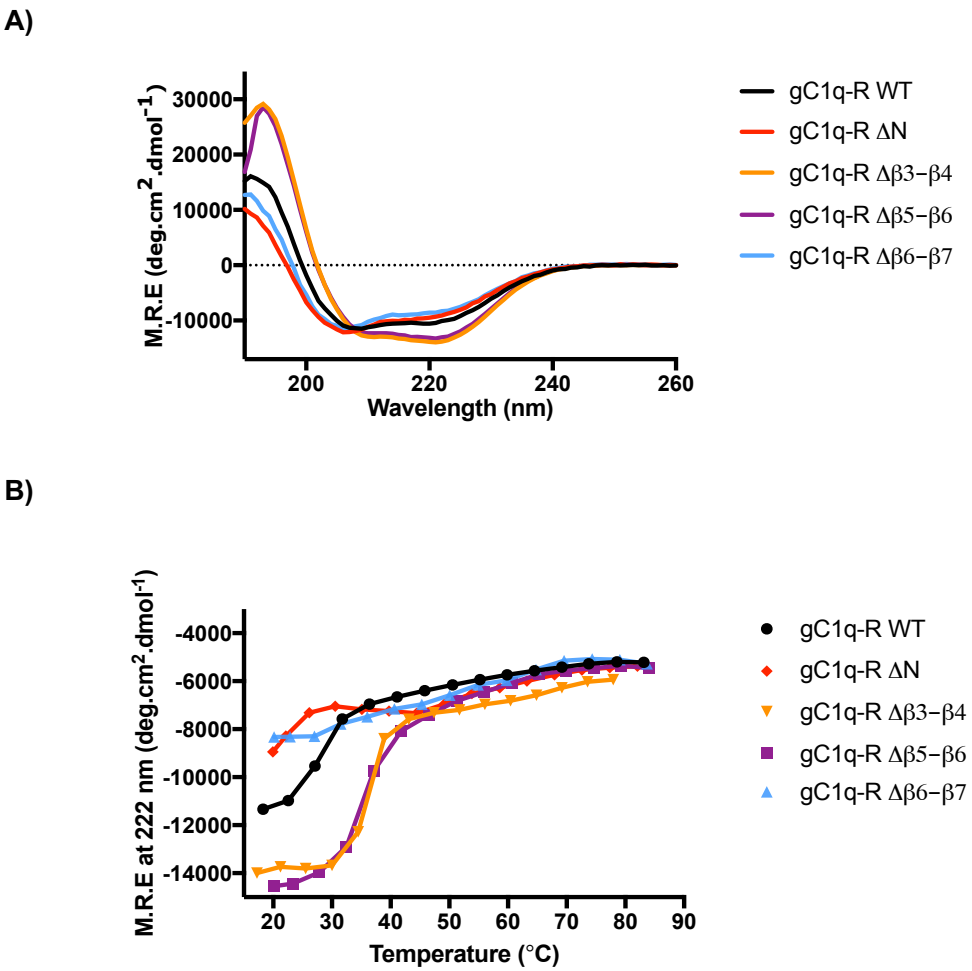


Figure 6.5: *Secondary structure analysis of the gC1q-R mutants compared with gC1q-R WT using CD.* ***A)*** *Far UV CD spectra of the gC1q-R mutants and gC1q-R WT (5 µM each) at 298 K.* ***B)*** *Thermal denaturation profiles, where M.R.E at 222nm is plotted against temperature, of the gC1q-R mutants and gC1q-R WT*

6.2.2.2 Trimer Formation

Analytical gel filtration elution profiles for the gC1q-R mutants were all comparable to gC1q-R WT (Figure 6.6) and confirmed that the mutations were still able to form the trimeric species. The elution peaks for the mutants all had small, high molecular weight shoulders corresponding to the formation of higher molecular weight species. These high molecular weight peaks eluted within the void

volume of the column and were caused by the formation of larger gC1q-R species. The identity of this species is unclear but is likely due to the formation of large protein aggregates or the assembly of multiple gC1q-R trimers. These peaks were small compared with the major trimeric peak and did not represent the major species present. The combination of CD and analytical gel filtration both provided strong evidence that the structure of each mutant remained largely unchanged and consistent with gC1q-R WT.

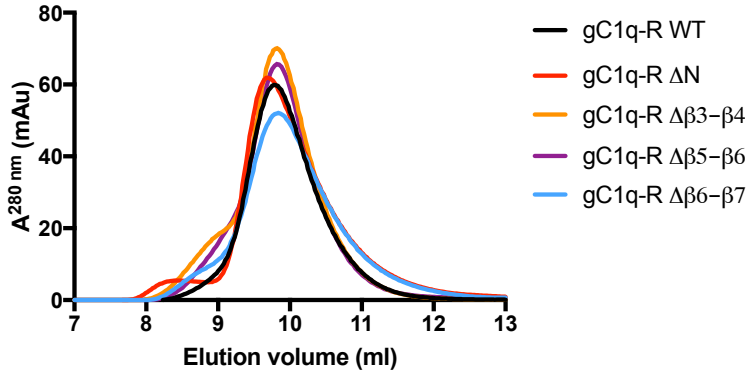


Figure 6.6: *Analytical gel filtration elution profiles for the gC1q-R mutants and gC1q-R WT (10 μM each). All elution profiles were consistent with gC1q-R WT, indicating that each mutant was still able to form a stable trimer.*

gC1q-R ΔN was still able to form a stable trimer in solution, and no monomeric species could be detected by gel filtration. This is in contrast with previous studies that suggested the removal of the αN helix disrupted trimer formation [155], but it is consistent with a solved X-Ray crystallography structure of this mutant [255] which further confirms the ability of this mutant to form a trimer in the absence of the αA helix. A comparison of the crystal structures for gC1q-R ΔN and gC1q-R WT is presented in Figure 6.7. There is very little distortion of the overall fold of gC1q-R ΔN when compared with gC1q-R WT suggesting that the αA helix is not as structurally integral as previously predicted. The most significant structural change is the elongation of the αB helix which causes it to shift away from the centre of the structure. This is coupled with another small shift within the β6 and β7 strands, but combined these changes have little effect

on the overall structure. Combining the CD data and analytical gel filtration data with the already solved crystal structure of gC1q-R ΔN, it is now clear that the major trimer stabilising interactions are not within the αA helix and likely lie within the αC helix interactions with the β sheet and αB helix of the adjacent monomer.

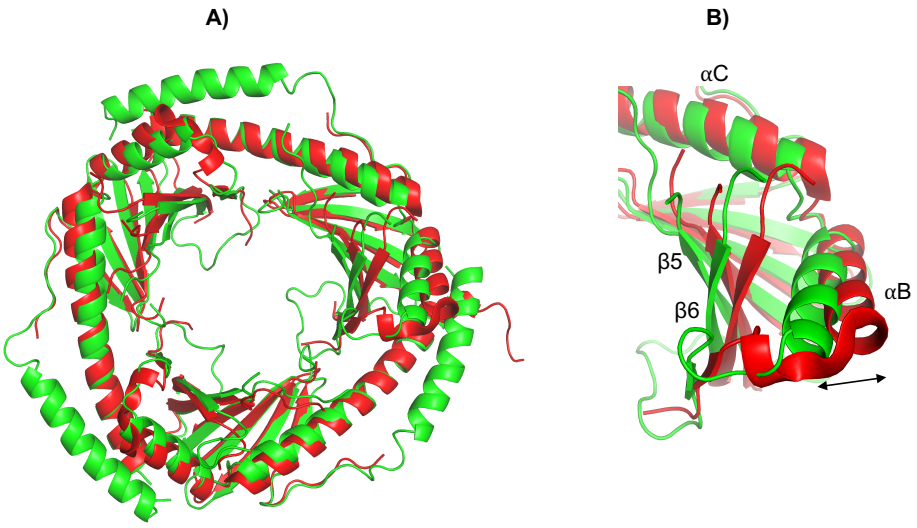


Figure 6.7: *gC1q-R ΔN crystal structure (PDB ID: 3RPX [255]) compared with gC1q-R WT.* ***A)*** *gC1q-R ΔN trimeric structure (red) overlaid with gC1q-R WT (green).* ***B)*** *Zoomed in view of the extended and slightly shifted αB helix, and the shifted β6 and β7 strands.*

6.2.3 Binding of gC1q-R Mutants to D5.

The ability of D5 to bind the gC1q-R mutants was tested by ITC (Figure 6.8). All the mutants, excluding gC1q-R Δβ6-β7, gave comparable isotherms to gC1q-R WT when titrated with full length D5 suggesting that these mutations had not targeted key D5 binding residues. The titration of D5 with gC1q-R Δβ6-β7 revealed a drastic change in binding with a large reduction in ΔH and an overall distortion of the sequential binding curve. An accurate fit could not be made for this binding curve due to the distorted nature of the isotherm.

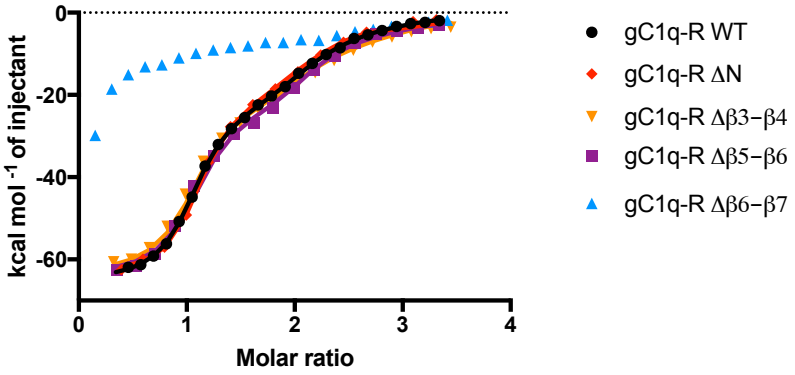


Figure 6.8: *ITC binding curves produced upon the titration of D5 (100 µM) with the various gC1q-R mutants (5 µM each) compared with gC1q-R WT. All binding curves, excluding gC1q-R Δβ6-β7, were fit to a three site sequential binding model. The gC1q-R Δβ6-β7 binding curve could not be fit to a binding model.*

	K_D1 (nM)	K_D2 (nM)	K_D3 (nM)
gC1q-R WT	1.9 (± 0.1)	64.9 (± 1.9)	1011.1 (± 75.8)
gC1q-R ΔN	3.4 (± 0.4)	120.3 (± 11.5)	1083.4 (± 176.6)
gC1q-R Δβ3-β4	1.2 (± 0.1)	33.4 (± 5.7)	724.6 (± 178.3)
gC1q-R Δβ5-β6	1.5 (± 0.2)	44.8 (± 6.0)	1283.7 (± 238.8)
gC1q-R Δβ6-β7		No fit	

Table 6.2: *ITC-derived K_D values for the three site sequential binding interaction of D5 with the gC1q-R mutants compared with gC1q-R WT. Excluding gC1q-R Δβ6-β7, all mutant K_D values were comparable with gC1q-R WT, and K_D values were not affected by more than 2-fold.*

The gC1q-R mutants were further tested for their binding to both D5-1 and D5-2 (Figure 6.9). Similar to D5 binding, all gC1q-R mutants, excluding gC1q-R *β6-β7*, gave comparable binding affinities when binding to D5-1 or D5-2. gC1q-R *β6-β7* gave no detectable signs of binding with D5-1, whereas D5-2 bound this mutant but with significant changes in binding isotherm when compared with gC1q-R WT. There was a two fold reduction in ΔH and a reduced N value of 0.5 implying that only partial binding was detected between D5-2 and gC1q-R *β6-β7*. The K_D, however, was only 3 fold smaller than gC1q-R WT suggesting that this mutation had not significantly affected the binding affinity, and the

major changes were the large reduction in binding stoichiometry. The complete knock out of D5-1 binding combined with partial binding of D5-2 is consistent with the reduced binding response detected between full length D5 and gC1q-R *β6-β7*. These results suggest that the *β6-β7* loop is the major binding site for the N-terminus of D5 but only a partial site for the C-terminus.

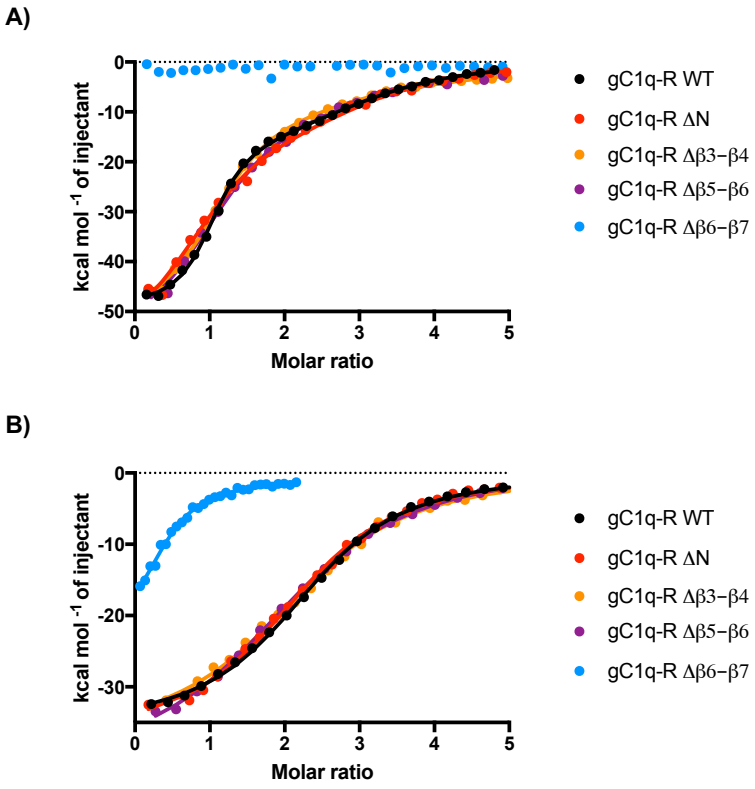


Figure 6.9: A) *ITC binding curves produced upon the titration of D5-1 (100 μM) with the various gC1q-R mutants (3 μM each) compared with gC1q-R WT. All isotherms, excluding gC1q-R Δβ6-β7, were fit to a three site sequential binding model. gC1q-R Δβ6-β7 produced no binding curve when titrated with D5-1. All experiments were performed after saturation with zinc.* ***B)*** *ITC binding curves produced upon the titration of D5-2 (100 μM) with the various gC1q-R mutants (3 μM each) compared with gC1q-R WT. All isotherms were fit to a single site binding model and, excluding Δβ6-β7, binding curves were all consistent with gC1q-R WT. gC1q-R Δβ6-β7 showed a significant decrease in ΔH and stoichiometry when compared with gC1q-R WT and the other gC1q-R mutants.*

	K_D1 (nM)	K_D2 (nM)	K_D3 (nM)
gC1q-R WT	73.0 (± 4.0)	1579.8 (± 145.3)	2673.8 (± 257.7)
gC1q-R /*Delta*N	120.8 (± 55.4)	680.3 (± 420.0)	1577.3 (± 369.1)
gC1q-R Δβ5-β6	131.1 (± 43.0)	1084.6 (± 517.4)	3663.0 (± 1018.3)
gC1q-R Δβ3-β4	135.8 (± 37.3)	1228.5 (± 407.9)	3690.0 (± 638.4)
gC1q-R Δβ6-β7		No binding	

Table 6.3: *ITC-derived K_D values for the three site sequential binding of D5-1 with the gC1q-R mutants compared with gC1q-R WT. Excluding gC1q-R Δβ6-β7, all mutant K_D values were comparable with gC1q-R WT, and K_D values were not affected by more than 3-fold.*

	N	K_D (nM)
gC1q-R WT	2.3 (± 0.01)	763.4(± 27.5)
gC1q-R ΔN	2.3 (± 0.02)	800.0 (± 45.6)
gC1q-R Δβ5-β6	2.1 (± 0.02)	885.0 (± 31.9)
gC1q-R Δβ3-β4	2.3 (± 0.03)	724.6 (± 50.0)
gC1q-R Δβ6-β7	0.5 (±0.02)	2032.5 (± 233.7)

Table 6.4: *ITC-derived N and K_D values for the single site binding of D5-2 with the gC1q-R mutants compared with gC1q-R WT. Excluding gC1q-R Δβ6-β7, all mutant K_D values were comparable with gC1q-R WT, and K_D values were not affected by more than 3-fold. gC1q-R Δβ6-β7 was associated with a significant reduction in K_D and N.*

6.3 Discussion

A range of different mutants were studied but only one resulted in a detectable change in D5 binding. The β6-β7 loop was identified as a major binding site for the N-terminus of D5, and the deletion of this site also had an effect on the assembly of multiple D5-2 ligands at the receptor. Results documented in chapter 5 suggest that both D5 regions bind to separate binding sites, so the fact this one 10 aa deletion had different effects on each region provides further evidence for the presence of two D5 binding sites on gC1q-R. The reasons for only partial disruption of D5-2 binding are unclear, but it is possible that this loop is critical for the correct arrangement and orientation of the three monomers required for

binding multiple D5-2 ligands. Both CD and analytical gel filtration confirmed that the deletion of this loop had little effect on the secondary structure and elution profile of the receptor, so the changes in gC1q-R structure would have been small. Alternatively, the β6-β7 loop could comprise a partial D5-2 binding site which is why D5-2 binding is not completely removed. More mutation studies must be performed in order to fully knock out D5-2 binding and map the entire D5 binding site on the receptor.

Mutating acidic rich regions within the β3-β4 and β5-β6 loops, present on the negative face, had no effect on the binding with D5, and the D5 binding site appears to be localised towards the central cavity of the neutral face. This data is in contrast with previous data where HMWK binding was most affected by deletions within the αA helix, and the β3-β4 and β5-β6 loops [155]. Discrepancies between the results documented in this chapter and previous mutation studies could be due to the indirect binding methods used by Ghebrehiwet et al; this previous study used solid phase ELISA to detect the degree of HMWK binding with the mutants. This method involved the binding of gC1q-R and gC1q-R mutants to microtiter plates, and the degree of HMWK binding was then detected using an antibody specific to HMWK. ITC is a more reliable method as it measures protein-protein interactions directly and does not require a secondary interaction for detection. Moreover, the interaction of gC1q-R with D5 gives a unique ITC profile which is characteristic of this system, and any disruption in binding is clearly visible.

By highlighting each β6-β7 loop (Figure 6.10) the sequential nature of D5 binding can be better understood. All three β6-β7 loops point towards the central cavity where the distances between the three monomers is at its shortest, therefore, the binding of multiple D5 ligands would be predicted to result in steric clashes between each site. The subsequent binding of a third D5 to the final site would exhibit steric clashes between both the first and second D5 ligand.

This might explain the negative cooperativity shown with both full length D5 and D5-1. D5-2 does not exhibit the same cooperative due to the fact the *β6-β7* loop is only partially involved in D5-2 binding, and the full binding site is likely further removed from the central cavity where steric clashes will not be as extreme.

