
THE UNIVERSITY *of* EDINBURGH

# Investigations into polymorphisms within complement receptor type 1 (CD35) thought to protect against severe malaria

**PATIENCE BORKOR TETTEH-QUARCOO**


**PhD**

**The University of Edinburgh**

**2011**

## *Abstract*

The human immune-regulatory protein, complement receptor type 1 (CR1, CD35), occurs on erythrocytes where it serves as the immune adherence receptor. It interacts with C3b, C4b, C1q and mannan-binding lectin (MBL). It additionally binds the *Plasmodium falciparum* protein, Rh4, in the non-sialic acid-dependent erythrocye-invasion pathway, and is also important for rosetting, via an interaction with *P. falciparum* erythrocyte membrane protein 1 (PfEMP1). A C3b/C4b, and PfEMP1 binding site lies in CCP modules 15-17 (out of 30 in CR1), while polymorphisms that afford advantage to some populations in dealing with severe malaria occur in CCPs 24-25, begging the question central to this thesis – do these polymorphism modulate function, and if so how? We hypothesized that the CR1 architecture apposes CCPs 15-17 and CCPs 24-25 using the exceptionally long linker between CCPs 21 and 22 as a hinge, thus polymorphic variants in CCPs 24-25 modulate functionality in CCPs 15-17. To test this, a panel of recombinant CR1 protein fragments (CCPs 21, 21-22, 20-23, 15-17, 17, 10-11, 17-25, 15-25 and 24-25) were produced in *Pichia pastoris* along with polymorphic forms of the relevant constructs. After purification, biophysical and biological methods were used to assess whether the linker does indeed act as a hinge, and the comparative abilities of the CCPs 15-25 variants (along with soluble CR1 (sCR1), CCPs 1-3 and the panel of CR1 fragments) to interact with a range of ligands were measured.

We found no evidence from NMR for face-to-face contacts between CCPs 21 and 22 that would be consistent with the long linker permitting a 180-degree bend between them. Indeed, based on scattering and analytical ultracentrifugation data, CCPs 20-23 form an extended rather than a bent-back structure. All of the four Knops blood-group variants of the CCPs 15-25 proteins produced similar results according to dynamic light scattering and AUC indicating no structural difference or change in self-association state between variants. In addition, based on the data collected from surface plasmon resonance (SPR), ELISA and fluid-phase cofactor (for factor I) assays, there were no evidence of any difference between the polymorphic forms with respect to their interactions with C3b, C4b, C1q and MBL. Only weak interaction was observed for sCR1, and all CCPs 15-25 variants, with the relevant part of PfEMP1, and there was no measurable difference amongst the variants in disrupting rosettes. The sCR1-Rh4.9 interaction was confirmed by SPR; affinities measured between the binding domain of Rh4 and the panel of CR1 fragments identified CCPs 1-3 (site 1) as the main interaction site. It seemed unlikely therefore that CCPs 24 and 25 could modulate Rh4 binding; indeed none of the four CR1 15-25 variants bound Rh4.9 appreciably. Thus we concluded that allotypic variations in CCPs 24-25 have no measurable effect on the architecture as well as binding of CR1 to its host or parasite ligands The inferred selective pressure acting on these variants likely arise from some other (*i.e.* besides malaria) geographically localised infectious diseases.

## *Acknowledgments*

I am grateful to my supervisors, Professor Paul N. Barlow and Professor Alex Rowe for giving me the unique privilege of working with them on this project and for all the encouraging support that I have received throughout my stay in Edinburgh. They were everything for me not only assisting with my academic work but also my social life. I especially appreciate their immense patience, understanding, support and encouragement during the later part of my course which proved very difficult for me as well as them.

My profound gratitude goes to the Darwin Trust of Edinburgh for funding without which, my academic life would have been cut short and never have had the opportunity to such lovely people in Edinburgh**.**

The assistance and timely suggestions that I received from members of the BioNMR group (the Barlow's group) in School of Chemistry Edinburgh are very much appreciated. Among these are the individual contributions of Christoph, Andy, Mara, Dinesh, Janice, Dušan, Juraj, John, Isabell, Marie, Carla, Conny, Maria, Elias and Eliza, Nicky, Fern, Bärbel, Yoshi, Henry, Dave, and Carina. Dr. Christoph Q. Schmidt is being acknowledged for helping immensely throughout the project from the onset with molecular work through protein expression SPR and NMR. I am greatful Christoph for the additional friendship you afforded my family when we were away from our home country. Dr.Andy Herbert is appreciatiated for all the assistance in both wet laboratory work and NMR, Dr.Dinesh Sores for the Bioinfomatics assistance and for always being available to help with structural models and domain boundaries, Dr. Mara Guariento for all the assistance, especially with the NMR. Mr. Juraj Belle and Dr. Dušan Uhrin for their patience and help with all the NMR experiments, Dr Janice Bramham for assistance with SPR experiments and Dr. John White for the immense assistance with the fermentations.

I also acknowledge the numerous help from Members of Prof.Alex Rowe's group I especially want to mention the immense assistance given by Mr. Ahmed Raza for all the technical support given with starting and maintaining malaria cultures, Dr. Ashfag Ghumra for assistance and suggestions concerning the handling of DBL's, Dr. Monica Arman for all the encouragements during the course, and to all then PhD students of the group (such as Antoine Claessens) for timely encouragments.Many thanks to Dr Mattew Higgins for timely providing most of the DBLs used in this study and for collaborating with us through the Rowe group.