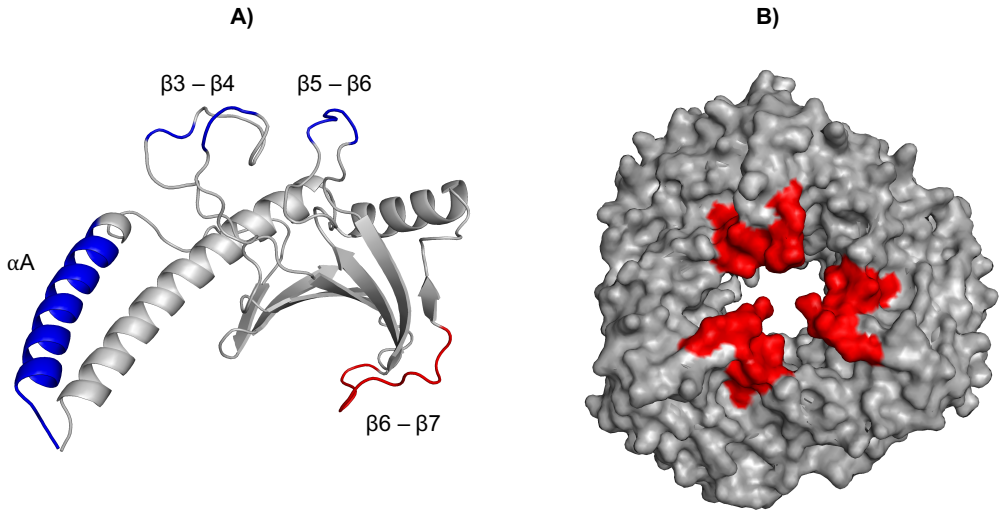


Figure 6.10: ***A)*** *gC1q-R monomer with the mutated regions highlighted in either red or blue. Regions highlighted in red correspond to residues directly involved in D5 binding, whereas regions highlighted in blue correspond to regions not involved in D5 binding.* ***B)*** *Surface diagram of the neutral face of the gC1q-R trimer with the key D5 binding site, within the β6-β7 loop, highlighted in red.*

One major question that arises from these data is how gC1q-R is orientated on the cell membrane so that both faces are accessible for ligand binding . The FXII binding site, along with a portion of the HMWK binding site, is situated on the negative face of gC1q-R [155], whereas in this study the D5-1 binding site of HMWK has been mapped to the neutral face of gC1q-R. It is hard to visualise a way in which membrane bound gC1q-R could orientate itself to expose both faces simultaneously. One option is that the 73 aa N-terminal amino acid sequences present on each monomer not only tether gC1q-R to the membrane but also act like a set of three stilts lifting the gC1q-R trimer away from the

cell membrane and making the neutral face more accessible for ligand binding. Due to the flexible nature of D5, this domain could wrap round the outside of the receptor and make contact with both faces whilst FXII binds the the more accessible negative face of gC1q-R.

CHAPTER 7

Crystallising gC1q-R in Complex With D5 Ligands

7.1 Introduction

Solving the crystal structure in 1999 provided major structural detail for gC1q-R and acted as a framework for additional gC1q-R biochemical experiments. Additionally, due to the discovery of the trimeric nature of the receptor, new models were proposed for potential mechanisms of gC1q-R mediated contact activation. Solving the structure of gC1q-R bound to D5 has the potential to further our understanding of contact activation to a similar extend that solving the initial gC1q-R structure did.

There is no structure available for full length HMWK or any of its isolated domains, but crystal structures of HMWK-derived peptides bound to other proteins have been solved. A short peptide derived from D6 has been co-crystallised with FXI [256] and bradykinin has been crystallised in complex with bacterial exopeptidases [257] and oligopeptide binding proteins involved

in the transportation of peptides across cell membranes [258]. There are, however, no previous reports of any solved crystal structures involving D5-derived peptides and this could largely be due to the difficulty involved in crystallising flexible proteins [259]. It is possible that, due to the tight, low nM affinity complex of D5 and gC1q-R, both these proteins could be co-crystallised if the complex structure is sufficiently ordered.

7.2 Aims

The aim of this section was to co-crystallise and solve the structure of the gC1q-R-D5 complex. NMR data presented in section 4.2.4.3 suggested that D5 remains largely disordered after binding to gC1q-R, so the smaller N and C-terminal D5 truncations, as well as the Lys-rich HMWK 493-516 peptide, were also used for co-crystallisation trials. It was believed that the smaller binding ligands would prove more likely to crystallise due to the reduction in flexible unbound regions. Once crystals were produced, attempts to collect X-ray diffraction data and solve the crystal structure were made.

7.3 Results

7.3.1 Crystallising the gC1q-R-D5 Complex

Full length D5 was the first ligand used in attempts to co-crystallise gC1q-R with a D5-derived ligand. gC1q-R was saturated with D5 by adding a 5 fold excess of D5 ligand. This was then loaded onto a gel filtration column equilibrated in AGF Buffer in order to separate any unbound ligand. Fractions from the gel filtration were analysed using SDS-PAGE to confirm that only peaks

corresponding to the co-migration of gC1q-R and D5 were used and, therefore, the only species present would be the gC1q-R-D5 complex. The fractions corresponding to the gC1q-R-D5 complex were pooled and spin concentrated to 7.5-15 mg/mL. A summary of the screens tested is shown in Table 7.1.

Protein Concentration (mg / ml)	Screen	Wells with Crystals
7.5	PACT	None
8.3	PEG	None
8.3	Morpheus	None
8.3	MPD	None
8.3	Pro Complex	None
15	Pro Complex	None
8.3	Grid Screen Salt	B6, B12

Table 7.1: *Summary of the crystal screens used in the attempts to crystallise the complex of gC1q-R and D5. All crystals were grown at 20 °C and took between 1-6 weeks to fully form.*

Attempts to co-crystallise these two proteins were largely unsuccessful and very few crystals were obtained. The only conditions that yielded any crystals were the high ammonium sulphate, basic pH conditions found in B6 and B12 of the Grid Screen Salt screen (Figure 7.1). Crystals from condition B6 gave no X-ray diffraction when subject to X-rays at the synchrotron, whereas B12 gave diffraction to 6.2 Å which was too poor to solve the structure.

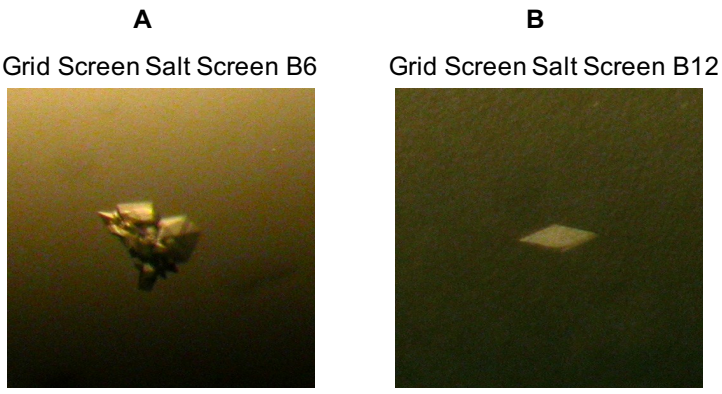


Figure 7.1: *Crystals produced from the gC1q-R-D5 Grid Screen Salt screen.* ***A)*** *Crystal cluster from B6 produced in the presence 0.1 M bicine (pH 9.0), 2.4 M ammonium sulphate.* ***B)*** *Single crystal from B12 produced in the presence of 0.1 M bicine (pH 9.0), 3.0 M ammonium sulphate.*

In order to improve the diffraction resolution for the crystals produced in the Grid Screen Salt screen, optimisation trials were performed based on the conditions found in B6 and B12 in attempts to grow larger, higher quality crystals. The pH was varied from pH 8.4-9.8 and the ammonium sulphate concentration was varied from 1.4-3.3 M. The optimisation screen was mostly unsuccessful and no crystals comparable to the original B6 and B12 crystals (Figure 7.1) were produced. Single rod-like crystals were produced in the presence of 3 M ammonium sulphate at pH 9.4 (Figure 7.3) but these crystals gave no X-ray diffraction.

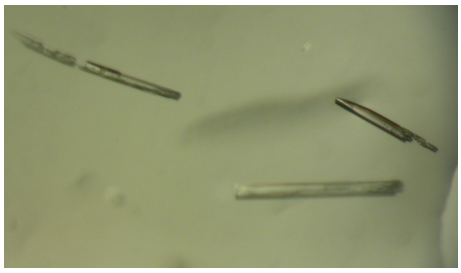


Figure 7.2: *Rod like crystals from F9 of the gC1q-R-D5 optimisation tray produced in the presence of 0.1 M bicine (pH 9.4), 3 M ammonium sulphate.*

7.3.2 Crystallising the gC1q-R-D5 Truncation Complex

Due to the lack of success in co-crystallising the gC1q-R-D5 complex, efforts were diverted to the co-crystallisation of gC1q-R with the D5-1 and D5-2 truncations. The larger full length D5 ligand is less likely to self organise into a crystal due to ligand flexibility present even in the bound form [260]; therefore, the smaller ligands would be more likely to crystallise due to reduced dynamic variability.

Similar to full length D5, complexes were prepared by saturating gC1q-R with either D5-1 or D5-2 at 5-fold excess. Once again this was followed by gel filtration in AGF Buffer to remove any unbound ligand. Fractions were analysed by SDS-PAGE to confirm the presence of the complex, pooled together and spin concentrated to 3-5 mg/mL. Due to the zinc-dependent nature of the interaction of D5-1 and gC1q-R, the gel filtration for the gC1q-R-D5-1 complex was performed in the presence of 50 µM $ZnCl_2$. A summary of the screens used is shown in Table 7.2.

Protein Concentration (mg / ml)	Ligand	Screen	Wells with Crystals
3.0	D5-1 with zinc	Cationic	None
3.0	D5-2	Cationic	None
5.0	D5-1 with zinc	Grid Screen Salt	None
3.1	D5-2	Grid Screen Salt	None
5.0	D5-1 with zinc	MPD	None
3.1	D5-2	MPD	E1
5.0	D5-1 with zinc	Pro Complex	None
5.0	D5-2	Pro Complex	None

Table 7.2: *Summary of the crystal screens used in the attempts to crystallise the complex of gC1q-R and either D5-1 or D5-2. All crystals were grown at 20 °C and took between 1-6 weeks to fully form.*

No crystals were obtained for the gC1q-R-D5-1 complex, but the MPD screen yielded large crystals for the gC1q-R-D5-2 complex (Figure 7.3 A) which gave

X-ray diffraction at 2.2 Å resolution. This crystal had the same space group as the 1P32 crystal with a similar unit cell (Table 7.3), and molecular replacement revealed no extra electron density corresponding to D5-2. Furthermore, the gC1q-R structure was identical to that of the unbound gC1q-R suggesting that D5-2 was not present in this crystal.

	Unit Cell Length (Å)			**Unit Cell Angle (°)**			**Space Group**
1P32	a = 58.62	b = 56.48	c = 93.83	$\alpha = 90$	$\beta = 95.99$	$\gamma = 90$	P1 21 1
MPD E1	a = 59.36	b = 56.75	c = 94.75	$\alpha = 90$	$\beta = 96.37$	$\gamma = 90$	P1 21 1

Table 7.3: *Unit cell parameters and space group of the 1P32 gC1q-R structure and the MPD screen E1 crystal.*

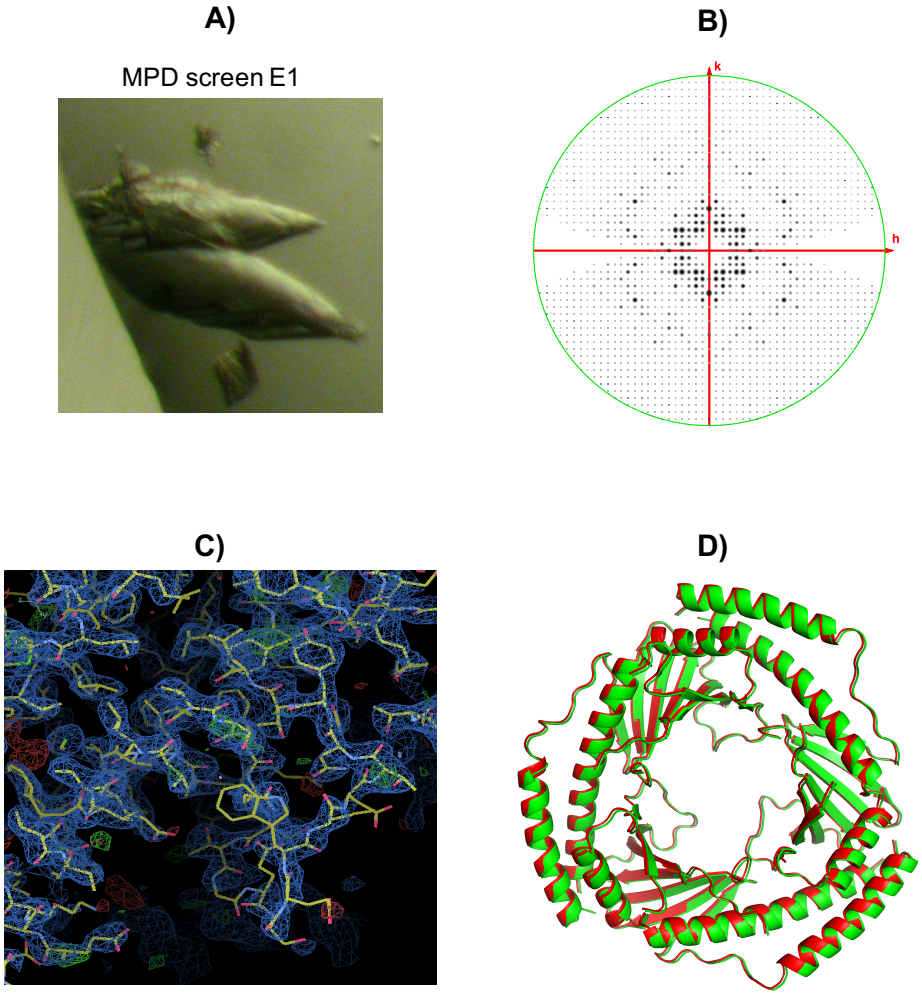


Figure 7.3: ***A)*** *Large crystals from E1 of the gC1q-R-D5-2 MPD screen produced in the presence of 0.1 M citric acid (pH 9.0), 10 % MPD (v/v).* ***B)*** *2.2 Å resolution crystal diffraction pattern from the E1 crystal collected from the Diamond Light Source.* ***C)*** *|Fo|-|Fc| electron density omit map of the E1 crystal after molecular replacement with the 1P32 crystal structure. The electron map is focussed on the β6-β7 loop shown to be involved in D5 binding. There is no extra electron density present within the electron map suggesting that the ligand was not present in the structure.* ***D)*** *Overlay of the 1P32 crystal structure (green) and the E1 crystal structure (cyan). Both structures overlay with one another which further implies the E1 crystal corresponds to unbound gC1q-R.*

7.3.3 Crystallising the gC1q-R-HMWK 493-516 Complex

Finally, crystallisation trials of the HMWK 493-516 peptide were performed. This peptide binds with similar strength to D5-2, but in the bound form it would potentially have fewer flexible, unbound regions due to its reduced size and, therefore, might be easier to crystallise as a complex. Once again gC1q-R was saturated with the ligand by mixing with 10 fold excess of peptide. The sample was then gel filtrated in AGF buffer to remove any unbound ligand, and the fractions corresponding to gC1q-R were spin concentrated to 3-5 mg/mL. Due to the small size of the peptide, it was not possible to confirm that HMWK 493-516 had co-migrated with gC1q-R using SDS-PAGE; therefore, a second protein stock solution was also used where the peptide was mixed with gC1q-R without an additional gel filtration step. The non-gel filtrated samples were dialysed against AGF buffer to ensure these were in comparable conditions to the gel filtrated samples. A summary of the screens tested is shown in Table 7.4.

Protein Concentration (mg / ml)	Ligand : Receptor Ratio	Gel Filtrated	Screen	Wells with Crystals
3.1	10 : 1	No	Cationic	None
5.0	Unknown	Yes	Grid Screen Salt	None
3.1	10 : 1	No	Grid Screen Salt	None
5.0	Unknown	Yes	MPD	None
3.1	10 : 1	No	MPD	A6, A8, B6, C1
5.0	Unknown	Yes	Morpheus	None
5.0	10 : 1	No	Morpheus	None
5.0	Unknown	Yes	Pro Complex	A2
5.0	10 : 1	No	Pro Complex	None

Table 7.4: *Summary of the crystal screens used in the attempts to crystallise the complex of gC1q-R and HMWK 493-516. All crystals were grown at 20 °C and took between 1-6 weeks to fully form.*

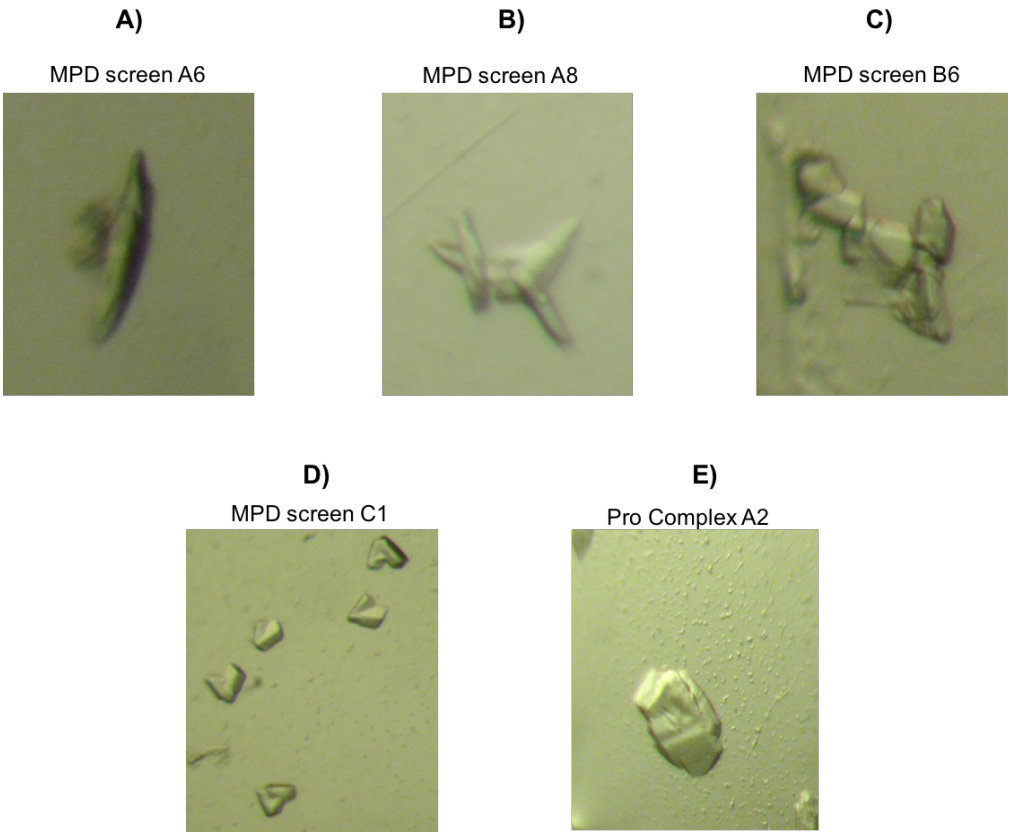


Figure 7.4: *Crystals produced from the gC1q-R-HMWK 493-516 MPD and Pro Complex screens.* ***A)*** *Crystal from A6 of the MPD screen produced in the presence of 0.2 M sodium chloride, 40 % MPD (v/v).* ***B)*** *Crystal cluster from A8 of the MPD screen produced in the presence of 0.2 M potassium chloride, 40 % MPD (v/v).* ***C)*** *Crystal cluster from B6 of the MPD screen produced in the presence of 0.2 M potassium nitrate, 40 % MPD (v/v).* ***D)*** *Rock like crystals from C1 of the MPD screen produced in the presence of 0.2 M magnesium acetate, 40 % MPD (v/v).* ***E)*** *Rock like crystal from A2 of the Pro Complex screen produced in the presence of 0.1 M calcium acetate, 0.1 M MES (pH 6.0), 15 % PEG 400 (v/v).*

The only screens which produced any crystals were the MPD and morpheus screens. Crystals from MPD A6, A8 and B6 (Figure 7.4 A-C) gave no X-ray diffraction, whereas the crystals from C1 (Figure 7.4 D) gave diffraction to 6 Å resolution which, once again, was too poor to resolve the structure. The crystal from A2 of the Pro Complex screen (Figure 7.4 E), although big in size, had not crystallised as a single crystal form; therefore, optimisation trays based on these conditions were set up in order to improve crystal quality. The calcium

acetate concentration was varied from 50-300 mM and PEG 400 concentration was varied from 5-30 % (v/v). From this optimisation tray, a large, single crystal formed in the presence of 0.1 M calcium acetate and 10 % PEG 400 (v/v) (Figure 7.5). Conditions for this crystal were similar to the original A2 Pro Complex screen crystal, but the PEG 400 concentration had been reduced from 15 % (v/v) to 10 % (v/v). This crystal gave poor X-Ray diffraction with a resolution of 8.8 Å which was too low to solve the structure.