Many thanks also go to Professor John Atkinson, Richard Hauhart, Thom Allen and all group members, (of Washington University School of Medicine, St.Louis, USA) for the supply of the CR1 cDNA and all their collaborative supports. Dr. Tham Wai-Hong, Prof. Alan Cowman, Prof. Arthur Rowe and Dr. Haydyn Mertens for their nice collaborative time.

Extreme gratitude is what I extend to Prof. Adjei, Prof Lawson, Prof Newman, and Dr Ayeh-Kumi, and also all members of the University of Ghana Medical School (especially Department of Microbiology), College of health sciences, for their heart warming support.

My heartfelt appreciation goes to my husband, family members and friends who tirelessly supported me throughout the difficult times, not forgetting Emmanuel Afutu.

## *Declaration*

With the exception of those duly acknowledged , referenced or clearly stated in the text as done by others, the work described in this thesis is mine and has not been submitted in whole or part for any other degree or qualification in this university or any other.

……

PATIENCE BORKOR TETTEH-QUARCOO

## *Table of Contents*

## *Table of Figures*

## *List of Table*

## *List of flow charts*

## *Abbreviations*

| | |
|---|---|
| BMG | Buffered minimal glycerol |
| BMM | Buffered minimal methanol |
| Bp | base pairs |
| BSA | Bovine serum albumin |
| C4bp | C4b binding protein |
| CCP | Complement control protein |
| CR1 | Complement receptor1 |
| DAF | Decay Accelerating Factor |
| DBLα | Duffy-Like binding domain alpha |
| DMSO | dimethyl sulfoxide |
| DNA | deoxyribonucleic acid |
| dNTP | any deoxyribonucleoside triphosphate |
| DTPA | diethylenetriaminpentaacetic acid |
| *E. coli* | *Escherichia coli* |
| EDTA | ethylenediaminetetraacetic acid |
| fI | Factor I |
| H | hour(s) |
| kDa | Kilodalton |
| LB | Luria-Bertani medium |
| MAC | Membrane attack complex |
| Min | Minutes |
| Mw | molecular weight |
| NEB | New England Biolabs |
| $OD_{600}$ | optical density at wavelength= 600 nm |
| PfEMP1 | *Plasmodium falciparum* erythrocyte membrane protein 1 |
| pI | isoelectric point |
| RCA | regulators of complement activation |
| Rpm | revolutions per minute |
| SDS-PAGE | Sodium dodecyl sulphate – polyacrylamide gel |
| TAE | tris acetate EDTA |
| TCA | Trichloroacetic acid |
| Tris | Tris(hydoxymethyl)aminomethane |
| w/v | Weight by volume |
| Wt | wild type |
| YNB | Yeast nitrogen Base medium |
| YPD | Yeast Extract Peptone Dextrose |
| YPDS | Yeast Extract Peptone Dextrose Medium with Sorbitol |

# CHAPTER 1 INTRODUCTION

# CHAPTER ONE

# INTRODUCTION

# CHAPTER 1 INTRODUCTION

## 1.1 Polymorphisms and disease

Polymorphisms in genes have been studied, over many years, in search of possible associations with disease conditions. A polymorphism is a variation in the DNA that is too common to be due merely to a new mutation. A polymorphism must have a frequency of at least 1% in the population (Manning, 2010). Examples of very well-studied polymorphisms include those found in genes responsible for sickle-cell disease (Allison, 1956; Ford, 1973 ; Aidoo *et al*., 2002, Williams *et al*., 2005), the ABO blood group antigens (Clark, 1964; Crow, 1993; Meade and Earickson 2005 ; Cserti and Dzik, 2007) and glucose-6-phosphate dehydrogenase deficiency (Beutler, 1994; Verrelli *et al*., 2002). In each of these cases the possession by an individual of a certain variant appears to offer protection against malaria or other infectious diseases, and it is thought that these polymorphisms have become widespread in populations native to disease endemic areas as a result of selective pressure **(**Cooke and Hill, 2001; Sykes, 1999**).** While some polymorphisms have been strongly linked to increased or decreased predisposition to specific diseases, others appear unconnected to any known pathology. A third category consists of polymorphisms for which circumstantial evidence supports a disease link but further work is needed to confirm this and establish its molecular basis.

### 1.1.1 Debate on Knops blood group polymorphisms and their association with severe malaria

A good example of this third category are the Knops blood group polymorphisms, the best studied of which encodes sequence variations found in long homologous repeat (LHR)-D of the complement receptor type 1 (CR1)/immune adherence receptor. The strikingly non-uniform geographical distribution of these polymorphisms (Rowe *et al.*, 2009; Stoute, 2005) has given rise to the widely accepted hypothesis that they are also connected to a survival advantage in malaria-endemic areas. More specifically the possession of certain Knops blood group antigens seem to be linked to protection against severe malarial anaemia and cerebral malaria (Rowe *et al.*, 2009**;** Stoute, 2005**,** 2011) background material on malaria may be found below).

This proposition had been under debate for many years. In particular, the identification of CR1 as an erythrocyte-borne ligand for *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) in field isolates, and the demonstration that this interaction contributes to the formation of rosettes containing infected and uninfected red blood cells – a phenomenon that, based on several studies, has been associated with disease severity - seemed to bolster the case for such a link (Rowe *et al*., 1995; Heddini *et al*., 2001, Carlson *et al*., 1990 and Treutiger *et al*., 1992).

On the other hand, other studies (Zimmerman *et al.,* 2003; Jallow *et al.*, 2009; Bellamy *et al.*, 1998) concluded that, Knops blood group alleles are not associated with protection against severe malaria in the Gambia. These authors suggested that the alleles might confer a selective advantage against infectious disease in general with no specificity for *P. falciparum* malaria. These findings, however, were complicated by variations between the various alleles in terms of CR1 copy number on erythrocytes.