Figure 7.5: *Large single crystal from the gC1q-R-HMWK 493-516 Pro Complex A2 optimisation tray produced in the presence of 0.1 M calcium acetate, 0.1 M MES pH 6.0, 10 % PEG 400 (v/v).*

7.4 Discussion

A summary for all the crystals used for diffraction experiments is shown in Table 7.5. Five out of the nine crystals subject to X-rays showed no signs of diffraction, whilst a further three yielded poor diffraction resolutions. This is a common feature for many protein crystals and can be due to poor order of molecules and high solvent volume within the crystal [261]. A number of techniques can be implemented in order to increase diffraction, including crystal dehydration, which is used to reduce solvent volume and improve crystal packing [262]; and seeding, which involves using a crystal as a nucleation point for the growth of larger crystals [263]. These techniques can be performed in future

crystallisation trials along with more crystal condition optimisations in order to increase diffraction resolutions.

The only crystal that gave sufficient X-ray diffraction for structure determination was the gC1q-R-D5-2 crystal, but no ligand was present in the structure. This could be due to the low pH conditions disrupting the interaction of gC1q-R and D5-2. Furthermore, citrate has been shown to directly bind Lys residues in a number of different proteins [264][265][266] and, therefore, could be directly inhibiting the interaction of the Lys-rich region of D5-2 with gC1q-R. Future crystallisation attempts will focus on more neutral pH conditions in the absence of salts such as citrate in order to increase the likelihood that the ligand remains bound to the receptor during the crystallisation process.

Despite not solving the crystal structure of gC1q-R in complex with one the D5 ligands, the work documented in this chapter has narrowed down potential crystallisation conditions for future trials. Furthermore, a number of crystals gave poor diffraction, but these conditions can be used as templates for future optimisation experiments in attempts to increase the diffraction resolution.

Complex	**Crystallisation conditions**	**Diffraction Resolution (Å)**	**Comments**
gC1q-R-D5	0.1 M bicine (pH 9.0), 2.4 M ammonium sulphate	No diffraction	-
gC1q-R-D5	0.1 M bicine (pH 9.0), 3.0 M ammonium sulphate	6.2	Poor diffraction
gC1q-R-D5	0.1 M bicine (pH 9.4), 3.0 M ammonium sulphate	No diffraction	-
gC1q-R-D5-2	0.1 M citric acid (pH 4.0), 10 % MPD (v/v)	2.2	No ligand
gC1q-R-HMWK 493-516	0.2 M sodium chloride, 40 % MPD (v/v)	No diffraction	-
gC1q-R-HMWK 493-516	0.2 M potassium chloride, 40 % MPD (v/v)	No diffraction	-
gC1q-R-HMWK 493-516	0.2 M potassium nitrate, 40 % MPD (v/v)	No diffraction	-
gC1q-R-HMWK 493-516	0.2 M magnesium acetate, 40 % MPD (v/v)	6.0	Poor Diffraction
gC1q-R-HMWK 493-516	0.1 M MES (pH 6.0), 0.1 M calcium acetate, 10 % PEG 400	8.8	Poor diffraction

Table 7.5: *Summary of all the crystals subjected to X-ray diffraction at the synchrotron.*

CHAPTER 8

General Discussion : Further Understanding Contact Activation

The data presented in this thesis are consistent with the original hypothesis that gC1q-R is able to assemble multiple contact factors simultaneously. Multiple binding events, showcasing negative cooperativity, were detected between D5 and gC1q-R by ITC. Not only did this provide the first evidence for multiple ligands binding to gC1q-R but the negatively cooperative binding observed also highlights a key potential mechanism for controlling FXII binding to the gC1q-R trimer. One of the main flaws with this work is the fact all experiments were performed on isolated D5 and not the full length protein; this poses the question of whether the multiple D5 binding events are biologically significant and if the much larger full length HMWK would showcase similar binding properties. Future work will focus on full length HWMK in attempts to observe the same multi-step binding and recreate a more biologically relevant system.

8.1 gC1q-R-D5-FXII Complex

In addition to studying the binding of full length HMWK with gC1q-R, another major step needed to further understand the role of gC1q-R in contact activation is to detect the simultaneous binding of HMWK and FXII and understand the interplay between these two gC1q-R ligands. The hypothesis is that, due to the trimeric structure of gC1q-R, HMWK and FXII are able to bind to one gC1q-R receptor, facilitating contact activation through the assembly of FXII, FXI and PK at one surface. Preliminary gel filtration studies were performed in attempts to observe a ternary complex between gC1q-R, D5 and the FnII domain of FXII. Results were consistent with previous work reporting that HMWK and FXII compete with one another for gC1q-R binding [107] as D5 readily out competed FnII for binding to the receptor. Only at low concentrations of D5 did both D5 and FnII comigrate with gC1q-R. The same experiment will be repeated using the full length proteins in attempts to observe the much larger complex of gC1q-R with full length HMWK and FXII.

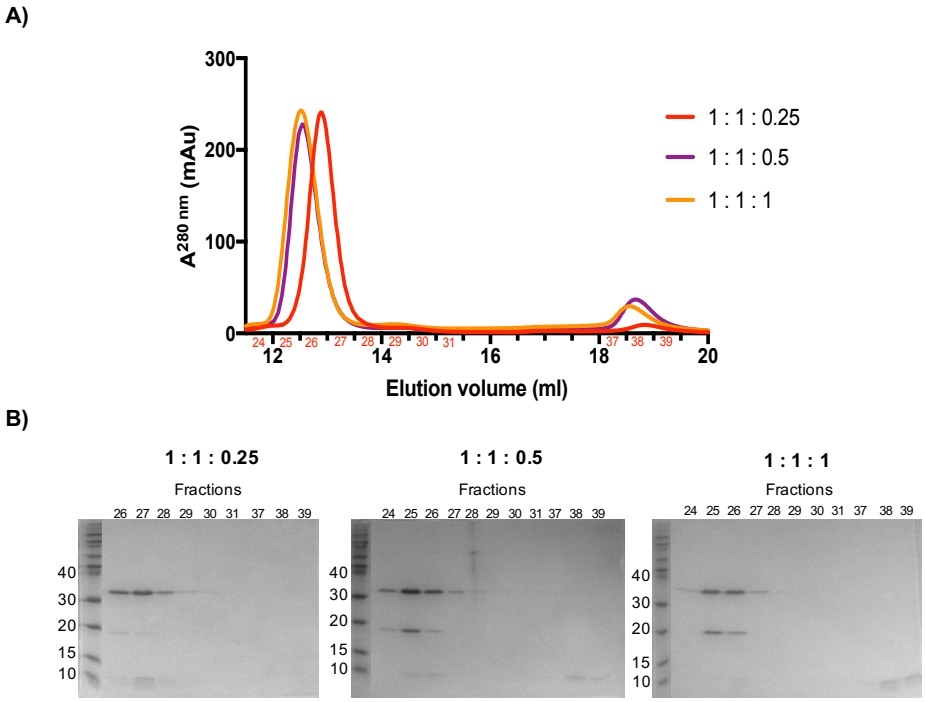


Figure 8.1: *Analytical gel filtration of gC1q-R and the FnII domain of FXII (20 µM each) combined with increasing concentrations of D5.* ***A)*** *Elution profiles of the various complexes with fraction numbers indicated in red.* ***B)*** *SDS-PAGE gels of the 1 ml fractions collected showing the protein content of each peak. The first lane of each gel shows the protein markers with relevant sizes labelled in kDa. gC1q-R, D5 and FnII can be seen at roughly 32, 20 and 10 kDa respectively. As the concentration of D5 is increased, FnII shifts from the high molecular weight peak to the low molecular weight peak suggesting D5 is outcompeting FnII for gC1q-R binding.*

The biological significance of the negatively cooperative binding of D5 with gC1q-R can only be fully appreciated when factoring in FXII. D5 outcompetes the initial FnII domain of FXII due to the strong, low nM binding affinity of D5 for the receptor. However, due to the significantly reduced binding affinities of the second and third binding events, it is possible that FnII could out compete additional D5 ligands for these secondary sites. Unpublished SPR studies looking at the interaction of FnII and gC1q-R calculated a binding affinity in high nM range; therefore, FXII will be outcompeted by the first two D5 ligands, which have greater affinity for gC1q-R, but would compete with the third D5, which has a similar binding affinity. This allows FXII to compete for the final binding site on gC1q-R whilst the other two sites are occupied by HMWK, in-

creasing the chances of having FXI, PK and FXII at one surface. In the absence of negative cooperativity, FXII would be fully outcompeted by D5 and contact activation would not be initiated.

A revised hypothesis therefore, is that binding of HMWK is regulated through the presence of negative cooperativity produced upon the binding of N-terminal region of D5. This allows FXII binding to be specific for the final unbound monomer of gC1q-R and contact initiation occurs.

8.2 Controlling Contact Activation

The mechanisms involved in controlling the assembly of the contact initiators at the cell surface are poorly understood. If the interactions of HMWK and FXII with gC1q-R are not regulated, the contact pathway would constantly be activated; therefore, mechanisms controlling these interactions must be in place. One potential control mechanism could be through the regulation of gC1q-R expression. Studies have shown that endothelial cell expression of gC1q-R is upregulated in the presence of inflammatory cytokines [267]. Cytokines are known to regulate expression of TF [268] [269] [270], and it is possible that these cytokines play similar roles in the contact pathway through gC1q-R expression. Alternatively, increased contact activation could be due to increased presentation or direct targeting of gC1q-R to the cell surface. Upon the damage of endothelial cells, intracellular gC1q-R becomes exposed to the blood which increases binding of FXII and HMWK. This may be biologically relevant during thrombosis where increased contact activation results in uncontrolled fibrin deposition.

An alternative method for controlling contact activation is through the control of Zn^{2+} levels. Both HMWK and FXII bind gC1q-R in a zinc-dependent man-

ner [106], and it is now understood that HMWK zinc dependency is due to the N-terminal region of D5. Therefore, in order to assemble both FXII and HMWK with gC1q-R, a sufficient source of Zn^{2+} must be present. Concentrations of Zn^{2+} in blood plasma are between 10-20 µM [271], most of which is bound to serum albumin [272], resulting in free Zn^{2+} concentrations of approximately 0.5 µM [273]. FXII binding to HUVEC cells was shown to increase in the presence of Zn^{2+}, and binding plateaued at zinc concentrations of 50 µM, 250 fold excess over FXII concentrations [152]. This zinc concentration is 100 fold larger than the free Zn^{2+} concentration found in the blood, and for sufficient binding of FXII to the cell surface, an increase in blood Zn^{2+} is required. Zn^{2+} concentrations were shown to increase by 15-20 % at injured sites within rats [274] which supports Zn^{2+} concentrations as a potential control mechanism for contact activation. Platelets are known to release Zn^{2+} upon activation [275] [46], resulting in higher Zn^{2+} concentrations at sites of platelet aggregation and subsequent contact activation through the zinc-dependent interactions of FXII and HMWK with gC1q-R.

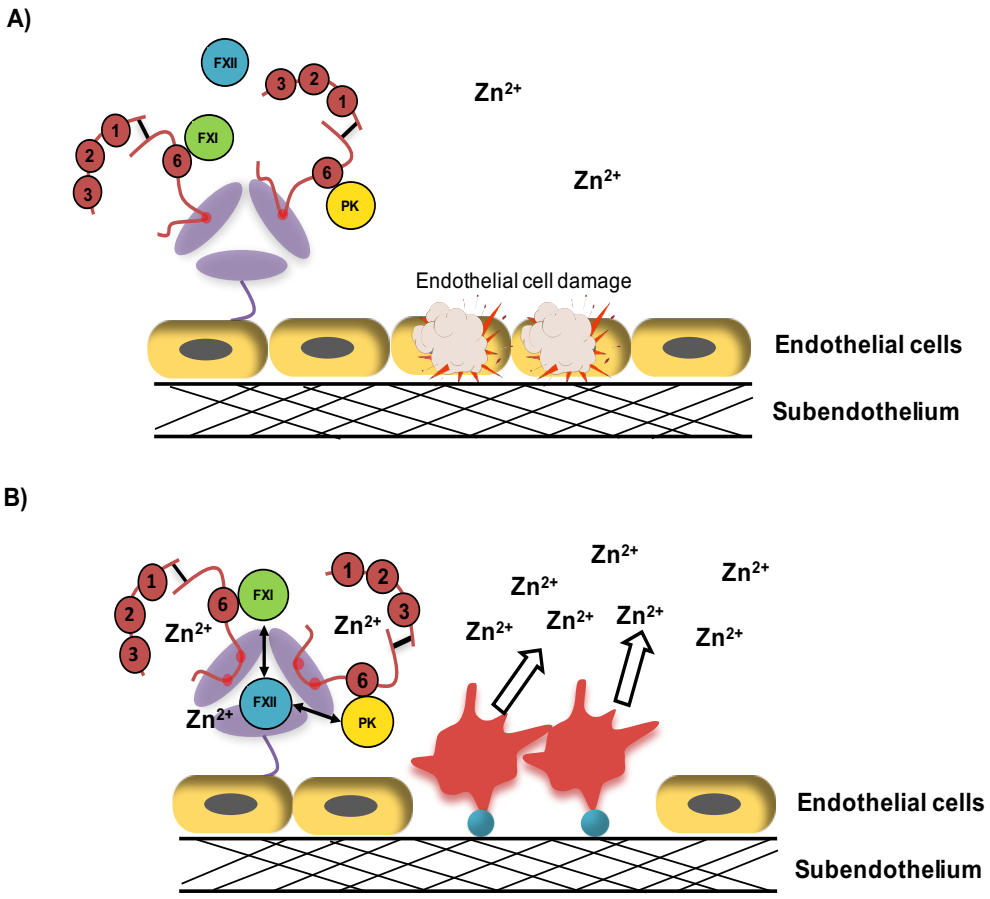


Figure 8.2: *A model for the initial activation of the contact pathway by Zn^{2+}. **A)** Standing concentrations of free Zn^{2+} in the blood are not sufficient for the binding of FXII and HMWK with gC1q-R, and the contact pathway is not activated. HMWK is able to bind gC1q-R through the C-terminal region of D5 in the absence of Zn^{2+} but cannot bind fully through both the N and C-termini; therefore, the binding affinity is reduced. Activation of the contact pathway cannot occur. **B)** Upon endothelial cell damage, platelets are recruited to the site of injury though the interaction with vWF on the subendothelium. Platelet activation results in the release of Zn^{2+} which allows FXII and both the C and N-termini of HMWK to bind with gC1q-R, resulting in contact activation.*

8.3 Contact Activation as a Therapeutic Target for Blood Disorders

There are many blood disorders caused by deficiencies of proteins of the contact pathway. The most well known blood disorder is haemophilia, which is described as excessive bleeding and easy bruising. The two main types of haemophilia are type A and type B, and both subtypes are caused by the compromised activation of FX during the contact pathway [276]. Haemophilia type A is due to genetic deficiency or disfunction in FVIII, which is a co-factor for the FIXa dependent cleavage of FX, whereas haemophilia type B is due to a deficiency in FIX. Treatment of these disorders usually involves supplementation of the missing or dysfunctional clotting factors in order to restore blood coagulation [277].

Disorders involving the contact initiator proteins also exist. Deficiencies in FXII and FXI, unlike other clotting factors, do not result in excessive bleeding [58][278]. This observation has led to the hypothesis that the major role of the contact pathway is bradykinin production and the extrinsic coagulation pathway is the major pathway responsible for the initiation of blood coagulation. However, experiments showing that FXI and FXII deficient mice exhibit defects in the formation and stabilisation of platelet-rich occlusive thrombi suggested that contact activation is responsible for maintaining the thrombus, whilst the extrinsic pathway is involved in the initial formation of the thrombus [279]. As a result, the early stages of contact activation could potentially provide key therapeutic targets for vessel-occluding diseases without increasing the risk of excessive bleeding. Hypercoagulation is an abnormality in blood clotting that results in increased clot formation and can result in thrombosis [280]. Hypercoagulation has traditionally been treated with antiocoagulants such as warfarin. This drug works by inhibiting vitamin K epoxide reductase and reduces the production of vitamin K [281], a crucial cofactor for FIX, FX, FVII and pro-

thrombin [282]. Other anticoagulant drugs include the FXa inhibitors apixaban and rivaroxaban [283] and the thrombin inhibitor dabigatran [284]. One major drawback to these drugs, due to the fact they inhibit proteins late on in the coagulation cascade, is they increase the risk of excessive bleeding and have been linked with brain haemorrhages within patients [285]. Targeting the contact activator proteins as anticoagulant therapies would reduce the stability of the active thrombus without entirely inhibiting blood coagulation. This pathway offers a route to selectively treating thrombosis without affecting normal haemostasis and would have the major benefit of preventing hypercoagulation whilst also reducing risks of uncontrollable bleeding [286]. HMWK, FXII, FXI and PK are all potential drug targets for the inhibition of contact activation.

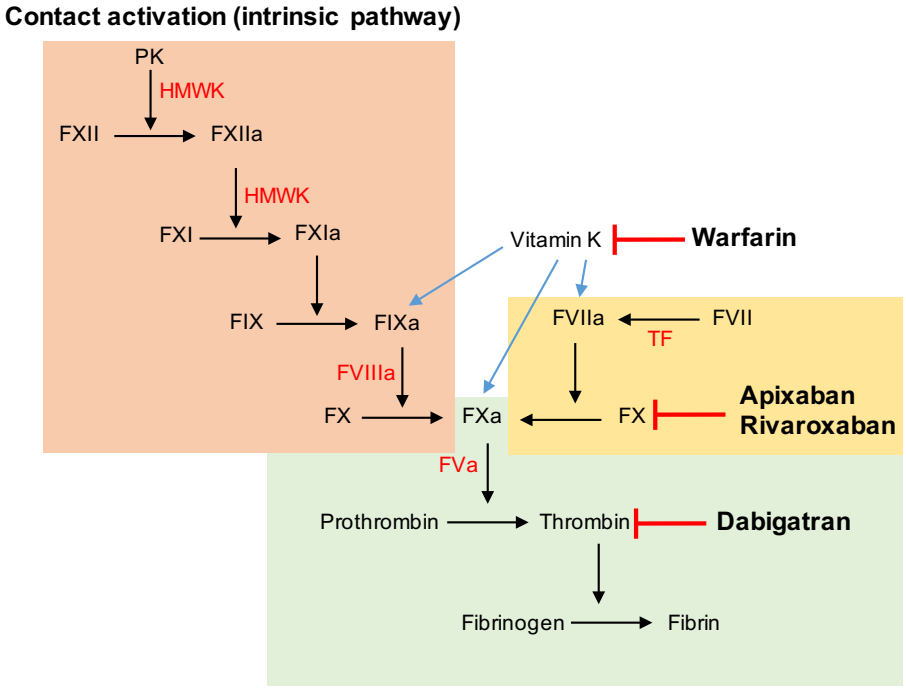


Figure 8.3: *The major antiocoagulant drugs, used to prevent hypercoagulation of the blood, target proteins involved in late stages of the coagulation cascade or the extrinsic pathway. Each drug is highlighted in bold, and the proteins or non protein cofactors that they target are highlighted in red. The use of these drugs is accompanied with an increased risk of uncontrolled bleeding due to the complete inhibition of blood coagulation. Targeting events involved in contact initiation would reduce this risk, as hypercoagulation is prevented but coagulation can still occur through the extrinsic pathway.*

One approach for inhibiting contact activation would be to specifically inhibit the protease activity of FXII, PK or FXI, preventing the initiation of the cleavage cascade. Selective protease inhibitors for PK [287], FXI [288] and FXII [289] have been documented which could be used as anticoagulant therapies in the future. Alternatively, instead of targeting enzyme activity, preventing the presentation of these enzymes to the cell surface would also inhibit contact activation. This strategy has the potential to be far more specific than directly targeting the enzyme activity as the enzymes in the coagulation pathway are all serine proteases which represents an extremely diverse and common class of enzymes; therefore, targeting one specific serine protease can be extremely tricky. Targeting protein presentation at the cell surface could be achieved through inhibiting the interaction of HMWK D6 with PK and FXI or by inhibiting the direct interaction of HMWK D5 with gC1q-R. Small molecules that mimic specific protein binding regions have proved a successful approach when attempting to target and inhibit PPIs [290], and this method could be used to disrupt the interactions of HMWK with gC1q-R, PK and FXI.

The work presented in this thesis has provided valuable mechanistic and structural detail into the interaction of D5 with gC1q-R and can be used as a starting point for the design of novel drugs targeting the assembly of contact initiators. Due to the presence of two independent gC1q-R binding regions, it would be possible to partially inhibit the interaction of D5 and gC1q-R by targeting only one of these regions. This offers a much more sophisticated level of control over the interaction of HMWK and gC1q-R which could be adjusted depending on the degree of hypercoagulation. A starting point for drug design could be analogue mimics of the Lys-rich C-terminal region of D5. These analogues would compete with the C-terminal D5 region of HMWK for gC1q-R binding, decreasing the rate of contact activation, and could be the basis for a novel class of drugs used for the treatment of various coagulation disorders.

8.4 Conclusions

The major objectives for this thesis were to observe the interaction of gC1q-R and D5 of HMWK, map the interaction surfaces and link this interaction with contact activation at the cell surface. The interaction of these two proteins was detected using gel filtration and further quantified using ITC. These data were used to help further understand the mechanisms involved in contact activation, as the observed negatively cooperative, three site sequential binding mechanism would allow multiple HMWK ligands to bind to one receptor whilst also allowing FXII to bind to the low affinity site.

The production of two D5 truncations revealed that D5 has multiple gC1q-R binding sites: a zinc-dependent N-terminal binding site and a zinc-independent C-terminal binding site. The C-terminal binding region was assigned to a 24 long Lys-rich peptide consisting of residues 493-516 of HMWK. Further competition experiments revealed that these two binding regions bind to two distinct sites on gC1q-R. Additionally, D5 binding was located to the central β6-β7 loop of gC1q-R. Attempts to crystallise gC1q-R in complex with the various D5 ligands have so far proved unsuccessful, and the exact binding interface is still not fully understood. Despite this, key interaction patches on gC1q-R have been determined as well as multiple D5 binding regions.

Future work will involve further mapping the exact interaction surface of D5 and gC1q-R, focussing on further crystallisation trials to crystallise the complex of gC1q-R with the D5 ligands. The formation of the gC1q-R-HMWK-FXII ternary complex will be investigated in order to understand the interplay of FXII and HMWK at the cell surface. Long term objectives will be to eventually assemble the full gC1q-R-HMWK-FXII-FXI-PK complex in order to validate the hypothesis for contact initiation. By mapping each individual interaction site involved in the formation of this complex, it will be possible to design novel

therapeutic drugs for the prevention of thrombosis and thrombosis associated diseases. The results presented in this thesis provide the initial ground work for understanding the formation of an endothelial cell multiprotein complex responsible for the activation of the contact pathway.

CHAPTER 9

Appendix

9.1 Appendix 1

9.1.1 Protein Sequences

MKLITILFLCSRLLLSLTQESQSEEIDCNDKDLFKAVDAALKKYNSQNQSNNQFVLY
RITEATKTVGSDTFYSFKYEIKEGDCPVQSGKTWQDCEYKDAAKAATGECTATVGKR
SSTKFSVATQTCQITPAEGPVVTAQYDCLGCVHPISTQSPDLEPILRHGIQYFNNNT
QHSSLFMLNEVKRAQRQVVAGLNFRITYSIVQTNCSKENFLFLTPDCKSLWNGDTGE
CTDNAYIDIQLRIASFSQNCDIYPGKDFVQPPTKICVGCPRDIPTNSPELEETLTHT
ITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESNEELTESCET
KKLGQSLDCNAEVYVVPWEKKIYPTVNCQPLGMISLMKRPPGFSPFRSSRIGEIKEE
TTVSPPHTSMAPAQDEERDSGKEQGHTRRHDWGHEKQRKHNLGHGHKHERDQGHGHQ
RGHGLGHGHEQQHGLGHGHKFKLDDDLEHQGGHVLDHGHKHKHGHGHGKHKNKGKKN
GKHNGWKTEHLASSSEDSTTPSAQTQEKTEGPTPIPSLAKPGVTVTFSDFQDSDLIA
TMMPPISPAPIQSDDDWIPDIQIDPNGLSFNPISDFPDTTSPKCPGRPWKSVSEINP
TTQMKESYYFDLTDGLS

D1 D2 D3 D4 D5 D6

Figure 9.1: *Full length HMWK aa sequence with each domain highlighted a different colour.*

The protein DNA sequencing results were translated using EXPASY Translate tool (http://web.expasy.org/translate/) in order to generated DNA/aa alignments for the protein constructs used.

9.1.1.1 D5

```
atgggcagcagccatcatcatcatcacagcagcggcctggtgccgcgc|ggcagccat
 M  G  S  S  H  H  H  H  H  H  S  S  G  L  V  P  R| G  S  H
atggctagcatgactggtggacagcaaatgggtcgcggatccactgtaagtccaccccac
 M  A  S  M  T  G  G  Q  Q  M  G  R  G  S  T  V  S  P  P  H
acttccatggcacctgcacaagatgaagagcgggattcaggaaaagaacaagggcatact
 T  S  M  A  P  A  Q  D  E  E  R  D  S  G  K  E  Q  G  H  T
cgtagacatgactggggccatgaaaaacaaagaaaacataatcttggccatggccataaa
 R  R  H  D  W  G  H  E  K  Q  R  K  H  N  L  G  H  G  H  K
catgaacgtgaccaagggcatgggcaccaaagaggacatggccttggccatggacacgaa
 H  E  R  D  Q  G  H  G  H  Q  R  G  H  G  L  G  H  G  H  E
caacagcatggtcttggtcatggacataagttcaaacttgatgatgatcttgaacaccaa
 Q  Q  H  G  L  G  H  G  H  K  F  K  L  D  D  D  L  E  H  Q
gggggccatgtccttgaccatggacataagcataagcatggtcatggccacggaaaacat
 G  G  H  V  L  D  H  G  H  K  H  K  H  G  H  G  H  G  K  H
aaaaataaaggcaaaaagaatggaaagcacaatggttggaaaacagagcatttggcaagc
 K  N  K  G  K  K  N  G  K  H  N  G  W  K  T  E  H  L  A  S
tcttctgaagacagttga
 S  S  E  D  S  -
```

Figure 9.2: *Sequencing data for the D5 protein construct with the DNA and aa sequences aligned. The construct used consisted of residues Thr401-Ser531 of HMWK with a cleavable N-terminal His tag, highlighted in red, and a non-cleavable T7 tag, highlighted in blue. The exact point of thrombin cleavage is represented with a black line.*

9.1.1.2 gC1q-R

```
atgcacaccgacggagacaaagcttttgttgatttcctgagtgatgaaattaaggaggaa
 M  H  T  D  G  D  K  A  F  V  D  F  L  S  D  E  I  K  E  E
agaaaaattcagaagcataaaaccctccctaagatgtctggaggttgggagctggaactg
 R  K  I  Q  K  H  K  T  L  P  K  M  S  G  G  W  E  L  E  L
aatgggacagaagcgaaattagtgcggaaagttgccggggaaaaaatcacggtcactttc
 N  G  T  E  A  K  L  V  R  K  V  A  G  E  K  I  T  V  T  F
aacattaacaacagcatcccaccaacatttgatggtgaggaggaaccctcgcaagggcag
 N  I  N  N  S  I  P  P  T  F  D  G  E  E  E  P  S  Q  G  Q
aaggttgaagaacaggagcctgaactgacatcaactcccaatttcgtggttgaagttata
 K  V  E  E  Q  E  P  E  L  T  S  T  P  N  F  V  V  E  V  I
aagaatgatgatggcaagaaggcccttgtgttggactgtcattatccagaggatgaggtt
 K  N  D  D  G  K  K  A  L  V  L  D  C  H  Y  P  E  D  E  V
ggacaagaagacgaggctgagagtgacatcttctctatcagggaagttagctttcagtcc
 G  Q  E  D  E  A  E  S  D  I  F  S  I  R  E  V  S  F  Q  S
actggcgagtctgaatggaaggatactaattatacactcaacacagattccttggactgg
 T  G  E  S  E  W  K  D  T  N  Y  T  L  N  T  D  S  L  D  W
gccttatatgaccacctaatggatttccttgccgaccgaggggtggacaacacttttgca
 A  L  Y  D  H  L  M  D  F  L  A  D  R  G  V  D  N  T  F  A
gatgagctggtggagctcagcacagccctggagcaccaggagtacattacttttcttgaa
 D  E  L  V  E  L  S  T  A  L  E  H  Q  E  Y  I  T  F  L  E
gacctcaagagttttgtcaagagccagtag
 D  L  K  S  F  V  K  S  Q  -
```

Figure 9.3: *Sequencing data for the gC1q-R protein construct with the DNA and aa sequences aligned. The construct used consisted of residues Met74-Gln282 of gC1q-R.*

9.1.1.3 D5-1

```
atgtcccctatactaggttattggaaaattaagggccttgtgcaacccactcgacttctt
 M  S  P  I  L  G  Y  W  K  I  K  G  L  V  Q  P  T  R  L  L
ttggaatatcttgaagaaaaatatgaagagcatttgtatgagcgcgatgaaggtgataaa
 L  E  Y  L  E  E  K  Y  E  E  H  L  Y  E  R  D  E  G  D  K
tggcgaaacaaaaagtttgaattgggtttggagtttcccaatcttccttattatattgat
 W  R  N  K  K  F  E  L  G  L  E  F  P  N  L  P  Y  Y  I  D
ggtgatgttaaattaacacagtctatggccatcatacgttatatagctgacaagcacaac
 G  D  V  K  L  T  Q  S  M  A  I  I  R  Y  I  A  D  K  H  N
atgttgggtggttgtccaaaagagcgtgcagagatttcaatgcttgaaggagcggttttg
 M  L  G  G  C  P  K  E  R  A  E  I  S  M  L  E  G  A  V  L
gatattagatacggtgtttcgagaattgcatatagtaaagactttgaaactctcaaagtt
 D  I  R  Y  G  V  S  R  I  A  Y  S  K  D  F  E  T  L  K  V
acatatttaaatggtgatcatgtaacccatcctgacttcatgttgtatgacgctcttgat
 T  Y  L  N  G  D  H  V  T  H  P  D  F  M  L  Y  D  A  L  D
gttgttttatacatggacccaatgtgcctggatgcgttcccaaaattagtttgttttaaa
 V  V  L  Y  M  D  P  M  C  L  D  A  F  P  K  L  V  C  F  K
aaacgtattgaagctatcccacaaattgataagtacttgaaatccagcaagtatatagca
 K  R  I  E  A  I  P  Q  I  D  K  Y  L  K  S  S  K  Y  I  A
tggcctttgcagggctggcaagccacgtttggtggtggcgaccatcctccaaaatcggat
 W  P  L  Q  G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D
Ctggttccgcgt|ggatccactgtaagtccaccccacacttccatggcacctgcacaagat
 L  V  P  R  |  G  S  T  V  S  P  P  H  T  S  M  A  P  A  Q  D
gaagagcgggattcaggaaaagaacaagggcatactcgtagacatgactggggccatgaa
 E  E  R  D  S  G  K  E  Q  G  H  T  R  R  H  D  W  G  H  E
aaacaaagaaaacataatcttggccatggccataaacatgaacgtgaccaagggcatggg
 K  Q  R  K  H  N  L  G  H  G  H  K  H  E  R  D  Q  G  H  G
caccaaagaggacatggccttggccatggacacgaacaacagcatggtcttggtcattag
 H  Q  R  G  H  G  L  G  H  G  H  E  Q  Q  H  G  L  G  H  -
```

Figure 9.4: *Sequencing data for the D5-1 protein construct with the DNA and aa sequences aligned. The construct used consisted of residues Thr401-Gly474 of HMWK with a cleavable N-terminal GST tag, highlighted in red. The exact point of thrombin cleavage is represented with a black line.*

9.1.1.4 D5-2

```
atgtcccctatactaggttattggaaaattaagggccttgtgcaacccactcgacttctt
 M  S  P  I  L  G  Y  W  K  I  K  G  L  V  Q  P  T  R  L  L
ttggaatatcttgaagaaaaatatgaagagcatttgtatgagcgcgatgaaggtgataaa
 L  E  Y  L  E  E  K  Y  E  E  H  L  Y  E  R  D  E  G  D  K
tggcgaaacaaaaagtttgaattgggtttggagtttcccaatcttccttattatattgat
 W  R  N  K  K  F  E  L  G  L  E  F  P  N  L  P  Y  Y  I  D
ggtgatgttaaattaacacagtctatggccatcatacgttatatagctgacaagcacaac
 G  D  V  K  L  T  Q  S  M  A  I  I  R  Y  I  A  D  K  H  N
atgttgggtggttgtccaaaagagcgtgcagagatttcaatgcttgaaggagcggttttg
 M  L  G  G  C  P  K  E  R  A  E  I  S  M  L  E  G  A  V  L
gatattagatacggtgtttcgagaattgcatatagtaaagactttgaaactctcaaagtt
 D  I  R  Y  G  V  S  R  I  A  Y  S  K  D  F  E  T  L  K  V
gattttcttagcaagctacctgaaatgctgaaaatgttcgaagatcgtttatgtcataaa
 D  F  L  S  K  L  P  E  M  L  K  M  F  E  D  R  L  C  H  K
gttgttttatacatggacccaatgtgcctggatgcgttcccaaaattagtttgttttaaa
 V  V  L  Y  M  D  P  M  C  L  D  A  F  P  K  L  V  C  F  K
aaacgtattgaagctatcccacaaattgataagtacttgaaatccagcaagtatatagca
 K  R  I  E  A  I  P  Q  I  D  K  Y  L  K  S  S  K  Y  I  A
tggcctttgcagggctggcaagccacgtttggtggtggcgaccatcctccaaaatcggat
 W  P  L  Q  G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D
Ctggttccgcgt|ggatccggacataagttcaaacttgatgatgatcttgaacaccaaggg
 L  V  P  R  |  G  S  G  H  K  F  K  L  D  D  D  L  E  H  Q  G
ggccatgtccttgaccatggacataagcataagcatggtcatggccacggaaaacataaa
 G  H  V  L  D  H  G  H  K  H  K  H  G  H  G  H  G  K  H  K
aataaaggcaaaaagaatggaaagcacaatggttggaaaacagagcatttggcaagctct
 N  K  G  K  K  N  G  K  H  N  G  W  K  T  E  H  L  A  S  S
tctgaagacagttag
 S  E  D  S  -
```

Figure 9.5: *Sequencing data for the D5-2 protein construct with the DNA and aa sequences aligned. The construct used consisted of residues His475-Ser531 of HMWK with a cleavable N-terminal GST tag, highlighted in red. The exact point of thrombin cleavage is represented with a black line.*

9.1.1.5 HMWK 401-438

```
atgtcccctatactaggttattggaaaattaagggccttgtgcaacccactcgacttctt
 M  S  P  I  L  G  Y  W  K  I  K  G  L  V  Q  P  T  R  L  L
ttggaatatcttgaagaaaaatatgaagagcatttgtatgagcgcgatgaaggtgataaa
 L  E  Y  L  E  E  K  Y  E  E  H  L  Y  E  R  D  E  G  D  K
tggcgaaacaaaaagtttgaattgggtttggagtttcccaatcttccttattatattgat
 W  R  N  K  K  F  E  L  G  L  E  F  P  N  L  P  Y  Y  I  D
ggtgatgttaaattaacacagtctatggccatcatacgttatatagctgacaagcacaac
 G  D  V  K  L  T  Q  S  M  A  I  I  R  Y  I  A  D  K  H  N
atgttgggtggttgtccaaaagagcgtgcagagatttcaatgcttgaaggagcggttttg
 M  L  G  G  C  P  K  E  R  A  E  I  S  M  L  E  G  A  V  L
gatattagatacggtgtttcgagaattgcatatagtaaagactttgaaactctcaaagtt
 D  I  R  Y  G  V  S  R  I  A  Y  S  K  D  F  E  T  L  K  V
acatatttaaatggtgatcatgtaacccatcctgacttcatgttgtatgacgctcttgat
 T  Y  L  N  G  D  H  V  T  H  P  D  F  M  L  Y  D  A  L  D
gttgttttatacatggacccaatgtgcctggatgcgttcccaaaattagtttgttttaaa
 V  V  L  Y  M  D  P  M  C  L  D  A  F  P  K  L  V  C  F  K
aaacgtattgaagctatcccacaaattgataagtacttgaaatccagcaagtatatagca
 K  R  I  E  A  I  P  Q  I  D  K  Y  L  K  S  S  K  Y  I  A
tggcctttgcagggctggcaagccacgtttggtggtggcgaccatcctccaaaatcggat
 W  P  L  Q  G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D
Ctggttccgcgt|ggatccactgtaagtccaccccacacttccatggcacctgcacaagat
 L  V  P  R  |  G  S  T  V  S  P  P  H  T  S  M  A  P  A  Q  D
gaagagcgggattcaggaaaagaacaagggcatactcgtagacatgactggggccatgaa
 E  E  R  D  S  G  K  E  Q  G  H  T  R  R  H  D  W  G  H  E
aaacaaagaaaatag
 K  Q  R  K  -
```

Figure 9.6: *Sequencing data for the HMWK 401-438 protein construct with DNA and aa sequences aligned. The construct used consisted of residues Thr401-Lys438 of HMWK with a cleavable N-terminal GST tag, highlighted in red. The exact point of thrombin cleavage is represented with a black line.*

9.1.1.6 gC1q-R ΔN

```
atgattcagaagcataaaaccctccctaagatgtctggaggttgggagctggaactgaat
 M  I  Q  K  H  K  T  L  P  K  M  S  G  G  W  E  L  E  L  N
gggacagaagcgaaattagtgcggaaagttgccggggaaaaaatcacggtcactttcaac
 G  T  E  A  K  L  V  R  K  V  A  G  E  K  I  T  V  T  F  N
attaacaacagcatcccaccaacatttgatggtgaggaggaaccctcgcaagggcagaag
 I  N  N  S  I  P  P  T  F  D  G  E  E  E  P  S  Q  G  Q  K
gttgaagaacaggagcctgaactgacatcaactcccaatttcgtggttgaagttataaag
 V  E  E  Q  E  P  E  L  T  S  T  P  N  F  V  V  E  V  I  K
aatgatgatggcaagaaggcccttgtgttggactgtcattatccagaggatgaggttgga
 N  D  D  G  K  K  A  L  V  L  D  C  H  Y  P  E  D  E  V  G
caagaagacgaggctgagagtgacatcttctctatcagggaagttagctttcagtccact
 Q  E  D  E  A  E  S  D  I  F  S  I  R  E  V  S  F  Q  S  T
ggcgagtctgaatggaaggatactaattatacactcaacacagattccttggactgggcc
 G  E  S  E  W  K  D  T  N  Y  T  L  N  T  D  S  L  D  W  A
ttatatgaccacctaatggatttccttgccgaccgaggggtggacaacacttttgcagat
 L  Y  D  H  L  M  D  F  L  A  D  R  G  V  D  N  T  F  A  D
gagctggtggagctcagcacagccctggagcaccaggagtacattacttttcttgaagac
 E  L  V  E  L  S  T  A  L  E  H  Q  E  Y  I  T  F  L  E  D
ctcaagagttttgtcaagagccagtag
 L  K  S  F  V  K  S  Q  -
```

Figure 9.7: *Sequencing data for the gC1q-R ΔN mutant with the DNA and aa sequences aligned. This mutant consisted of Ile96-Gln282 of gC1q-R where residues His75-Lys95, corresponding to the N-terminal helix, were deleted from the sequence.*