Another study (Thathy *et al.*, 2005) suggested there was association between the polymorphic forms and severe forms of malaria in Kenya. These results supported the hypothesis that the *Swain-Langley(2)* (*Sl(2))* allele and, possibly, the *McCoyb* (*McCb*) allele (see below for details of the specific amino acid variations that comprise the set of Knops blood group antigens) evolved against a background of malarial transmission and that certain allotypes might confer a survival advantage.

As has been pointed out by others (Krych-Goldberg *et al.,* 2002), a detailed study at the molecular level of the interaction between PfEMP1 and CR1 could help understand whether the Knops blood group alleles have any direct effect on affinity between these proteins. Furthermore, a wider exploration of the functional properties of the Knops blood group antigenic variants is clearly justified. This would include examination of their ability to regulate complement and to interact with the multiple binding partners of CR1 including the complement proteins (see below) C3b, C4b, and C1q, as well as PfEMP1 and the more recently identified additional parasite ligand important for erythrocyte invasion, Rh 4 (Krych *et al.,* 1991, Krych-Goldberg *et al.,* 2002; Klickstein *et al.,* 1997; Rowe *et al*., 2000; Tham *et al.,* 2010).

## 1.2 The complement system

The operation of the human immune system involves a combination of both innate (natural) and acquired (adaptive) immune responses. Innate immunity is made up of a range of components that together comprise a non-specific means of defence against pathogens. This range includes: physical barriers such as the skin and mucous

membranes; physiological factors such as the maintenance of certain pH values, temperatures and oxygen tensions; the secretion of proteins such as lysozyme into external body fluids; soluble protein factors within the blood stream and interstitial fluids such as those of the complement cascade, interferons, collectins, and C-reactive protein; and phagocytic cells including macrophages and polymorphonuclear leucocytes **(**Roitt *et al.*, 1998**).**

The complement system is considered the key molecular component of innate immunity, even though it is now well established that various complement proteins also act to augment antibody production by B-cells and to stimulate T-cells. Moreover the complement cascade, which is potentially cytolytic, is an important effector arm of adaptive immunity. The complement system is now known to consist of more than 30 plasma and membrane-associated proteins (Jha *et al.*, 2003) and to include both activators and regulators. There are three main routes to activation of the complement cascade (Figure 1.1A).

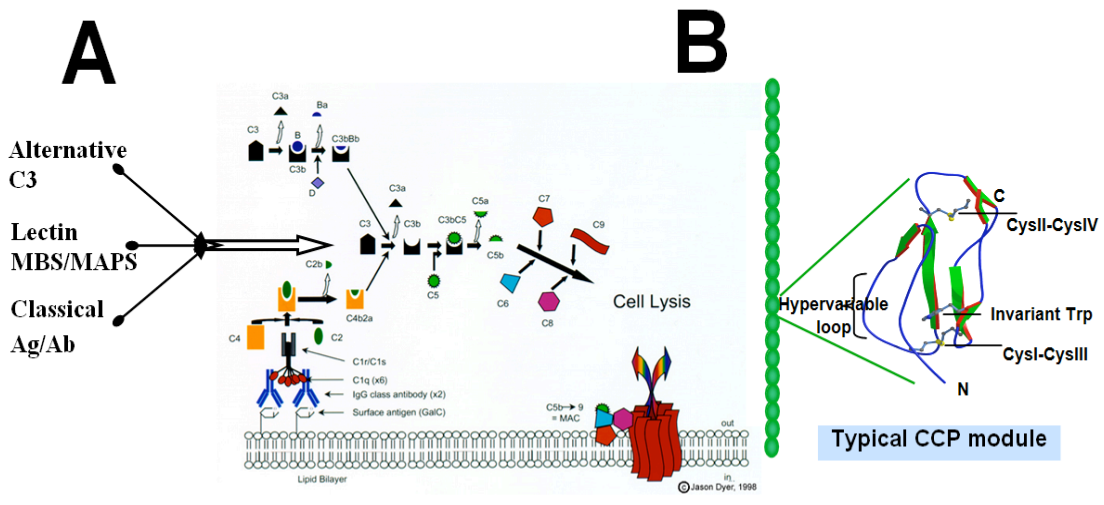



***Figure 1.1 The complement pathways and CCP module***

# CHAPTER 1 INTRODUCTION

**(A)** Schematic diagram of the complement cascade: The diagram draws attention to the central role that C3 plays in all three pathways. This highlights the need for its convertase – the enzyme that converts C3 to its activated product, C3b - to be properly regulated. **(B)** The green cartoon on the left represents the multiple complement control protein domains or modules (CCPs), joined together by short linking sequences like beads on a string, which comprise a typical member of the regulators of complement activation (RCA) family of proteins encoded on chromosome 1q32. Note that these proteins are very unlikely to be fully extended and rod-like as drawn here although in fact there are no structures available for any RCA or RCA fragment longer than four CCPs. To the right is shown a cartoon to illustrate the secondary and tertiary structures of a typical CCP.