9.1.1.7 gC1q-R Δβ3-β4

```
atgcacaccgacggagacaaagcttttgttgatttcctgagtgatgaaattaaggaggaa
 M  H  T  D  G  D  K  A  F  V  D  F  L  S  D  E  I  K  E  E
agaaaaattcagaagcataaaaccctccctaagatgtctggaggttgggagctggaactg
 R  K  I  Q  K  H  K  T  L  P  K  M  S  G  G  W  E  L  E  L
aatgggacagaagcgaaattagtgcggaaagttgccggggaaaaaatcacggtcactttc
 N  G  T  E  A  K  L  V  R  K  V  A  G  E  K  I  T  V  T  F
aacattaacaacagcatcccaccaacatttgatggtgctgctgctccctcgcaagggcag
 N  I  N  N  S  I  P  P  T  F  D  G  A  A  A  P  S  Q  G  Q
aaggttgctgctcaggagcctgaactgacatcaactcccaatttcgtggttgaagttata
 K  V  A  A  Q  E  P  E  L  T  S  T  P  N  F  V  V  E  V  I
aagaatgatgatggcaagaaggcccttgtgttggactgtcattatccagaggatgaggtt
 K  N  D  D  G  K  K  A  L  V  L  D  C  H  Y  P  E  D  E  V
ggacaagaagacgaggctgagagtgacatcttctctatcagggaagttagctttcagtcc
 G  Q  E  D  E  A  E  S  D  I  F  S  I  R  E  V  S  F  Q  S
actggcgagtctgaatggaaggatactaattatacactcaacacagattccttggactgg
 T  G  E  S  E  W  K  D  T  N  Y  T  L  N  T  D  S  L  D  W
gccttatatgaccacctaatggatttccttgccgaccgaggggtggacaacacttttgca
 A  L  Y  D  H  L  M  D  F  L  A  D  R  G  V  D  N  T  F  A
gatgagctggtggagctcagcacagccctggagcaccaggagtacattacttttcttgaa
 D  E  L  V  E  L  S  T  A  L  E  H  Q  E  Y  I  T  F  L  E
gacctcaagagttttgtcaagagccagtag
 D  L  K  S  F  V  K  S  Q  -
```

Figure 9.8: *Sequencing data for the gC1q-R Δβ3-β4 mutant with the DNA and aa sequences aligned. The region mutated is highlighted in red and corresponds to Glu146-148 and Glu156-157 substituted for Ala residues.*

9.1.1.8 gC1q-R Δβ5-β6

```
atgcacaccgacggagacaaagcttttgttgatttcctgagtgatgaaattaaggaggaa
 M  H  T  D  G  D  K  A  F  V  D  F  L  S  D  E  I  K  E  E
agaaaaattcagaagcataaaaccctccctaagatgtctggaggttgggagctggaactg
 R  K  I  Q  K  H  K  T  L  P  K  M  S  G  G  W  E  L  E  L
aatgggacagaagcgaaattagtgcggaaagttgccggggaaaaaatcacggtcactttc
 N  G  T  E  A  K  L  V  R  K  V  A  G  E  K  I  T  V  T  F
aacattaacaacagcatcccaccaacatttgatggtgaggaggaaccctcgcaagggcag
 N  I  N  N  S  I  P  P  T  F  D  G  E  E  E  P  S  Q  G  Q
aaggttgaagaacaggagcctgaactgacatcaactcccaatttcgtggttgaagttata
 K  V  E  E  Q  E  P  E  L  T  S  T  P  N  F  V  V  E  V  I
aagaatgatgatggcaagaaggcccttgtgttggactgtcattatccagaggatgaggtt
 K  N  D  D  G  K  K  A  L  V  L  D  C  H  Y  P  E  D  E  V
ggacaaagtgacatcttctctatcagggaagttagctttcagtccactggcgagtctgaa
 G  Q  S  D  I  F  S  I  R  E  V  S  F  Q  S  T  G  E  S  E
tggaaggatactaattatacactcaacacagattccttggactgggccttatatgaccac
 W  K  D  T  N  Y  T  L  N  T  D  S  L  D  W  A  L  Y  D  H
ctaatggatttccttgccgaccgaggggtggacaacacttttgcagatgagctggtggag
 L  M  D  F  L  A  D  R  G  V  D  N  T  F  A  D  E  L  V  E
ctcagcacagccctggagcaccaggagtacattacttttcttgaagacctcaagagtttt
 L  S  T  A  L  E  H  Q  E  Y  I  T  F  L  E  D  L  K  S  F
gtcaagagccagtag
 V  K  S  Q  -
```

Figure 9.9: *Sequencing data for the gC1q-R Δβ5-β6 mutant with the DNA and aa sequences aligned. This mutant had residues Glu196-Glu200 deleted from the sequence.*

9.1.1.9 gC1q-R Δβ6-β7

```
atgcacaccgacggagacaaagcttttgttgatttcctgagtgatgaaattaaggaggaa
 M  H  T  D  G  D  K  A  F  V  D  F  L  S  D  E  I  K  E  E
agaaaaattcagaagcataaaaccctccctaagatgtctggaggttgggagctggaactg
 R  K  I  Q  K  H  K  T  L  P  K  M  S  G  G  W  E  L  E  L
aatgggacagaagcgaaattagtgcggaaagttgccggggaaaaaatcacggtcactttc
 N  G  T  E  A  K  L  V  R  K  V  A  G  E  K  I  T  V  T  F
aacattaacaacagcatcccaccaacatttgatggtgaggaggaaccctcgcaagggcag
 N  I  N  N  S  I  P  P  T  F  D  G  E  E  E  P  S  Q  G  Q
aaggttgaagaacaggagcctgaactgacatcaactcccaatttcgtggttgaagttata
 K  V  E  E  Q  E  P  E  L  T  S  T  P  N  F  V  V  E  V  I
aagaatgatgatggcaagaaggcccttgtgttggactgtcattatccagaggatgaggtt
 K  N  D  D  G  K  K  A  L  V  L  D  C  H  Y  P  E  D  E  V
ggacaagaagacgaggctgagagtgacatcttctctatcagggaagttagctttcagtcc
 G  Q  E  D  E  A  E  S  D  I  F  S  I  R  E  V  S  F  Q  S
acactcaacacagattccttggactgggccttatatgaccacctaatggatttccttgcc
 T  L  N  T  D  S  L  D  W  A  L  Y  D  H  L  M  D  F  L  A
gaccgaggggtggacaacacttttgcagatgagctggtggagctcagcacagccctggag
 D  R  G  V  D  N  T  F  A  D  E  L  V  E  L  S  T  A  L  E
caccaggagtacattacttttcttgaagacctcaagagttttgtcaagagccagtag
 H  Q  E  Y  I  T  F  L  E  D  L  K  S  F  V  K  S  Q  -
```

Figure 9.10: *Sequencing data for the gC1q-R Δβ6-β7 mutant with the DNA and aa sequences aligned. This mutant had residues Thr214-Tyr224 deleted from the sequence.*

9.2 Appendix 2

9.2.1 Plasmid Maps

All plasmid maps were made using SnapGene viewer software (http://www.snapgene.com/products/snapgene_viewer/).

9.2.1.1 pET-28a D5

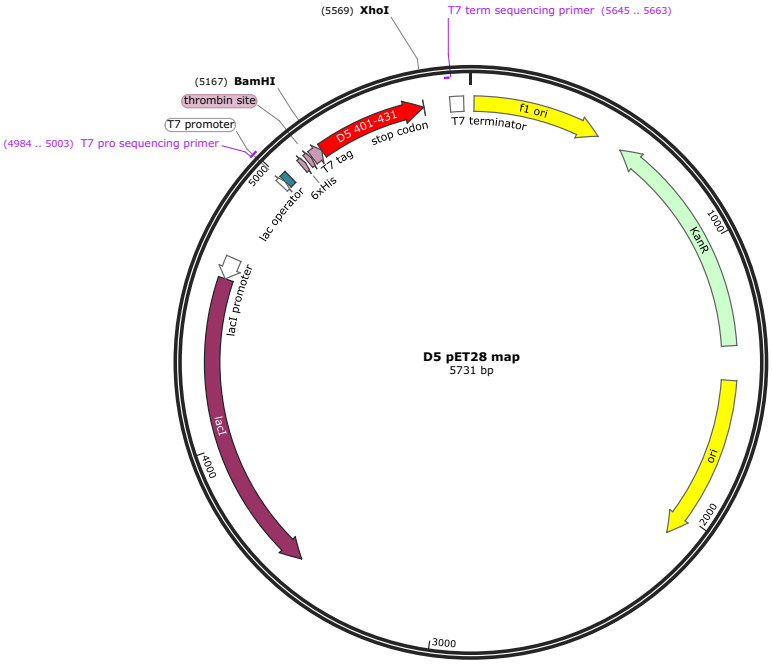


Figure 9.11: *Plasmid map for the D5 pET28-a plasmid construct showing the inserted D5 gene and other major features.*

9.2.1.2 pGEX 4T1 D5-1

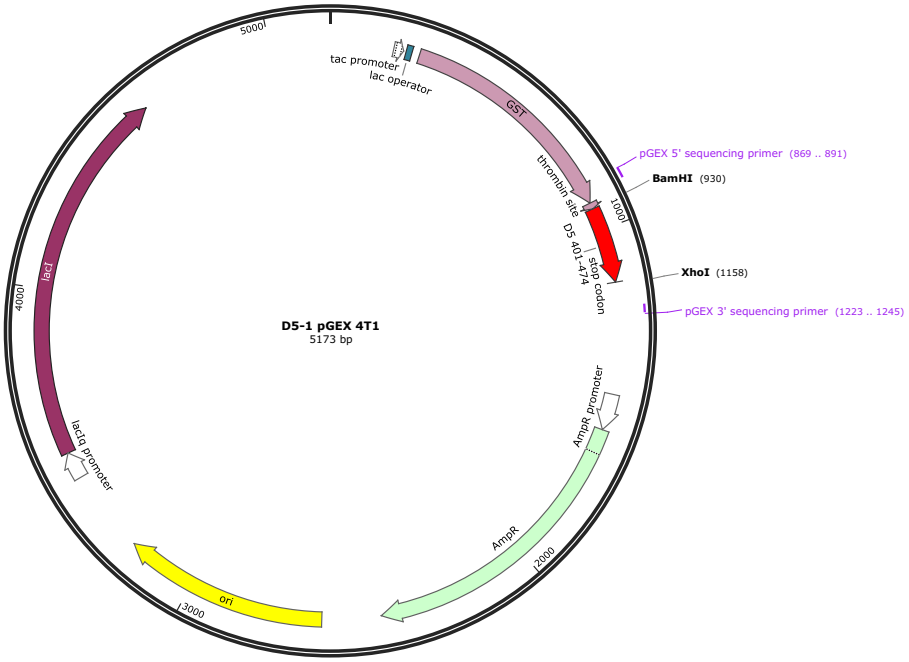


Figure 9.12: *Plasmid map for the D5-1 pGEX 4T1 plasmid construct showing the inserted D5-1 gene and other major features.*

9.2.1.3 pGEX 4T1 D5-2

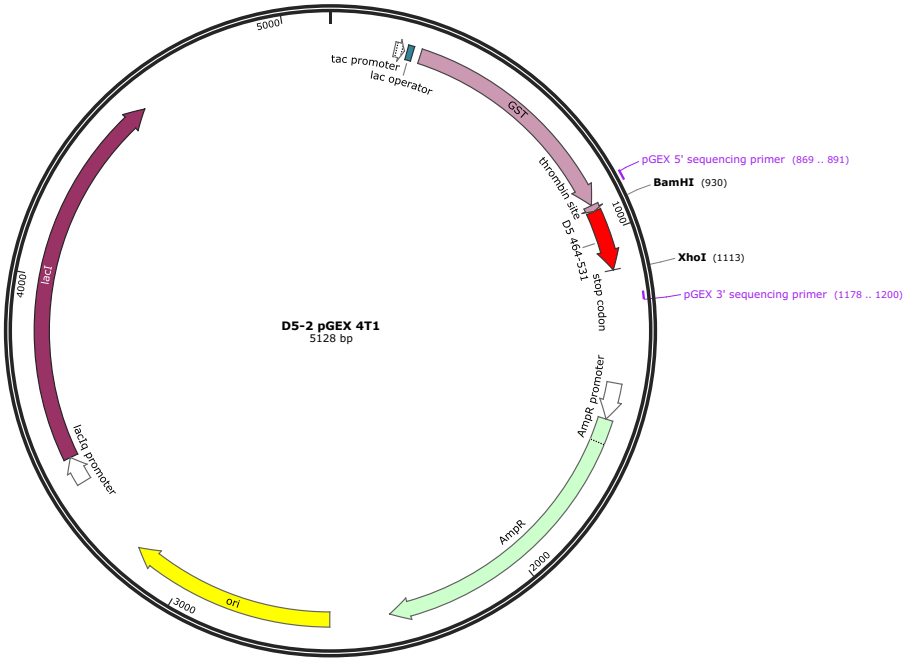


Figure 9.13: *Plasmid map for the D5-2 pGEX 4T1 plasmid construct showing the inserted D5-1 gene and other major features.*

9.2.1.4 pGEX 4T1 HMWK 401-438

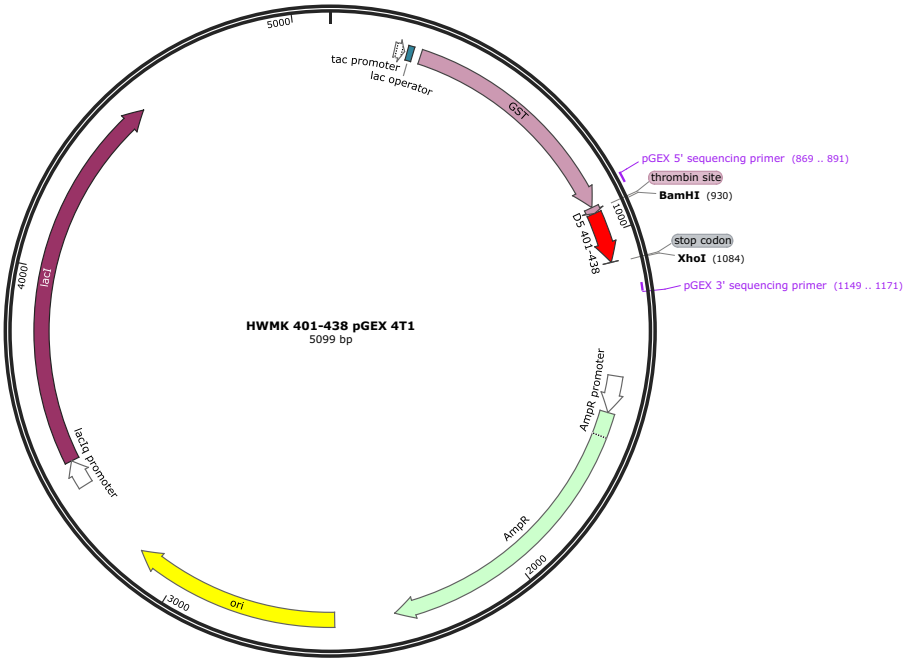


Figure 9.14: *Plasmid map for the HMWK 401-438 pGEX 4T1 plasmid construct showing the inserted D5-1 gene and other major features.*

References

[1] J. De Las Rivas and C. Fontanillo. Protein-protein interactions essentials: key concepts to building and analyzing interactome networks. *PLoS Comput Biol*, 6(6):e1000807, Jun 2010.

[2] M. W. Gonzalez and M. G. Kann. Chapter 4: Protein interactions and disease. *PLoS Comput Biol*, 8(12):e1002819, 2012.

[3] A. J. Gale. Continuing education course 2: current understanding of hemostasis. *Toxicol Pathol*, 39(1):273–80, Jan 2011.

[4] K. S. Sakariassen, P. A. Bolhuis, and J. J. Sixma. Human blood platelet adhesion to artery subendothelium is mediated by factor VIII-Von Willebrand factor bound to the subendothelium. *Nature*, 279(5714):636–8, Jun 1979.

[5] C. M. Heldebrant, R. J. Butkowski, S. P. Bajaj, and K. G. Mann. The activation of prothrombin. II. Partial reactions, physical and chemical characterization of the intermediates of activation. *J Biol Chem*, 248(20):7149–63, Oct 1973.

[6] B. Blombäck, B. Hessel, D. Hogg, and L. Therkildsen. A two-step fibrinogen–fibrin transition in blood coagulation. *Nature*, 275(5680):501–5, Oct 1978.

[7] J. W. Weisel and R. I. Litvinov. Mechanisms of fibrin polymerization and clinical implications. *Blood*, 121(10):1712–9, Mar 2013.

[8] E. W. Davie. A brief historical review of the waterfall/cascade of blood coagulation. *J Biol Chem*, 278(51):50819–32, Dec 2003.

[9] R. W. SHERMER, R. G. MASON, R. H. WAGNER, and K. M. BRINKHOUS. Studies on thrombin-induced platelet agglutination. *J Exp Med*, 114:905–20, Dec 1961.

[10] V. Sreeramkumar, J. M. Adrover, I. Ballesteros, M. I. Cuartero, J. Rossaint, I. Bilbao, M. Nácher, C. Pitaval, I. Radovanovic, Y. Fukui, R. P. McEver, M.-D. Filippi, I. Lizasoain, J. Ruiz-Cabello, A. Zarbock, M. A. Moro, and A. Hidalgo. Neutrophils scan for activated platelets to initiate inflammation. *Science*, 346(6214):1234–8, Dec 2014.

[11] P. H. Kuijper, H. I. Gallardo Torres, J. A. van der Linden, J. W. Lammers, J. J. Sixma, L. Koenderman, and J. J. Zwaginga. Platelet-dependent primary hemostasis promotes selectin- and integrin-mediated neutrophil adhesion to damaged endothelium under flow conditions. *Blood*, 87(8): 3271–81, Apr 1996.

[12] P. H. Kuijper, H. I. Gallardo Torres, J. W. Lammers, J. J. Sixma, L. Koenderman, and J. J. Zwaginga. Platelet and fibrin deposition at the damaged vessel wall: cooperative substrates for neutrophil adhesion under flow conditions. *Blood*, 89(1):166–75, Jan 1997.

[13] B. Furie and B. C. Furie. In vivo thrombus formation. *J Thromb Haemost*, 5 Suppl 1:12–7, Jul 2007.

[14] E. Kesieme, C. Kesieme, N. Jebbin, E. Irekpita, and A. Dongo. Deep vein thrombosis: a clinical review. *J Blood Med*, 2:59–69, 2011.

[15] S. Z. Goldhaber and H. Bounameaux. Pulmonary embolism and deep vein thrombosis. *Lancet*, 379(9828):1835–46, May 2012.

[16] S. Parikh, R. Shah, and P. Kapoor. Portal vein thrombosis. *Am J Med*, 123 (2):111–9, Feb 2010.

[17] M. J. O'Dea, R. S. Malek, R. M. Tucker, and R. E. Fulton. Renal vein thrombosis. *J Urol*, 116(4):410–4, Oct 1976.

[18] E. Previtali, P. Bucciarelli, S. M. Passamonti, and I. Martinelli. Risk factors for venous and arterial thrombosis. *Blood Transfus*, 9(2):120–38, Apr 2011.

[19] S. Palta, R. Saroa, and A. Palta. Overview of the coagulation system. *Indian J Anaesth*, 58(5):515–23, Sep 2014.

[20] G. L. Peters, P. B. Erwin, M. K. Pitlick, and Z. A. Stacy. Utilization of coagulation assays in clinical therapeutics. *Pharmacotherapy*, 33(11):1214–22, Nov 2013.

[21] A. D. Schecter, B. Spirn, M. Rossikhina, P. L. Giesen, V. Bogdanov, J. T. Fallon, E. A. Fisher, L. M. Schnapp, Y. Nemerson, and M. B. Taubman. Release of active tissue factor by human arterial smooth muscle cells. *Circ Res*, 87(2):126–32, Jul 2000.

[22] R. A. Fleck, L. V. Rao, S. I. Rapaport, and N. Varki. Localization of human tissue factor antigen by immunostaining with monospecific, polyclonal anti-human tissue factor antibody. *Thromb Res*, 59(2):421–37, Jul 1990.

[23] B. A. Bouchard, M. A. Shatos, and P. B. Tracy. Human brain pericytes differentially regulate expression of procoagulant enzyme complexes comprising the extrinsic pathway of blood coagulation. *Arterioscler Thromb Vasc Biol*, 17(1):1–9, Jan 1997.

[24] T. A. Drake, J. H. Morrissey, and T. S. Edgington. Selective cellular expression of tissue factor in human tissues. Implications for disorders of hemostasis and thrombosis. *Am J Pathol*, 134(5):1087–97, May 1989.

[25] H. ten Cate, K. A. Bauer, M. Levi, T. S. Edgington, R. D. Sublett, S. Barzegar, B. L. Kass, and R. D. Rosenberg. The activation of factor X and prothrombin by recombinant factor VIIa in vivo is mediated by tissue factor. *J Clin Invest*, 92(3):1207–12, Sep 1993.