The classical pathway is triggered by the binding to the complement C1 complex (composed of C1q, and two copies each of C1r and C1s) to antigen-antibody complexes (Morgan and Harris, 1999**)** via the globular heads of the trimeric C1q components. Specific sugars (mainly found on the surface of micro-organisms) trigger the lectin pathway via the mannose binding lectin (MBL)/MBL-associated serine protease (MASP) complex, which resembles C1 in terms of its moilecular architecture. In brief, activation of either C1 or MBL/MASP results in cleavage of the inert C4 component of the complement cascade to activated products C4a and C4b. Cleavage results in exposure of new binding sites on C4b and the activation of a thioester group such that it readily attaches covalently to nearby surfaces. Surface-bound C4b is an opsonin, targeting the cells and particles to which it is attached for phagocytosis. The nascent C4b additionally binds to C2, whereupon C2 becomes a substrate for the activated C1-complex and is thereby cleaved to 2a and 2b. The 2b fragment is released while the C4bC2a complex thus formed is the classical pathway C3 convertase; it cleaves the abundant inert plasma protein C3 into its activated form, C3b and the anaphylatoxin, C3a. C3 and C4 (both around 185 kD) are closely related paralogues and the conversion of C3 to C3b resembles that of C4 to C4b in terms of exposure of

new binding sites and attachment of C3b to surface via a covalent linkage. C3b, like C4b, is an opsonin.

The alternative pathway begins when low-rate but spontaneous C3 activation in plasma occurs, involving hydrolysis of its buried thioester group and generating C3($H_2O$). This protein thought to resemble C3b both structurally and functionally, binds to complement factor B to form the complex C3($H_2O$)B. In this context the factor B is cleaved by factor D forming C3(H2O)Bb. This (like C4b2a) is a C3 convertase (sometimes called the initiating c3 convertase), and unless it is regaled it will "kick-start" the complement cascade by generating more C3a and C3b molecules from C3. The activated product C3b (like C3(H2O)) binds to factor B and subsequently yields the C3bBb complex, which is yet another C3 convertase in this case called the alternative pathway convertase. Its action produces more C3b thus stoking a positive feed-back loop.

Since both C3b and C4b bind virtually indiscriminately to both foreign and self surfaces, regulators are essential to ensure that levels of C3b (and C3a) and C4b are controlled, on or near host cell membranes that would otherwise become opsonised or lysed by further steps in the complement cascade (see Figure 1.1). In these subsequent steps, both C3bBb and C4b2a can collect one (or possible more) C3b molecule to become the trimolecular C5 convertases, $(C3b)_2Bb$ and C4b2a3b, respectively. These undergo a specificity shift from cleavage of C3 to cleavage of C5. C5 is a paralogue of C3 and C4 although it does not contain a thioester. Cleavage yields C5b and C5a; the latter is a powerful anaphylatoxin. C5b interacts with C6 and then the C5bC6 complex nucleates formation of the C5b,6,7,8,$9_n$ membrane attack complex that leads to cell lysis **(**Law and Reid, 1995**;** Morgan and Harris, 1999). Thus three major activities of

complement are opsonization, anaphylatoxin release and cell lysis (Blom *et al.*, 2004). In more recent years it has become clear that further breakdown products of C3b – *e.g.* iC3b and C3d – are powerful opsonins and ligands for other cell-surface receptors such as CR2 and CR3 that trigger various cellular responses including the stimulation of antibody production by B-cells.

### 1.2.1 Regulators of complement activation (RCA)

Regulators of the complement cascade maintain the precise balance needed between activation and repression of complement amplification (Liszewski *et al.*, 1996). This is achieved by several mechanisms including the regulation of C3 and C5 convertases. The most important site of intervention is the regulation of the alternative pathway C3 convertases. Complement regulators can either be membrane-bound (for example CR1) or in the fluid phase (for example factor H). (Kirkitadze and Barlow, 2001; Law and Reid, 1995**;** Koolman and Rohm, 1998**;** Morgan and Harris, 1999**)**.

The largest family of complement inhibitors are a group of homologous proteins called the "regulators of complement activation" (RCA). These proteins are composed of various numbers of domains called complement control protein (CCP) modules, short consensus repeats or sushi domains (Figure 1.1B). CCP modules are built up of approximately 60 amino acid residues, and each contains four cysteine residues that form two disulphide bonds (Barlow *et al.*, 1991**;** Norman *et al.*, 1991**,** Figure 1.1B). Each CCP module also has an invariant tryptophan, conserved glycines and prolines, and a region of low conservation termed the hypervariable loop.

The modules are connected to one another by linkers of between three and eight residues. (Blein *et al.*, 2004). Some neighbouring CCP modules are known to stabilise each other and some are known to cooperate in order to form specific binding surfaces

(Kirkitadze and Barlow, 2001). Therefore the functions of these proteins are the product of the component CCP modules and of the relative orientations and intermodular interactions between CCP modules. The number of linker residues is a critical factor in this respect. Given the importance of the regulators in controlling and directing complement activation, it is clear that any deficiencies in function could lead to diseases in humans (Law and Reid, 1995).

### 1.2.2 Information on some examples of the RCAs

Factor H is a soluble glycoprotein that is elongated and made up entirely from 20 CCP modules. It competes with the binding of factor B to C3b, accelerates decay of C3bBb in the alternative pathway and acts as a cofactor in the cleavage of C3b by factor I. A set of five FH-related proteins (between five and nine CCPs) are probably also involved in various aspects of complement regulation but are less well characterised. Membrane cofactor protein (MCP) (or CD46) is a transmembrane protein acting as a cofactor for factor I-catalysed cleavage of C3b and C4b. Hence both pathways of the complement system are regulated by MCP that is comprised of four CCP modules, an O-glycosylated serine/threonine/proline-rich domain, a transmembrane region and an intracellular region. Decay accelerating factor (DAF) (or CD55) accelerates the decay of C3/C5-convertases of the classical and alternative pathways. It is composed of four CCP modules followed by a serine/threonine-rich domain, and a GPI-anchor used for attachment to the membrane. Complement receptor type 2 (CR2) (or CD21) is a member of the RCA family but not in fact a regulator. It is a cellular receptor composed of 15 CCP modules, a transmembrane region and an intracellular domain. CR2 on the surface of B-cells links the innate and acquired immune responses by stepping up