[26] L. V. Rao and S. I. Rapaport. Activation of factor VII bound to tissue factor: a key early step in the tissue factor pathway of blood coagulation. *Proc Natl Acad Sci U S A*, 85(18):6687–91, Sep 1988.

[27] M. Schapira, E. Despland, C. F. Scott, L. A. Boxer, and R. W. Colman. Purified human plasma kallikrein aggregates human blood neutrophils. *J Clin Invest*, 69(5):1199–202, May 1982.

[28] G. H. Goldsmith, H. Saito, and O. S. Ratnoff. The activation of plasminogen by Hageman factor (Factor XII) and Hageman factor fragments. *J. Clin. Invest.*, 62(1):54–60, Jul 1978.

[29] S. Selvarajan, L. R. Lund, T. Takeuchi, C. S. Craik, and Z. Werb. A plasma kallikrein-dependent plasminogen cascade required for adipocyte differentiation. *Nat Cell Biol*, 3(3):267–75, Mar 2001.

[30] C. G. Cochrane, S. D. Revak, and K. D. Wuepper. Activation of Hageman factor in solid and fluid phases. A critical role of kallikrein. *J Exp Med*, 138 (6):1564–83, Dec 1973.

[31] Y. Wu. Contact pathway of coagulation and inflammation. *Thromb J*, 13: 17, 2015.

[32] A. T. Long, E. Kenne, R. Jung, T. A. Fuchs, and T. Renné. Contact system revisited: an interface between inflammation, coagulation, and innate immunity. *J Thromb Haemost*, 14(3):427–37, Mar 2016.

[33] R. W. Colman. Surface-mediated defense reactions. The plasma contact activation system. *J Clin Invest*, 73(5):1249–53, May 1984.

[34] O. D. Ratnoff and H. Saito. Surface-mediated reactions. *Curr Top Hematol*, 2:1–57, 1979.

[35] C. Naudin, E. Burillo, S. Blankenberg, L. Butler, and T. Renné. Factor XII Contact Activation. *Semin Thromb Hemost*, Mar 2017.

[36] R. E. Thompson, R. Mandle, Jr, and A. P. Kaplan. Association of factor XI and high molecular weight kininogen in human plasma. *J Clin Invest*, 60 (6):1376–80, Dec 1977.

[37] H. L. Meier, J. V. Pierce, R. W. Colman, and A. P. Kaplan. Activation and function of human Hageman factor. The role of high molecular weight kininogen and prekallikrein. *J Clin Invest*, 60(1):18–31, Jul 1977.

[38] S. Reddigari and A. P. Kaplan. Cleavage of human high-molecular weight kininogen by purified kallikreins and upon contact activation of plasma. *Blood*, 71(5):1334–40, May 1988.

[39] R. E. Thompson, R. Mandle, Jr, and A. P. Kaplan. Characterization of human high molecular weight kininogen. Procoagulant activity associated with the light chain of kinin-free high molecular weight kininogen. *J Exp Med*, 147(2):488–99, Feb 1978.

[40] C. F. Scott, L. D. Silver, M. Schapira, and R. W. Colman. Cleavage of human high molecular weight kininogen markedly enhances its coagulant activity. Evidence that this molecule exists as a procofactor. *J Clin Invest*, 73(4):954–62, Apr 1984.

[41] S. de Maat, P. G. de Groot, and C. Maas. Contact system activation on endothelial cells. *Semin Thromb Hemost*, 40(8):887–94, Nov 2014.

[42] L. M. Henderson, C. D. Figueroa, W. Müller-Esterl, and K. D. Bhoola. Assembly of contact-phase factors on the surface of the human neutrophil membrane. *Blood*, 84(2):474–82, Jul 1994.

[43] A. Ben Nasr, A. Olsén, U. Sjöbring, W. Müller-Esterl, and L. Björck. Assembly of human contact phase proteins and release of bradykinin at the surface of curli-expressing Escherichia coli. *Mol Microbiol*, 20(5):927–35, Jun 1996.

[44] A. Ben Nasr, H. Herwald, U. Sjöbring, T. Renné, W. Müller-Esterl, and L. Björck. Absorption of kininogen from human plasma by Streptococcus

pyogenes is followed by the release of bradykinin. *Biochem J*, 326 (Pt 3): 657–60, Sep 1997.

[45] E. S. Gershom, M. R. Sutherland, P. Lollar, and E. L. G. Pryzdial. Involvement of the contact phase and intrinsic pathway in herpes simplex virus-initiated plasma coagulation. *J Thromb Haemost*, 8(5):1037–43, May 2010.

[46] F. Mahdi, Z. S. Madar, C. D. Figueroa, and A. H. Schmaier. Factor XII interacts with the multiprotein assembly of urokinase plasminogen activator receptor, gC1qR, and cytokeratin 1 on endothelial cell membranes. *Blood*, 99(10):3585–96, May 2002.

[47] J. W. Yau, H. Teoh, and S. Verma. Endothelial cell control of thrombosis. *BMC Cardiovasc Disord*, 15:130, Oct 2015.

[48] K. Joseph, Y. Shibayama, B. Ghebrehiwet, and A. P. Kaplan. Factor XII-dependent contact activation on endothelial cells and binding proteins gC1qR and cytokeratin 1. *Thromb Haemost*, 85(1):119–24, Jan 2001.

[49] J. H. Morrissey, S. H. Choi, and S. A. Smith. Polyphosphate: an ancient molecule that links platelets, coagulation, and inflammation. *Blood*, 119 (25):5972–9, Jun 2012.

[50] T. J. Gould, T. T. Vu, L. L. Swystun, D. J. Dwivedi, S. H. C. Mai, J. I. Weitz, and P. C. Liaw. Neutrophil extracellular traps promote thrombin generation through platelet-dependent and platelet-independent mechanisms. *Arterioscler Thromb Vasc Biol*, 34(9):1977–84, Sep 2014.

[51] D. Noubouossie, M. F. Whelihan, D. M. Monroe, and N. S. Key. Purified DNA, but Not Neutrophil Extracellular Traps (NETs), Promotes Contact Activation of Coagulation. *Blood*, 126(23):767–767, 2015.

[52] A. H. Schmaier. The elusive physiologic role of Factor XII. *J Clin Invest*, 118(9):3006–9, Sep 2008.

[53] C. Oschatz, C. Maas, B. Lecher, T. Jansen, J. Björkqvist, T. Tradler, R. Sedlmeier, P. Burfeind, S. Cichon, S. Hammerschmidt, W. Müller-Esterl, W. A. Wuillemin, G. Nilsson, and T. Renné. Mast cells increase vascular permeability by heparin-initiated bradykinin formation in vivo. *Immunity*, 34(2):258–68, Feb 2011.

[54] P. E. J. van der Meijden, I. C. A. Munnix, J. M. Auger, J. W. P. Govers-Riemslag, J. M. E. M. Cosemans, M. J. E. Kuijpers, H. M. Spronk, S. P. Watson, T. Renné, and J. W. M. Heemskerk. Dual role of collagen in factor XII-dependent thrombus formation. *Blood*, 114(4):881–90, Jul 2009.

[55] P. L. F. Giangrande. Six characters in search of an author: the history of the nomenclature of coagulation factors. *Br J Haematol*, 121(5):703–12, Jun 2003.

[56] D. J. Campbell. The kallikrein-kinin system in humans. *Clin Exp Pharmacol Physiol*, 28(12):1060–5, Dec 2001.

[57] A. Gruber. The role of the contact pathway in thrombus propagation. *Thromb Res*, 133 Suppl 1:S45–7, May 2014.

[58] T. Renné, M. Pozgajová, S. Grüner, K. Schuh, H.-U. Pauer, P. Burfeind, D. Gailani, and B. Nieswandt. Defective thrombus formation in mice lacking coagulation factor XII. *J Exp Med*, 202(2):271–81, Jul 2005.

[59] J. E. Bird, P. L. Smith, X. Wang, W. A. Schumacher, F. Barbera, J.-P. Revelli, and D. Seiffert. Effects of plasma kallikrein deficiency on haemostasis and thrombosis in mice: murine ortholog of the Fletcher trait. *Thromb Haemost*, 107(6):1141–50, Jun 2012.

[60] A. S. Revenko, D. Gao, J. R. Crosby, G. Bhattacharjee, C. Zhao, C. May, D. Gailani, B. P. Monia, and A. R. MacLeod. Selective depletion of plasma prekallikrein or coagulation factor XII inhibits thrombosis in mice without increased risk of bleeding. *Blood*, 118(19):5302–11, Nov 2011.

[61] H. Zhang, E. C. Löwenberg, J. R. Crosby, A. R. MacLeod, C. Zhao, D. Gao, C. Black, A. S. Revenko, J. C. M. Meijers, E. S. Stroes, M. Levi, and B. P. Monia. Inhibition of the intrinsic coagulation pathway factor XI by antisense oligonucleotides: a novel antithrombotic strategy with lowered bleeding risk. *Blood*, 116(22):4684–92, Nov 2010.

[62] D. Gailani and T. Renné. Intrinsic pathway of coagulation and arterial thrombosis. *Arterioscler Thromb Vasc Biol*, 27(12):2507–13, Dec 2007.

[63] N. Kitamura, Y. Takagaki, S. Furuto, T. Tanaka, H. Nawa, and S. Nakanishi. A single gene for bovine high molecular weight and low molecular weight kininogens. *Nature*, 305(5934):545–9, 1983.

[64] D. M. Kerbiriou, B. N. Bouma, and J. H. Griffin. Immunochemical studies of human high molecular weight kininogen and of its complexes with plasma prekallikrein or kallikrein. *J Biol Chem*, 255(9):3952–8, May 1980.

[65] R. J. Mandle, R. W. Colman, and A. P. Kaplan. Identification of prekallikrein and high-molecular-weight kininogen as a complex in human plasma. *Proc Natl Acad Sci U S A*, 73(11):4179–83, Nov 1976.

[66] R. E. Thompson, R. Mandle, Jr, and A. P. Kaplan. Studies of binding of prekallikrein and Factor XI to high molecular weight kininogen and its light chain. *Proc Natl Acad Sci U S A*, 76(10):4862–6, Oct 1979.

[67] J. W. Weisel, C. Nagaswami, J. L. Woodhead, R. A. DeLa Cadena, J. D. Page, and R. W. Colman. The shape of high molecular weight kininogen. Organization into structural domains, changes with activation, and interactions with prekallikrein, as determined by electron microscopy. *J Biol Chem*, 269(13):10100–6, Apr 1994.

[68] M. Rocha E Silva, W. T. Berlado, and G. Rosenfeld. Bradykinin, a hypotensive and smooth muscle stimulating factor released from plasma globulin by snake venoms and by trypsin. *Am J Physiol*, 156(2):261–73, Feb 1949.

[69] S. P. Kunapuli, R. A. DeLa Cadena, and R. W. Colman. Bacterial expression of biologically active high molecular weight kininogen light chain. *Thromb Haemost*, 67(4):428–33, Apr 1992.

[70] S. P. Kunapuli, R. A. DeLa Cadena, and R. W. Colman. Deletion mutagenesis of high molecular weight kininogen light chain. Identification of two anionic surface binding subdomains. *J Biol Chem*, 268(4):2486–92, Feb 1993.

[71] K. Mori and S. Nagasawa. Studies on human high molecular weight (HMW) kininogen. II. Structural change of HMW kininogen by the action of human plasma kallikrein. *J Biochem*, 89(5):1465–73, May 1981.

[72] R. W. Colman. Structural biology of high molecular weight kininogen. *Pure and Appl. Chem Appl. Chem*, 66(1):27–34, 1994.

[73] C. F. Scott, L. D. Silver, A. D. Purdon, and R. W. Colman. Cleavage of human high molecular weight kininogen by factor XIa in vitro. Effect on structure and function. *J Biol Chem*, 260(19):10856–63, Sep 1985.

[74] H. N. Bradford, B. A. Jameson, A. A. Adam, R. P. Wassell, and R. W. Colman. Contiguous binding and inhibitory sites on kininogens required for the inhibition of platelet calpain. *J Biol Chem*, 268(35):26546–51, Dec 1993.

[75] H. Herwald, A. A. Hasan, J. Godovac-Zimmermann, A. H. Schmaier, and W. Müller-Esterl. Identification of an endothelial cell binding site on kininogen domain D3. *J Biol Chem*, 270(24):14634–42, Jun 1995.

[76] Y. T. Wachtfogel, R. A. DeLa Cadena, S. P. Kunapuli, L. Rick, M. Miller, R. L. Schultze, D. C. Altieri, T. S. Edgington, and R. W. Colman. High molecular weight kininogen binds to Mac-1 on neutrophils by its heavy chain (domain 3) and its light chain (domain 5). *J Biol Chem*, 269(30): 19307–12, Jul 1994.

[77] F. J. Meloni and A. H. Schmaier. Low molecular weight kininogen binds

to platelets to modulate thrombin-induced platelet activation. *J Biol Chem*, 266(11):6786–94, Apr 1991.

[78] A. A. Hasan, D. B. Cines, H. Herwald, A. H. Schmaier, and W. Müller-Esterl. Mapping the cell binding site on high molecular weight kininogen domain 5. *J Biol Chem*, 270(33):19256–61, Aug 1995.

[79] D. F. Noubouossie, M. F. Whelihan, Y.-B. Yu, E. Sparkenbaugh, R. Pawlinski, D. M. Monroe, and N. S. Key. In vitro activation of coagulation by human neutrophil DNA and histone proteins but not neutrophil extracellular traps. *Blood*, 129(8):1021–1029, Feb 2017.

[80] J. M. Gajsiewicz, S. A. Smith, and J. H. Morrissey. Polyphosphate and RNA Differentially Modulate the Contact Pathway of Blood Clotting. *J Biol Chem*, 292(5):1808–1814, Feb 2017.

[81] F. van der Graaf, F. J. Keus, R. A. Vlooswijk, and B. N. Bouma. The contact activation mechanism in human plasma: activation induced by dextran sulfate. *Blood*, 59(6):1225–33, Jun 1982.

[82] M. Siebeck, J. C. Cheronis, E. Fink, J. Kohl, B. Spies, M. Spannagl, M. Jochum, and H. Fritz. Dextran sulfate activates contact system and mediates arterial hypotension via B2 kinin receptors. *J Appl Physiol (1985)*, 77(6):2675–80, Dec 1994.

[83] S. Zhu and S. L. Diamond. Contact activation of blood coagulation on a defined kaolin/collagen surface in a microfluidic assay. *Thromb Res*, 134 (6):1335–43, Dec 2014.

[84] R. A. DeLa Cadena and R. W. Colman. The sequence HGLGHGHEQQHGLGHGH in the light chain of high molecular weight kininogen serves as a primary structural feature for zinc-dependent binding to an anionic surface. *Protein Sci*, 1(1):151–60, Jan 1992.

[85] J. S. Greengard and J. H. Griffin. Receptors for high molecular weight

kininogen on stimulated washed human platelets. *Biochemistry*, 23(26): 6863–9, Dec 1984.

[86] J. M. Zini, A. H. Schmaier, and D. B. Cines. Bradykinin regulates the expression of kininogen binding sites on endothelial cells. *Blood*, 81(11): 2936–46, Jun 1993.

[87] R. Røjkaer and I. Schousboe. Partial identification of the Zn2+-binding sites in factor XII and its activation derivatives. *Eur J Biochem*, 247(2): 491–6, Jul 1997.

[88] J. Howl and S. J. Payne. Bradykinin receptors as a therapeutic target. *Expert Opin Ther Targets*, 7(2):277–85, Apr 2003.

[89] K. Ishihara, M. Kamata, I. Hayashi, S. Yamashina, and M. Majima. Roles of bradykinin in vascular permeability and angiogenesis in solid tumor. *Int Immunopharmacol*, 2(4):499–509, Mar 2002.

[90] A. Dray and M. Perkins. Bradykinin and inflammatory pain. *Trends Neurosci*, 16(3):99–104, Mar 1993.

[91] E. Sato, S. Koyama, H. Nomura, K. Kubo, and M. Sekiguchi. Bradykinin stimulates alveolar macrophages to release neutrophil, monocyte, and eosinophil chemotactic activity. *J Immunol*, 157(7):3122–9, Oct 1996.

[92] I. Paegelow, S. Trzeczak, S. Böckmann, and G. Vietinghoff. Migratory responses of polymorphonuclear leukocytes to kinin peptides. *Pharmacology*, 66(3):153–61, Nov 2002.

[93] L. M. F. Leeb-Lundberg, F. Marceau, W. Müller-Esterl, D. J. Pettibone, and B. L. Zuraw. International union of pharmacology. XLV. Classification of the kinin receptor family: from molecular mechanisms to pathophysiological consequences. *Pharmacol Rev*, 57(1):27–77, Mar 2005.

[94] A. Ahluwalia and M. Perretti. B1 receptors as a new inflammatory target. Could this B the 1? *Trends Pharmacol Sci*, 20(3):100–4, Mar 1999.

[95] C. D. Figueroa, P. Ehrenfeld, and K. D. Bhoola. Kinin receptors as targets for cancer therapy. *Expert Opin Ther Targets*, 16(3):299–312, Mar 2012.

[96] I.-M. Frick, L. Björck, and H. Herwald. The dual role of the contact system in bacterial infectious disease. *Thromb Haemost*, 98(3):497–502, Sep 2007.

[97] A. J. Chu. Blood coagulation as an intrinsic pathway for proinflammation: a mini review. *Inflamm Allergy Drug Targets*, 9(1):32–44, Mar 2010.

[98] J. F. Tait and K. Fujikawa. Identification of the binding site for plasma prekallikrein in human high molecular weight kininogen. A region from residues 185 to 224 of the kininogen light chain retains full binding activity. *J Biol Chem*, 261(33):15396–401, Nov 1986.

[99] E. Hooley, P. A. McEwan, and J. Emsley. Molecular modeling of the prekallikrein structure provides insights into high-molecular-weight kininogen binding and zymogen activation. *J Thromb Haemost*, 5(12): 2461–6, Dec 2007.

[100] T. Renné, D. Gailani, C. M. Meijers, J, and W. Müller-Esterl. Characterization of the H-kininogen-binding site on factor XI: a comparison of factor XI and plasma prekallikrein. *J Biol Chem*, 277(7):4892–9, Feb 2002.

[101] M. Delvaeye and E. M. Conway. Coagulation and innate immune responses: can we view them separately? *Blood*, 114(12):2367–74, Sep 2009.

[102] E. J. Gustafson, A. H. Schmaier, Y. T. Wachtfogel, N. Kaufman, U. Kucich, and R. W. Colman. Human neutrophils contain and bind high molecular weight kininogen. *J Clin Invest*, 84(1):28–35, Jul 1989.

[103] L. P. Fernando, S. Natesan, K. Joseph, and A. P. Kaplan. High molecular weight kininogen and factor XII binding to endothelial cells and astrocytes. *Thromb Haemost*, 90(5):787–95, Nov 2003.

[104] R. W. Colman, B. A. Jameson, Y. Lin, D. Johnson, and S. A. Mousa. Domain 5 of high molecular weight kininogen (kininostatin) down-regulates

endothelial cell proliferation and migration and inhibits angiogenesis. *Blood*, 95(2):543–50, Jan 2000.

[105] J. C. Zhang, F. Donate, X. Qi, N. P. Ziats, J. C. Juarez, A. P. Mazar, Y.-P. Pang, and K. R. McCrae. The antiangiogenic activity of cleaved high molecular weight kininogen is mediated through binding to endothelial cell tropomyosin. *Proc Natl Acad Sci U S A*, 99(19):12224–9, Sep 2002.

[106] K. Joseph, B. Ghebrehiwet, E. I. Peerschke, K. B. Reid, and A. P. Kaplan. Identification of the zinc-dependent endothelial cell binding protein for high molecular weight kininogen and factor XII: identity with the receptor that binds to the globular "heads" of C1q (gC1q-R). *Proc Natl Acad Sci U S A*, 93(16):8552–7, Aug 1996.

[107] H. Herwald, J. Dedio, R. Kellner, M. Loos, and W. Müller-Esterl. Isolation and characterization of the kininogen-binding protein p33 from endothelial cells. Identity with the gC1q receptor. *J Biol Chem*, 271(22):13040–7, May 1996.

[108] R. W. Colman, R. A. Pixley, S. Najamunnisa, W. Yan, J. Wang, A. Mazar, and K. R. McCrae. Binding of high molecular weight kininogen to human endothelial cells is mediated via a site within domains 2 and 3 of the urokinase receptor. *J Clin Invest*, 100(6):1481–7, Sep 1997.

[109] A. Estreicher, J. Mühlhauser, J. L. Carpentier, L. Orci, and J. D. Vassalli. The receptor for urokinase type plasminogen activator polarizes expression of the protease to the leading edge of migrating monocytes and promotes degradation of enzyme inhibitor complexes. *J Cell Biol*, 111(2):783–92, Aug 1990.

[110] V. Ellis and K. Danø. Plasminogen activation by receptor-bound urokinase. *Semin Thromb Hemost*, 17(3):194–200, Jul 1991.

[111] E. Bacharach, A. Itin, and E. Keshet. In vivo patterns of expression of urokinase and its inhibitor PAI-1 suggest a concerted role in regulating

physiological angiogenesis. *Proc Natl Acad Sci U S A*, 89(22):10686–90, Nov 1992.

[112] J. Pöllänen, R. W. Stephens, and A. Vaheri. Directed plasminogen activation at the surface of normal and malignant cells. *Adv Cancer Res*, 57: 273–328, 1991.

[113] A. Nykjaer, B. Møller, R. F. Todd, 3rd, T. Christensen, P. A. Andreasen, J. Gliemann, and C. M. Petersen. Urokinase receptor. An activation antigen in human T lymphocytes. *J Immunol*, 152(2):505–16, Jan 1994.

[114] Y. Liu, D. J. Cao, I. M. Sainz, Y.-L. Guo, and R. W. Colman. The inhibitory effect of HKa in endothelial cell tube formation is mediated by disrupting the uPA-uPAR complex and inhibiting its signaling and internalization. *Am J Physiol Cell Physiol*, 295(1):C257–67, Jul 2008.

[115] F. Mahdi, Z. Shariat-Madar, A. Kuo, M. Carinato, D. B. Cines, and A. H. Schmaier. Mapping the interaction between high molecular mass kininogen and the urokinase plasminogen activator receptor. *J Biol Chem*, 279 (16):16621–8, Apr 2004.

[116] A. A. Hasan, T. Zisman, and A. H. Schmaier. Identification of cytokeratin 1 as a binding protein and presentation receptor for kininogens on endothelial cells. *Proc Natl Acad Sci U S A*, 95(7):3615–20, Mar 1998.

[117] R. A. Pixley, R. G. Espinola, B. Ghebrehiwet, K. Joseph, A. Kao, K. Bdeir, D. B. Cines, and R. W. Colman. Interaction of high-molecular-weight kininogen with endothelial cell binding proteins suPAR, gC1qR and cytokeratin 1 determined by surface plasmon resonance (BiaCore). *Thromb Haemost*, 105(6):1053–9, Jun 2011.

[118] K. Joseph, B. Ghebrehiwet, and A. P. Kaplan. Cytokeratin 1 and gC1qR mediate high molecular weight kininogen binding to endothelial cells. *Clin Immunol*, 92(3):246–55, Sep 1999.

[119] D. G. Wright and J. I. Gallin. Secretory responses of human neutrophils: exocytosis of specific (secondary) granules by human neutrophils during adherence in vitro and during exudation in vivo. *J Immunol*, 123(1):285–94, Jul 1979.

[120] A. P. Kaplan, A. B. Kay, and K. F. Austen. A prealbumin activator of prekallikrein. 3. Appearance of chemotactic activity for human neutrophils by the conversion of human prekallikrein to kallikrein. *J Exp Med*, 135(1):81–97, Jan 1972.

[121] Y. T. Wachtfogel, U. Kucich, H. L. James, C. F. Scott, M. Schapira, M. Zimmerman, A. B. Cohen, and R. W. Colman. Human plasma kallikrein releases neutrophil elastase during blood coagulation. *J Clin Invest*, 72(5): 1672–7, Nov 1983.

[122] D. C. Altieri, R. Bader, P. M. Mannucci, and T. S. Edgington. Oligospecificity of the cellular adhesion receptor Mac-1 encompasses an inducible recognition specificity for fibrinogen. *J Cell Biol*, 107(5):1893–900, Nov 1988.

[123] M. S. Diamond, D. E. Staunton, S. D. Marlin, and T. A. Springer. Binding of the integrin Mac-1 (CD11b/CD18) to the third immunoglobulin-like domain of ICAM-1 (CD54) and its regulation by glycosylation. *Cell*, 65 (6):961–71, Jun 1991.

[124] R. Rothlein and T. A. Springer. Complement receptor type three-dependent degradation of opsonized erythrocytes by mouse macrophages. *J Immunol*, 135(4):2668–72, Oct 1985.

[125] O. F. Ramos, C. Kai, E. Yefenof, and E. Klein. The elevated natural killer sensitivity of targets carrying surface-attached C3 fragments require the availability of the iC3b receptor (CR3) on the effectors. *J Immunol*, 140(4): 1239–43, Feb 1988.

[126] M. A. Arnaout, L. L. Lanier, and D. V. Faller. Relative contribution of

the leukocyte molecules Mo1, LFA-1, and p150,95 (LeuM5) in adhesion of granulocytes and monocytes to vascular endothelium is tissue- and stimulus-specific. *J Cell Physiol*, 137(2):305–9, Nov 1988.

[127] D. C. Anderson, L. J. Miller, F. C. Schmalstieg, R. Rothlein, and T. A. Springer. Contributions of the Mac-1 glycoprotein family to adherence-dependent granulocyte functions: structure-function assessments employing subunit-specific monoclonal antibodies. *J Immunol*, 137(1):15–27, Jul 1986.

[128] D. I. Simon, Z. Chen, H. Xu, C. Q. Li, J. f. Dong, L. V. McIntire, C. M. Ballantyne, L. Zhang, M. I. Furman, M. C. Berndt, and J. A. López. Platelet glycoprotein ibalpha is a counterreceptor for the leukocyte integrin Mac-1 (CD11b/CD18). *J Exp Med*, 192(2):193–204, Jul 2000.

[129] N. Sheng, M. B. Fairbanks, R. L. Heinrikson, G. Canziani, I. M. Chaiken, D. M. Mosser, H. Zhang, and R. W. Colman. Cleaved high molecular weight kininogen binds directly to the integrin CD11b/CD18 (Mac-1) and blocks adhesion to fibrinogen and ICAM-1. *Blood*, 95(12):3788–95, Jun 2000.

[130] E. J. Gustafson, H. Lukasiewicz, Y. T. Wachtfogel, K. J. Norton, A. H. Schmaier, S. Niewiarowski, and R. W. Colman. High molecular weight kininogen inhibits fibrinogen binding to cytoadhesins of neutrophils and platelets. *J Cell Biol*, 109(1):377–87, Jul 1989.

[131] L. Y. Yung, F. Lim, M. M. Khan, S. P. Kunapuli, L. Rick, R. W. Colman, and S. L. Cooper. Neutrophil adhesion on surfaces preadsorbed with high molecular weight kininogen under well-defined flow conditions. *Immunopharmacology*, 32(1-3):19–23, May 1996.

[132] M. M. Khan, S. P. Kunapuli, Y. Lin, A. Majluf-Cruz, R. A. Cadena, S. L. Cooper, and R. W. Colman. Three noncontiguous peptides comprise binding sites on high-molecular-weight kininogen to neutrophils. *Am J Physiol*, 275(1 Pt 2):H145–50, Jul 1998.

[133] R. N. Puri, F. Zhou, C. J. Hu, R. F. Colman, and R. W. Colman. High molecular weight kininogen inhibits thrombin-induced platelet aggregation and cleavage of aggregin by inhibiting binding of thrombin to platelets. *Blood*, 77(3):500–7, Feb 1991.

[134] E. I. Peerschke, K. B. Reid, and B. Ghebrehiwet. Identification of a novel 33-kDa C1q-binding site on human blood platelets. *J Immunol*, 152(12): 5896–901, Jun 1994.

[135] E. I. B. Peerschke, T. K. Murphy, and B. Ghebrehiwet. Activation-dependent surface expression of gC1qR/p33 on human blood platelets. *Thromb Haemost*, 89(2):331–9, Feb 2003.

[136] E. G. Huizinga, S. Tsuji, R. A. P. Romijn, M. E. Schiphorst, P. G. de Groot, J. J. Sixma, and P. Gros. Structures of glycoprotein Ibalpha and its complex with von Willebrand factor A1 domain. *Science*, 297(5584):1176–9, Aug 2002.

[137] F. Adam, M.-C. Guillin, and M. Jandrot-Perrus. Glycoprotein Ib-mediated platelet activation. A signalling pathway triggered by thrombin. *Eur J Biochem*, 270(14):2959–70, Jul 2003.

[138] K. Joseph, Y. Nakazawa, W. F. Bahou, B. Ghebrehiwet, and A. P. Kaplan. Platelet glycoprotein Ib: a zinc-dependent binding protein for the heavy chain of high-molecular-weight kininogen. *Mol Med*, 5(8):555–63, Aug 1999.

[139] T. Chavakis, S. Santoso, K. J. Clemetson, U. J. H. Sachs, I. Isordia-Salas, R. A. Pixley, P. P. Nawroth, R. W. Colman, and K. T. Preissner. High molecular weight kininogen regulates platelet-leukocyte interactions by bridging Mac-1 and glycoprotein Ib. *J Biol Chem*, 278(46):45375–81, Nov 2003.

[140] H. N. Bradford, R. A. Dela Cadena, S. P. Kunapuli, J. F. Dong, J. A. Löpez, and R. W. Colman. Human kininogens regulate thrombin binding to

platelets through the glycoprotein Ib-IX-V complex. *Blood*, 90(4):1508–15, Aug 1997.

[141] K.-B. Kim, B.-W. Kim, H.-J. Choo, Y.-C. Kwon, B.-Y. Ahn, J.-S. Choi, J.-S. Lee, and Y.-G. Ko. Proteome analysis of adipocyte lipid rafts reveals that gC1qR plays essential roles in adipogenesis and insulin signal transduction. *Proteomics*, 9(9):2373–82, May 2009.

[142] E. I. B. Peerschke, J. O. Minta, S. Z. Zhou, A. Bini, A. Gotlieb, R. W. Colman, and B. Ghebrehiwet. Expression of gC1q-R/p33 and its major ligands in human atherosclerotic lesions. *Mol Immunol*, 41(8):759–66, Jul 2004.

[143] S. K. Petersen-Mahrt, C. Estmer, C. Ohrmalm, D. A. Matthews, W. C. Russell, and G. Akusjärvi. The splicing factor-associated protein, p32, regulates RNA splicing by inhibiting ASF/SF2 RNA binding and phosphorylation. *EMBO J*, 18(4):1014–24, Feb 1999.

[144] E. I. B. Peerschke and B. Ghebrehiwet. cC1qR/CR and gC1qR/p33: observations in cancer. *Mol Immunol*, 61(2):100–9, Oct 2014.

[145] T. Muta, D. Kang, S. Kitajima, T. Fujiwara, and N. Hamasaki. p32 protein, a splicing factor 2-associated protein, is localized in mitochondrial matrix and is functionally important in maintaining oxidative phosphorylation. *J Biol Chem*, 272(39):24363–70, Sep 1997.

[146] J. Dedio, W. Jahnen-Dechent, M. Bachmann, and W. Müller-Esterl. The multiligand-binding protein gC1qR, putative C1q receptor, is a mitochondrial protein. *J Immunol*, 160(7):3534–42, Apr 1998.

[147] B. J. Soltys, D. Kang, and R. S. Gupta. Localization of P32 protein (gC1q-R) in mitochondria and at specific extramitochondrial locations in normal tissues. *Histochem Cell Biol*, 114(3):245–55, Sep 2000.

[148] M. Majumdar, J. Meenakshi, S. K. Goswami, and K. Datta. Hyaluronan binding protein 1 (HABP1)/C1QBP/p32 is an endogenous substrate for

MAP kinase and is translocated to the nucleus upon mitogenic stimulation. *Biochem Biophys Res Commun*, 291(4):829–37, Mar 2002.

[149] H. C. van Leeuwen and P. O'Hare. Retargeting of the mitochondrial protein p32/gC1Qr to a cytoplasmic compartment and the cell surface. *J Cell Sci*, 114(Pt 11):2115–23, Jun 2001.

[150] G. Natsoulis, F. Hilger, and G. R. Fink. The HTS1 gene encodes both the cytoplasmic and mitochondrial histidine tRNA synthetases of S. cerevisiae. *Cell*, 46(2):235–43, Jul 1986.

[151] I. Stein, Y. Peleg, S. Even-Ram, and O. Pines. The single translation product of the FUM1 gene (fumarase) is processed in mitochondria before being distributed between the cytosol and mitochondria in Saccharomyces cerevisiae. *Mol Cell Biol*, 14(7):4770–8, Jul 1994.

[152] S. R. Reddigari, Y. Shibayama, T. Brunnée, and A. P. Kaplan. Human Hageman factor (factor XII) and high molecular weight kininogen compete for the same binding site on human umbilical vein endothelial cells. *J Biol Chem*, 268(16):11982–7, Jun 1993.

[153] B. Ghebrehiwet, C. CebadaMora, L. Tantral, J. Jesty, and E. I. B. Peerschke. gC1qR/p33 serves as a molecular bridge between the complement and contact activation systems and is an important catalyst in inflammation. *Adv Exp Med Biol*, 586:95–105, 2006.

[154] J. Jiang, Y. Zhang, A. R. Krainer, and R. M. Xu. Crystal structure of human p32, a doughnut-shaped acidic mitochondrial matrix protein. *Proc Natl Acad Sci U S A*, 96(7):3572–7, Mar 1999.

[155] B. Ghebrehiwet, J. Jesty, S. Xu, R. Vinayagasundaram, U. Vinayagasundaram, Y. Ji, A. Valentino, K. K. Hosszu, S. Mathew, K. Joseph, A. P. Kaplan, and E. I. B. Peerschke. Structure-function studies using deletion mutants identify domains of gC1qR/p33 as potential therapeutic targets for vascular permeability and inflammation. *Front Immunol*, 2, Nov 2011.

[156] K. Moffat and B. T. Chait. Biophysical methods: doing more with less. *Curr Opin Struct Biol*, 13(5):535–7, Oct 2003.

[157] C. Chung. The use of biophysical methods increases success in obtaining liganded crystal structures. *Acta Crystallogr D Biol Crystallogr*, 63(Pt 1): 62–71, Jan 2007.

[158] P. Schuck and H. Zhao. Biophysical methods for the study of protein interactions. *Methods*, 59(3):259–60, Mar 2013.

[159] S. M. Kelly, T. J. Jess, and N. C. Price. How to study proteins by circular dichroism. *Biochim Biophys Acta*, 1751(2):119–39, Aug 2005.

[160] W. A. Baase and W. C. Johnson, Jr. Circular dichroism and DNA secondary structure. *Nucleic Acids Res*, 6(2):797–814, Feb 1979.

[161] E. S. Stevens. Vacuum UV circular dichroism of polysaccharides. *Photochem Photobiol*, 44(3):287–93, Sep 1986.

[162] N. J. Greenfield. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc*, 1(6):2876–90, 2006.

[163] B. Jirgensons. *Molecular Biology Biochemistry and Biophysics*, volume 5, chapter Optical Activity of Amino Acids, Peptides, and Proteins, pages 47–56. Springer Berlin Heidelberg, Berlin, Heidelberg, 1973.

[164] A. Micsonai, F. Wien, L. Kernya, Y.-H. Lee, Y. Goto, M. Réfrégiers, and J. Kardos. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc Natl Acad Sci U S A*, 112(24): E3095–103, Jun 2015.

[165] G. Holzwarth and P. Doty. The Ultraviolet Cicrcular Dichrosim Of Polypeptides. *J Am Chem Soc*, 87:218–28, Jan 1965.

[166] N. Greenfield and G. D. Fasman. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry*, 8(10):4108–16, Oct 1969.

[167] V. SYu, I. A. Baikalov, Z. M. Shen, C. S. Wu, and J. T. Yang. Circular dichroic analysis of denatured proteins: inclusion of denatured proteins in the reference set. *Anal Biochem*, 214(1):17–24, Oct 1993.

[168] N. J. Greenfield. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat Protoc*, 1(6):2527–35, 2006.

[169] C. N. Pace. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol*, 131:266–80, 1986.

[170] J. S. Valastyan and S. Lindquist. Mechanisms of protein-folding diseases at a glance. *Dis Model Mech*, 7(1):9–14, Jan 2014.

[171] Y. Liang. Applications of isothermal titration calorimetry in protein science. *Acta Biochim Biophys Sin (Shanghai)*, 40(7):565–76, Jul 2008.

[172] M. J. Cliff and J. E. Ladbury. A survey of the year 2002 literature on applications of isothermal titration calorimetry. *J Mol Recognit*, 16(6):383–91, 2003.

[173] M. J. Cliff, A. Gutierrez, and J. E. Ladbury. A survey of the year 2003 literature on applications of isothermal titration calorimetry. *J Mol Recognit*, 17(6):513–23, 2004.

[174] O. Okhrimenko and I. Jelesarov. A survey of the year 2006 literature on applications of isothermal titration calorimetry. *J Mol Recognit*, 21(1):1–19, 2008.

[175] R. J. Falconer and B. M. Collins. Survey of the year 2009: applications of isothermal titration calorimetry. *J Mol Recognit*, 24(1):1–16, 2011.

[176] A. Cooper and C. M. Johnson. Introduction to microcalorimetry and biomolecular energetics. *Methods Mol Biol*, 22:109–24, 1994.

[177] Y. Tang, X. Zeng, and J. Liang. Surface Plasmon Resonance: An Introduction to a Surface Spectroscopy Technique. *J Chem Educ*, 87(7):742–746, Jul 2010.

[178] A. Brückner, C. Polge, N. Lentze, D. Auerbach, and U. Schlattner. Yeast two-hybrid, a powerful tool for systems biology. *Int J Mol Sci*, 10(6):2763–88, Jun 2009.

[179] X. Du, Y. Li, Y.-L. Xia, S.-M. Ai, J. Liang, P. Sang, X.-L. Ji, and S.-Q. Liu. Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int J Mol Sci*, 17(2), Jan 2016.

[180] J. C. KENDREW, G. BODO, H. M. DINTZIS, R. G. PARRISH, H. WYCKOFF, and D. C. PHILLIPS. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature*, 181(4610):662–6, Mar 1958.

[181] F. C. Bernstein, T. F. Koetzle, G. J. Williams, E. F. Meyer, Jr, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J Mol Biol*, 112(3):535–42, May 1977.

[182] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. The Protein Data Bank. *Nucleic Acids Res*, 28(1):235–42, Jan 2000.

[183] E. L. Forsythe, D. L. Maxwell, and M. Pusey. Vapor diffusion, nucleation rates and the reservoir to crystallization volume ratio. *Acta Crystallogr D Biol Crystallogr*, 58(Pt 10 Pt 1):1601–5, Oct 2002.

[184] I. Russo Krauss, A. Merlino, A. Vergara, and F. Sica. An overview of biological macromolecule crystallization. *Int J Mol Sci*, 14(6):11643–91, May 2013.

[185] M. A. Dessau and Y. Modis. Protein crystallization for X-ray crystallography. *J Vis Exp*, (47), Jan 2011.

[186] S. E. Ealick and R. L. Walter. Synchrotron beamlines for macromolecular crystallography. *Current Opinion in Structural Biology*, 3(5):725 – 736, 1993.

[187] S. M. Gruner and S. E. Ealick. Charge coupled device X-ray detectors for macromolecular crystallography. *Structure*, 3(1):13–5, Jan 1995.

[188] W. A. Hendrickson. Anomalous diffraction in crystallographic phase evaluation. *Q Rev Biophys*, 47(1):49–93, Feb 2014.

[189] M. F. Perutz. Isomorphous replacement and phase determination in non-centrosymmetric space groups. *Acta Crystallographica*, 9(11):867–873, Nov 1956.

[190] G. L. Taylor. Introduction to phasing. *Acta Crystallogr D Biol Crystallogr*, 66(Pt 4):325–38, Apr 2010.