antibody production upon binding to C3d-antigen complexes (C3d is the ultimate proteolytic degradation product of C3b and it remains covalently bound to the surface). CR2 is also known to be a receptor for the Epstein-Barr virus. C4b-binding protein (C4bp) is a soluble protein that acts similarly to factor H, but on the classical pathway. It is spider-like in structure with the major isoform made up of 59 CCP modules, grouped into seven α-chains (eight CCP modules per α-chain) and one β-chain (three CCP modules per β-chain) (Blom *et al.*, 2004). The C-termini of both chains have additional regions, which polymerise the single chains by disulfide formation to create the mature protein. CR1 is the final member of the RCA family and is the focus of this thesis.

### 1.2.3 Complement receptor 1 (CR1 or CD35), expression and functions

Complement receptor 1 is a cellular regulatory protein found on the surfaces of a variety of cells, although the majority is located on erythrocytes (Ahearn and Fearon, 1989). The CR1 gene is expressed by all peripheral blood cells except platelets, natural killer cells and most T lymphocytes; a small amount of soluble form founding plasma has been reported (Hamer *et al.,* 1998 ; Dantelsson *et al.,* 1994 ; Pascual *et al.,* 1993), but its physiological relevance remains unsubstantiated (Fearon, 1980; Tedder *et al.,* 1983; Yoon and Fearon*,* 1985). In tissues, it is expressed by follicular-dentritic cells, B lymphocytes, glomerular prodocytes and some astrocytes (Krych-Goldberg and Atkinson, 2001).

In its most common allotypic variant, CR1 is made up of 30 CCP modules, a transmembrane domain and an intracellular domain. Three of the 30 CCP modules

appear to be needed for each of three key C3b/C4b-binding site. The N-terminal 28 CCP modules are made up of four "long homologous repeats" (LHRs A-D), each consisting of seven CCP modules (Birmingham and Hebert, 2001; Klickstein *et al.,* 1987). The three well studied C3b/C4b-binding sites lie (one each) at the N-termini of LHRs A-C; CCPs 1-3 forms one region (called functional site 1), whereas the two other, almost identical, regions comprising CCPs 8-10 and 15-17 form two copies of functional site 2 (Goldberg *et al.,* 1989 ; Kalli *et al.,* 1991). Each of sites 1 and 2 encompasses a discrete functional activity (Figure 1.2B).

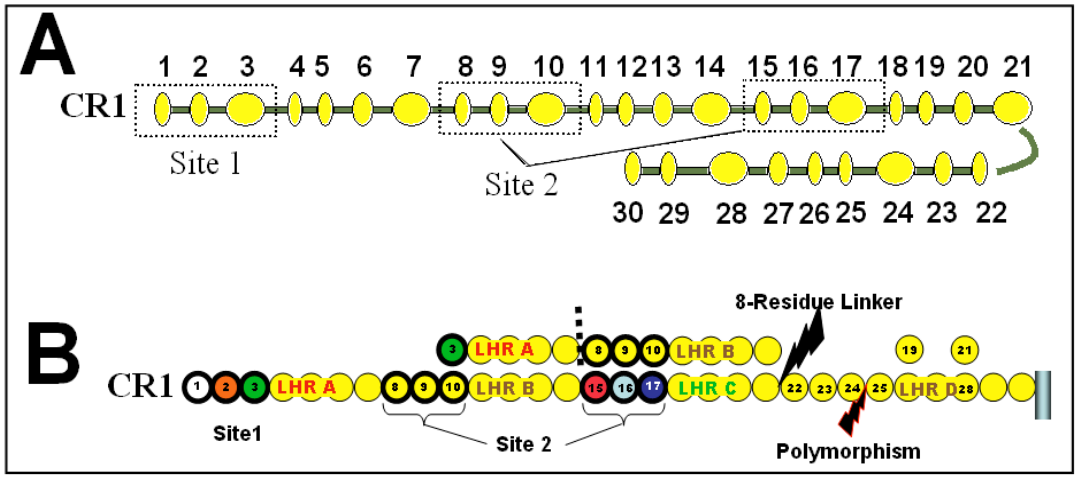



***Figure 1.2 Schematic presentations of different aspects of CR1***

(A) CR1 module and linker lengths. CCP modules are represented by oval shapes and their sizes are proportional to the number of residues they contain in excess of 50. Linker lengths are exaggerated to illustrate their variability. In this hypothetical structure the uniquely long (eight residues) linker between CCPs 21 and 22 (i.e. between LHRs C and D) is bent but in fact there is little reliable information available about the overall architecture of this 220-KDa glycoprotein. (B) CR1 internal homology and binding sites. The CCPs are represented by circles without reference to their variation in numbers of constituent residues. The diagram indicates the binding sites of CR1 – each is made up of the first three modules of a long homologous repeat (LHR). Modules placed on top of the full length CR1 are there to show the modules whose identities are greater than 90% (Portions of this diagram was modified from PhD thesis of Dr Dinesh Soares). Note that modules 3, 10 and 17 are nearly identical to one another; modules 3-14 are nearly identical to 10-21; these two triple-module fragments have the same functionality in complement regulation but interact differentially with *Pf*EMP1 (see text).

With regard to function, CR1 acts to regulate activation of both the classical and alternative pathways by serving (like MCP does) as a cofactor for factor I-mediated proteolysis of C3b and C4b to iC3b and iC4b (Krych *et al.,* 1994 **;** Subramanian *et al.,* 1996 ; Krych *et al.,* 2005) . It uniquely facilitates a further cleavage step of iC3b, to C3dg (or C4dg) (Ross *et al*., 1982 ; Medof *et al*., 1982; Iida and Nussenzweig, 1981; Medof and Nussenzweig, 1984); C3dg and C4dg are readily broken down non-specifically thus generating ligands for CR2 on B cells **(**Ricklin *et al.,* 2010**).** One product of C3b cleavage, iC3b, binds to complement receptor types 3 and 4 on phagocytic cells. CR1 also has decay-accelerating activity *i.e.* it speeds up the dissociation of all four of the convertases in the complement cascade (Farries, 1990 ; Fearon, 1979; Weisman *et al.*, 1990). By thus limiting the deposition of C3b and C4b, this decay-accelerating activity (Krych *et al.,* 1999) is thought to regulate the size of immune complexes *in vivo*, and prevent excessive complement activation. CR1 is also, importantly, the immune adherence receptor, that enables erythrocytes to harvest particles coated (opsonised) in C3b and/or C4b **(**Birmingham, 1995**),** and thus to ferry these opsonised entities to the liver and spleen (Krynch-Goldberg and Atkinson, 2001; Birmingham and Hebert, 2001; Dobson *et al.,* 1981) where they stimulate phagocytosis. In addition to the heavily studied C3b and C4-Binding sites CR1 has in the past been shown to bind both to C1q and to MBL, - initiating factors in the classical and alternative pathways respectively. There is also a possibility that they, too, contribute to CR1-mediated immune adherence (Klickstaein *et al.,* 1997 and Ghiran *et al.,* 2000). The region of CR1 implicated in binding C1q is found in LHR-D (Klickstein *et al.,* 1997)**.**

### 1.2.4 Functional localization within CR1

As mentioned above, functional sites 2 (i.e. CCP 8-10 and 15-17) is involved in C3b and to a lesser degree, C4b binding and has co-factor activity to factor I. Meanwhile, site 1 (CCP 1-3) binds C4b better than C3b. It is however involved in decay accerelation. Figure 1.3 indicates the modules involved in C3b, C4b, C1q and PfEMP 1 binding. It also shows that much more work will need to be done on CR1 to identify the specific region involved in the interaction with the other partners. Thirdly it shows that the binding sites of CR1 have different, although related, functions.

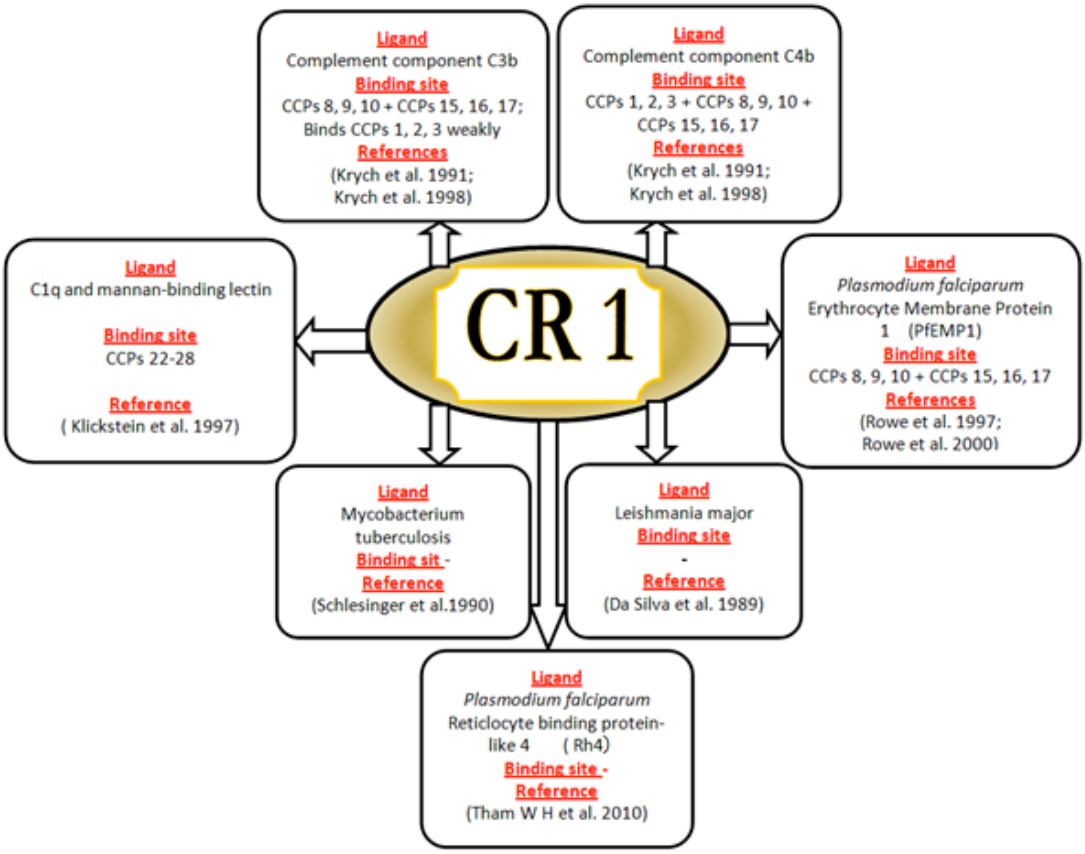



***Figure 1.3 Binding sites on CR1 for various ligands***

The CCP modules implicated in binding to each known ligand are given. Note that during the work described in this thesis it emerged that CCPs 1-3 are the major interacting partner for Rh4 – this is not shown on the diagram.