[191] P. Evans and A. McCoy. An introduction to molecular replacement. *Acta Crystallogr D Biol Crystallogr*, 64(Pt 1):1–10, Jan 2008.

[192] A. Ilari and C. Savino. Protein structure determination by x-ray crystallography. *Methods Mol Biol*, 452:63–87, 2008.

[193] A. J. C. Wilson. Largest likely values for the reliability index. *Acta Crystallographica*, 3(5):397–398, Sep 1950.

[194] N. Kirshenbaum, I. Michaelevski, and M. Sharon. Analyzing large protein complexes by structural mass spectrometry. *J Vis Exp*, (40), Jun 2010.

[195] S. J. Eyles and I. A. Kaltashov. Methods to study protein dynamics and folding by mass spectrometry. *Methods*, 34(1):88–99, Sep 2004.

[196] X. Han, A. Aslanian, and J. R. Yates, 3rd. Mass spectrometry for proteomics. *Curr Opin Chem Biol*, 12(5):483–90, Oct 2008.

[197] S. Kaveti and J. R. Engen. Protein interactions probed with mass spectrometry. *Methods Mol Biol*, 316:179–97, 2006.

[198] S. Banerjee and S. Mazumdar. Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *Int J Anal Chem*, 2012:282574, 2012.

[199] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246(4926):64–71, Oct 1989.

[200] M. Karas and F. Hillenkamp. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem*, 60(20):2299–301, Oct 1988.

[201] C. S. Ho, C. W. K. Lam, M. H. M. Chan, R. C. K. Cheung, L. K. Law, L. C. W. Lit, K. F. Ng, M. W. M. Suen, and H. L. Tai. Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin Biochem Rev*, 24(1):3–12, 2003.

[202] N. Mano and J. Goto. Biomedical and biological mass spectrometry. *Anal Sci*, 19(1):3–14, Jan 2003.

[203] R. J. Cotter. Time-of-flight mass spectrometry: an increasing role in the life sciences. *Biomed Environ Mass Spectrom*, 18(8):513–32, Aug 1989.

[204] I. A. Kaltashov, C. E. Bobst, and R. R. Abzalimov. Mass spectrometry-based methods to study protein architecture and dynamics. *Protein Sci*, 22(5):530–44, May 2013.

[205] L. P. McIntosh and F. W. Dahlquist. Biosynthetic incorporation of 15N and 13C for assignment and interpretation of nuclear magnetic resonance spectra of proteins. *Q Rev Biophys*, 23(1):1–38, Feb 1990.

[206] L. Kay, P. Keifer, and T. Saarinen. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *Journal of the American Chemical Society*, 114(26):10663–10665, 1992.

[207] M. R. O'Connell, R. Gamsjaeger, and J. P. Mackay. The structural analysis of protein-protein interactions by NMR spectroscopy. *Proteomics*, 9(23): 5224–32, Dec 2009.

[208] A. R. Krainer, A. Mayeda, D. Kozak, and G. Binns. Functional expression of cloned human splicing factor SF2: homology to RNA-binding proteins, U1 70K, and Drosophila splicing regulators. *Cell*, 66(2):383–94, Jul 1991.

[209] S. Zhang and M. D. Cahalan. Purifying plasmid DNA from bacterial colonies using the QIAGEN Miniprep Kit. *J Vis Exp*, (6):247, 2007.

[210] M. Mandel and A. Higa. Calcium-dependent bacteriophage DNA infection. *J Mol Biol*, 53(1):159–62, Oct 1970.

[211] A. G. Sabelnikov, A. V. Avdeeva, and Ilyashenko. Enhanced uptake of donor DNA by Ca2+ treated Escherichia coli cells. *Mol Gen Genet*, 138(4): 351–8, Jul 1975.

[212] H. E. Bergmans, I. M. van Die, and W. P. Hoekstra. Transformation in Escherichia coli: stages in the process. *J Bacteriol*, 146(2):564–70, May 1981.

[213] P. Y. Lee, J. Costumbrado, C.-Y. Hsu, and Y. H. Kim. Agarose gel electrophoresis for the separation of DNA fragments. *J Vis Exp*, (62), Apr 2012.

[214] P. Jajesniak and T. S. Wong. QuickStep-Cloning: a sequence-independent, ligation-free method for rapid construction of recombinant plasmids. *J Biol Eng*, 9:15, 2015.

[215] S. Lacks and B. Greenberg. A deoxyribonuclease of Diplococcus pneumoniae specific for methylated DNA. *J Biol Chem*, 250(11):4060–66, Jun 1975.

[216] P. McInerney, P. Adams, and M. Z. Hadi. Error Rate Comparison during Polymerase Chain Reaction by DNA Polymerase. *Mol Biol Int*, 2014: 287430, 2014.

[217] A. Griffiths, W. Gelbart, and J. Miller. *Modern Genetic Analysis*, chapter Making Recombinant DNA. New York: W. H. Freeman, 1999.

[218] M. A. Speer and T. L. Richard. Amplified insert assembly: an optimized approach to standard assembly of BioBrickTM genetic circuits. *J Biol Eng*, 5(1):17, Dec 2011.

[219] H. Lodish, A. Berk, and S. Zipursky. *Molecular Cell Biology. 4th edition*, chapter 7.6, Producing High Levels of Proteins from Cloned cDNAs. New York: W. H. Freeman, 2000.

[220] F. W. Studier, A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol*, 185:60–89, 1990.

[221] A. Sivashanmugam, V. Murray, C. Cui, Y. Zhang, J. Wang, and Q. Li. Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. *Protein Sci*, 18(5):936–48, May 2009.

[222] S. C. Gill and P. H. von Hippel. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem*, 182(2):319–26, Nov 1989.

[223] R. K. Scopes. Measurement of protein by spectrophotometry at 205 nm. *Anal Biochem*, 59(1):277–82, May 1974.

[224] S. Adhikari, P. V. Manthena, K. Sajwan, K. K. Kota, and R. Roy. A unified method for purification of basic proteins. *Anal Biochem*, 400(2):203–6, May 2010.

[225] J. A. Bornhorst and J. J. Falke. Purification of proteins using polyhistidine affinity tags. *Methods Enzymol*, 326:245–54, 2000.

[226] J. Porath. Immobilized metal ion affinity chromatography. *Protein Expr Purif*, 3(4):263–81, Aug 1992.

[227] K. C. Duong-Ly and S. B. Gabelli. Using ion exchange chromatography to purify a recombinantly expressed protein. *Methods Enzymol*, 541:95–103, 2014.

[228] J. Porath and P. Flodin. Gel filtration: a method for desalting and group separation. *Nature*, 183(4676):1657–9, Jun 1959.

[229] D. B. Smith and K. S. Johnson. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene*, 67(1):31–40, Jul 1988.

[230] A. A. Al-Tubuly. SDS-PAGE and Western Blotting. *Methods Mol Med*, 40: 391–405, 2000.

[231] K. Joseph, B. G. Tholanikunnel, B. Ghebrehiwet, and A. P. Kaplan. Interaction of high molecular weight kininogen binding proteins on endothelial cells. *Thromb Haemost*, 91(1):61–70, Jan 2004.

[232] A. J. Huhn, D. Parsonage, D. A. Horita, F. M. Torti, S. V. Torti, and T. Hollis. The high-molecular-weight kininogen domain 5 is an intrinsically unstructured protein and its interaction with ferritin is metal mediated. *Protein Sci*, 23(8):1013–22, Aug 2014.

[233] J. L. Cole, J. W. Lary, T. P Moody, and T. M. Laue. Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods Cell Biol*, 84:143–79, 2008.

[234] D. H. Williams, C. T. Calderone, D. P. O'Brien, and R. Zerella. Changes in motion vs. bonding in positively vs. negatively cooperative interactions. *Chem Commun (Camb)*, (12):1266–7, Jun 2002.

[235] K. P. Murphy and S. J. Gill. Solid model compounds and the thermodynamics of protein unfolding. *J Mol Biol*, 222(3):699–709, Dec 1991.

[236] G. I. Makhatadze and P. L. Privalov. On the entropy of protein folding. *Protein Sci*, 5(3):507–10, Mar 1996.

[237] D. Chandler. Interfaces and the driving force of hydrophobic assembly. *Nature*, 437(7059):640–7, Sep 2005.

[238] M. L. Bianconi. Avoiding Buffer Interference in ITC Experiments: A Case Study from the Analysis of Entropy-Driven Reactions of Glucose-6-Phosphate Dehydrogenase. *Methods Enzymol*, 567:237–56, 2016.

[239] A. G. Aslamkhan, A. Aslamkhan, and G. A. Ahearn. Preparation of metal ion buffers for biological experimentation: a methods approach with emphasis on iron and zinc. *J Exp Zool*, 292(6):507–22, May 2002.

[240] H. J. D. Peter E. Wright. Linking Folding and Binding. *Curr Opin Struct Biol*, 19(1):31–38, 2009.

[241] I. R. Kleckner and M. P. Foster. An introduction to NMR-based approaches for measuring protein dynamics. *Biochim Biophys Acta*, 1814(8): 942–68, Aug 2011.

[242] H. Herwald, M. Mörgelin, H. G. Svensson, and U. Sjöbring. Zinc-dependent conformational changes in domain D5 of high molecular mass kininogen modulate contact activation. *Eur J Biochem*, 268(2):396–404, Jan 2001.

[243] A. B. T. Ghisaidoobe and S. J. Chung. Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Förster resonance energy transfer techniques. *Int J Mol Sci*, 15(12):22518–38, Dec 2014.

[244] K. A. McCall, C. Huang, and C. A. Fierke. Function and mechanism of zinc metalloenzymes. *J Nutr*, 130(5S Suppl):1437S–46S, May 2000.

[245] G. Ding, P. Lorenz, M. Kreutzer, Y. Li, and H.-J. Thiesen. SysZNF: the C2H2 zinc finger gene database. *Nucleic Acids Res*, 37(Database issue): D267–73, Jan 2009.

[246] T. Fujii, Y. Hata, M. Oozeki, H. Moriyama, T. Wakagi, N. Tanaka, and T. Oshima. The crystal structure of zinc-containing ferredoxin from the

thermoacidophilic archaeon Sulfolobus sp. strain 7. *Biochemistry*, 36(6): 1505–13, Feb 1997.

[247] A. Liljas, K. K. Kannan, P. C. Bergstén, I. Waara, K. Fridborg, B. Strandberg, U. Carlbom, L. Järup, S. Lövgren, and M. Petef. Crystal structure of human carbonic anhydrase C. *Nat New Biol*, 235(57):131–7, Feb 1972.

[248] O. Dideberg, P. Charlier, G. Dive, B. Joris, J. M. Frère, and J. M. Ghuysen. Structure of a Zn2+-containing D-alanyl-D-alanine-cleaving carboxypeptidase at 2.5 A resolution. *Nature*, 299(5882):469–70, Sep 1982.

[249] S. Karlin and Z. Y. Zhu. Classification of mononuclear zinc metal sites in protein structures. *Proc Natl Acad Sci U S A*, 94(26):14231–6, Dec 1997.

[250] P. Tompa and M. Fuxreiter. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci*, 33(1): 2–8, Jan 2008.

[251] P. E. Wright and H. J. Dyson. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol*, 16(1):18–29, Jan 2015.

[252] V. J. Hilser and E. B. Thompson. Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc Natl Acad Sci U S A*, 104(20): 8311–5, May 2007.

[253] A. C. M. Ferreon, J. C. Ferreon, P. E. Wright, and A. A. Deniz. Modulation of allostery by protein intrinsic disorder. *Nature*, 498(7454):390–4, Jun 2013.

[254] R. A. Pixley, Y. Lin, I. Isordia-Salas, and R. W. Colman. Fine mapping of the sequences in domain 5 of high molecular weight kininogen (HK) interacting with heparin and zinc. *J Thromb Haemost*, 1(8):1791–8, Aug 2003.

[255] W. Tempel, Y. Tong, L. Crombet, Y. Shen, X. Guan, L. Nedyalkova, J. Walker, C. Arrowsmith, A. Edwards, C. Bountra, J. Weigelt, and

H. Park. 3RPX, Crystal structure of complement component 1, q subcomponent binding protein, c1qbp. Apr 2011.

[256] S. S. Wong, S. Østergaard, G. Hall, C. Li, P. M. Williams, H. Stennicke, and J. Emsley. A novel DFP tripeptide motif interacts with the coagulation factor XI apple 2 domain. *Blood*, 127(23):2915–23, Jun 2016.

[257] M. Akif, G. Masuyer, R. J. Bingham, E. D. Sturrock, R. E. Isaac, and K. R. Acharya. Structural basis of peptide recognition by the angiotensin-1 converting enzyme homologue AnCE from Drosophila melanogaster. *FEBS J*, 279(24):4525–34, Dec 2012.

[258] A. K. Mackenzie, K. Valegård, A. Iqbal, M. E. C. Caines, N. J. Kershaw, S. E. Jensen, C. J. Schofield, and I. Andersson. Crystal structures of an oligopeptide-binding protein from the biosynthetic pathway of the beta-lactamase inhibitor clavulanic acid. *J Mol Biol*, 396(2):332–44, Feb 2010.

[259] T. Le Gall, P. R. Romero, M. S. Cortese, V. N. Uversky, and A. K. Dunker. Intrinsic disorder in the Protein Data Bank. *J Biomol Struct Dyn*, 24(4): 325–42, Feb 2007.

[260] M. C. Deller, L. Kong, and B. Rupp. Protein stability: a crystallographer's perspective. *Acta Crystallogr F Struct Biol Commun*, 72(Pt 2):72–95, Feb 2016.

[261] B. Heras and J. L. Martin. Post-crystallization treatments for improving diffraction quality of protein crystals. *Acta Crystallogr D Biol Crystallogr*, 61(Pt 9):1173–80, Sep 2005.

[262] I. Russo Krauss, F. Sica, C. A. Mattia, and A. Merlino. Increasing the X-ray diffraction power of protein crystals by dehydration: the case of bovine serum albumin and a survey of literature data. *Int J Mol Sci*, 13(3):3782–800, 2012.

[263] T. Bergfors. Seeds to crystals. *J Struct Biol*, 142(1):66–76, Apr 2003.

[264] S. Remani, J. Sun, R. Kotaria, J. A. Mayor, J. M. Brownlee, D. H. T. Harrison, D. E. Walters, and R. S. Kaplan. The yeast mitochondrial citrate transport protein: identification of the Lysine residues responsible for inhibition mediated by Pyridoxal 5'-phosphate. *J Bioenerg Biomembr*, 40(6): 577–85, Dec 2008.

[265] N. Osheroff, D. L. Brautigan, and E. Margoliash. Mapping of anion binding sites on cytochrome c by differential chemical modification of lysine residues. *Proc Natl Acad Sci U S A*, 77(8):4439–43, Aug 1980.

[266] F. F. Syed, J. Y. Kim, K. Y. Ha, and M. S. Braiman. Citrate-binding site in Proteorhodopsin involves two lysines in the first cytoplasmic loop. *Biophysical Journal*, 98(3), Jan 2010.

[267] W. X. Guo, B. Ghebrehiwet, B. Weksler, K. Schweitzer, and E. I. Peerschke. Up-regulation of endothelial cell binding proteins/receptors for complement component C1q by inflammatory cytokines. *J Lab Clin Med*, 133(6): 541–50, Jun 1999.

[268] E. M. Scarpati and J. E. Sadler. Regulation of endothelial cell coagulant properties. Modulation of tissue factor, plasminogen activator inhibitors, and thrombomodulin by phorbol 12-myristate 13-acetate and tumor necrosis factor. *J Biol Chem*, 264(34):20705–13, Dec 1989.

[269] P. R. Conkling, C. S. Greenberg, and J. B. Weinberg. Tumor necrosis factor induces tissue factor-like activity in human leukemia cell line U937 and peripheral blood monocytes. *Blood*, 72(1):128–33, Jul 1988.

[270] M. Ernofsson, T. Tenno, and A. Siegbahn. Inhibition of tissue factor surface expression in human peripheral blood monocytes exposed to cytokines. *Br J Haematol*, 95(2):249–57, Nov 1996.

[271] M. Rükgauer, J. Klein, and J. D. Kruse-Jarres. Reference values for the trace elements copper, manganese, selenium, and zinc in the

serum/plasma of children, adolescents, and adults. *J Trace Elem Med Biol*, 11(2):92–8, Jun 1997.

[272] B. J. Scott and A. R. Bradwell. Identification of the serum binding proteins for iron, zinc, cadmium, nickel, and calcium. *Clin Chem*, 29(4):629–33, Apr 1983.

[273] S. Tubek, P. Grzanka, and I. Tubek. Role of zinc in hemostasis: a review. *Biol Trace Elem Res*, 121(1):1–8, Jan 2008.

[274] A. B. Lansdown, B. Sampson, and A. Rowe. Sequential changes in trace metal, metallothionein and calmodulin concentrations in healing skin wounds. *J Anat*, 195 (Pt 3):375–86, Oct 1999.

[275] B. Foley, S. A. Johnson, B. Hackley, J. C. Smith, Jr, and J. A. Halsted. Zinc content of human platelets. *Proc Soc Exp Biol Med*, 128(1):265–9, May 1968.

[276] D. J. Bowen. Haemophilia A and haemophilia B: molecular insights. *Mol Pathol*, 55(1):1–18, Feb 2002.

[277] E. Berntorp. Haemophilia treatment in 2030. *Haemophilia*, 22 Suppl 5: 15–9, Jul 2016.

[278] R. Asakai, D. W. Chung, O. D. Ratnoff, and E. W. Davie. Factor XI (plasma thromboplastin antecedent) deficiency in Ashkenazi Jews is a bleeding disorder that can result from three types of point mutations. *Proc Natl Acad Sci U S A*, 86(20):7667–71, Oct 1989.

[279] T. Renné, B. Nieswandt, and D. Gailani. The intrinsic pathway of coagulation is essential for thrombus stability in mice. *Blood Cells Mol Dis*, 36 (2):148–51, 2006.

[280] R. L. Bick and K. Ucar. Hypercoagulability and thrombosis. *Hematol Oncol Clin North Am*, 6(6):1421–31, Dec 1992.

[281] S. Shapiro. The hypoprothrombinemia-inducing activity of warfarin sodium (coumadin sodium). *J Kans Med Soc*, 55(12):687–91, Dec 1954.

[282] G. L. Nelsestuen, A. M. Shah, and S. B. Harvey. Vitamin K-dependent proteins. *Vitam Horm*, 58:355–89, 2000.

[283] T. Trujillo and P. P. Dobesh. Clinical use of rivaroxaban: pharmacokinetic and pharmacodynamic rationale for dosing regimens in different indications. *Drugs*, 74(14):1587–603, Sep 2014.

[284] B. E. Baetz and S. A. Spinler. Dabigatran etexilate: an oral direct thrombin inhibitor for prophylaxis and treatment of thromboembolic diseases. *Pharmacotherapy*, 28(11):1354–73, Nov 2008.

[285] J. N. Goldstein and S. M. Greenberg. Should anticoagulation be resumed after intracerebral hemorrhage? *Cleve Clin J Med*, 77(11):791–9, Nov 2010.

[286] D. Gailani and T. Renné. The intrinsic pathway of coagulation: a target for treating thromboembolic disease? *J Thromb Haemost*, 5(6):1106–12, Jun 2007.

[287] A. A. Stoop, R. V. Joshi, C. T. Eggers, and C. S. Craik. Analysis of an engineered plasma kallikrein inhibitor and its effect on contact activation. *Biol Chem*, 391(4):425–33, Apr 2010.

[288] W. Wu, H. Li, D. Navaneetham, Z. W. Reichenbach, R. F. Tuma, and P. N. Walsh. The kunitz protease inhibitor domain of protease nexin-2 inhibits factor XIa and murine carotid artery and middle cerebral artery thrombosis. *Blood*, 120(3):671–7, Jul 2012.

[289] S. Robert, C. Bertolla, B. Masereel, J.-M. Dogné, and L. Pochet. Novel 3-carboxamide-coumarins as potent and selective FXIIa inhibitors. *J Med Chem*, 51(11):3077–80, Jun 2008.

[290] M. R. Arkin, Y. Tang, and J. A. Wells. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chem Biol*, 21(9): 1102–14, Sep 2014.