### 1.2.5 CR1 polymorphisms

Three main categories of polymorphisms are exhibited by the CR1 gene. Size variations or molecular weight polymorphism were the first to be discovered and these might have arisen from LHR duplications and deletions (Holers *et al.,* 1987 ; Cohen *et al*., 1989 ; Dykman *et al*., 1983 ; Thomas *et al.*, 2005)**.** The numerous size variants are believed to be the outcome of unequal gene crossover (Ahearn and Fearon, 1989). Four allelic forms of CR1, containing three to six LHRs, have been characterized. One of these, type A (or F), contains four LHRs and is the most common (82%) allele in all human populations studied to date (Xiang *et al*., 1999 ; Fearon *et al.,* 1989).

The second class of polymorphisms identified, correlate with the quantitative expression of CR1, or CR1 copy numbers, on erythrocytes (Ruuska *et al.*, 1992 ; Kazatchkine *et al.*, 1987). This *Hind*III restriction-fragment-length polymorphism is found to be common in Caucasians but rare or absent in Africans. Homozygotes for the L (low expression) allele usually express fewer than 200 copies of CR1, whereas homozygotes for the H (high expression) allele express several times this number; heterozygotes are intermediate (Gibson and Waxman, 1994 ;Wilson *et al*., 1986; Xiang *et al*., 1999; Herrra *et al*., 1998 and Rowe *et al*., 2002 ; Dervillez *et al.*, 1997). The well-studied Knops blood-group antigenic variation is the third category of polymorphism (Moulds *et al.,* 1991). These variants are the focus of the current project. The Swain-Langley (Sl) and McCoy (McC) Knops blood-group antigens, are all restricted to LHR-D of CR1 (Moulds *et al.,* 1991; Rao *et al.,* 1991).

African-derived populations are characterized by a slightly higher incidence of a larger size – variant of CR1 **(**Moulds, 2002**),** and higher copy numbers of CR1 on

erytrocytes (Rowe *et al*., 2002) as well as increased frequency of the S/2 and McCb alleles (Moulds *et al.,* 2001). For example, elevated frequencies of McCb and S/2 alleles have been correlated to resistance to Mycobacterium tuberculosis infections (Noumsi *et al.,* 2011**)**

### 1.2.6 Knops blood group polymorphism

The Knops blood group set of polymorphisms consist of Knops a and b ($Kn^a/Kn^b$), McCoy a and b ($McC^a/McC^b$), Swain-Langley ($Sl^1/Sl^2$) and Villien (Vil) (Daniels *et al.*, 2003 and Moulds *et al.,* 2001). The $Knp^{a/b}$ pair was initially identified among previously transfused Caucasian women, in whom an unknown antibody, whose loci distinct from those already known (ABO or Rhesus) was detected (Helgeson *et al.,* 1970; Moulds *et al.,* 2004). In contrast to the anti-$Kn^b$, which was reported to be better represented in Caucasian donors, antibody producers to this newly discovered $McC^a$ antigen were prevalent in people of black African origin. (Molthan, 1983; Moulds *et al.,* 2004; Molthan and Giles, 1975). Interestingly, $Kn^a$, which seems to be associated with $McC^a$ was found predominantly in caucasoids (Molthan and Moulds, 1978) and this suggested that the ethnic background seems to influence the description of the Knop's polymorphism. Significant subsequent studies revealed that the Knops blood group antigenic variations are located on CR1 (Moulds *et al.,* 1991; Rao *et al.,* 1991 ; Reid, 2004). Furthermore, $McC^a$ and $McC^b$ and $Sl^{1/2}$ and Vil are each associated specifically with single nucleotide polymorphism in CCP 25 of CR1 (Moulds *et al.,* 2001). Thirdly, there was a renaming of $Sl^a$ antigen into Sl1 and Vil into Sl2 (Daniels *et al.,* 2003**).**

# CHAPTER 1 INTRODUCTION

It emerged that both the Swain-Langley (Sl) and McCoy (McC) blood group polymorphisms are antithetical pairs. The McC alleles code for either glutamic acid (Glu) or lysine (Lys) at position 1590 (*i.e.* creating K1590 or E1590) as a result of an Adenine (A) change to guanine (G) at base-pair position 4795 (c4795A>G). Such a substituion, of an acidic glutamic acid sidechain for a basic one, might change the electrostatics at the surface of module 25 **(**Soares *et al.,* 2005**)**, if it corresponds to an exposed position or could disrupt an internal salt bridge if it were buried. Similarly, Sl2 polymorphism corresponds to the change of Arginine (Arg) to glycine (Gly) at amino acid position 1601 (R1601 or G1601), also as a result of an Adenine (A) change to guanine (G) but this time at DNA base-pair position 4828 (c4828A>G). This also corresponds to a non-conservative substitution that could have significant effects upon the surface properties or structure of CCP 25 and hence upon its interaction with other proteins (in the complement system or from pathogens, its self-associative properties on the cell-surface, or its intramolecular interaction with other CR1 modules that may play a role in the overall architecture of the protein. Clearly, such effects might be adaptive and one could envisage potentially major differences between for example E1590, G1601 and its allotypic variant K1590, R1601. (Thathy *et al.,* 2005 ; Moulds *et al.,* 2000) (See figure 1.9 for possible structural implications).

| Phenotype | Frequency (%) | | | Amino Acid | Correlation with malaria |
|---|---|---|---|---|---|
| | Caucasians | African-Americans | Africans | | |
| McC(a+)* | 98 | 99 | 82-95 | Lys 1590 | Unknown |
| McC(b+)* | 1 | 44 | 44-56 | Glu 1590 | Unknown |
| Sl:1* | 99 | 65 | 37 | Arg 1601 | |
| Sl:2* | 1 | 39 | 68-72 | Gly 1601 | [a]Reduced rosetting |

*Table 1.1 Knops blood group polymorphisms*
* McC = McCoy; Sl = Swain-Langley. McCa and McCb are one allelic antigen pair. Sl1 and Sl2 are another pair. The corresponding phenotypes for the first pair are McC(a+) and McC(b+), and for the second pair are Sl:1 and Sl:2. [a]There is reduced rosetting of Sl:2 RBCs with *P. falciparum* infected Table adapted from (Krych-Goldberg *et al*., 2002 ; Cockburn I A *et al*., 2004)

### 1.2.7 Location of Knops blood group antigens with respect to C3b/C4-binding sites

Each functional site 2 of CR1 (in modules 8-10 and 15-17) is known to bind to both C3b and (to a lesser degree) C4b (The structure of CR1 15-17 is solved, Smith *et al*. 2002. On the other hand, the Knops blood group polymorphisms occur within CCP modules 25 (Moulds *et al*., 2001). Despite their distance from the nearest copy of functional site 2 within the primary sequence, it is entirely feasible that the Knops blood group antigens are brought into physical proximity with modules 15-17 when the exodomain of native CR1 adopts an energetically favoured three-dimensional arrangement of its multiple modules. If that were the case than the Knops blood group antigens could contribute to a contiguous binding site for C3b (or C4b) consisting predominantly of surfaces contributed by CCP modules 15-17. Alternatively, they could bind to C3b at some other site than do modules 15-17 in a weak secondary

interaction that is not easily detectable but modulates affinity none the less. A further possibility is that they modulate interactions with C3b or C4 via an intramolecular interaction between modules 24/25 and modules 15-17.

A less likely possibility is that there are direct or indirect interactions between the Knops antigens and either functional site 1 (CCPs 1-3) or the first copy of site 2 (CCPs 8-10). Note, however, that the presence of the exceptionally long linker between LHR-C and LHR–D (*i.e.* CCPs 21 sand 22) provides a convenient "hinge point" (see below for a more extensive discussion of this point) in CR1 that is suggestive of a folding back of LHR-C onto LHR-D and hence proximity and potential cooperation between CCP 24/25 and CCPs 15-17 rather than the other functional sites. Finally it is conceivable that interactions between these apparently distant regions of the molecule could be intermolecular and be linked to self-association of CR1 – clustering of CR1 upon ligation has been reported to accompany ligation.

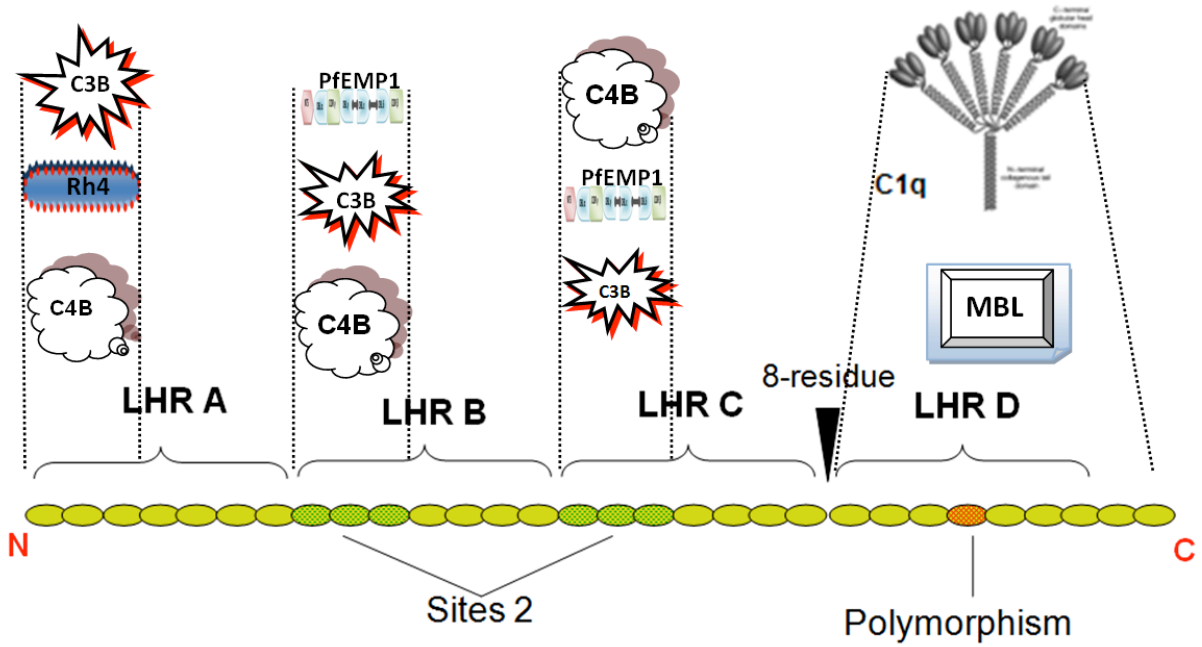



***Figure 1.4 Important regions of CR1***

This cartoon shows the positions of the binding sites for C3b, C4b, C1q/MBL as well as PfEMP1 and Rh4. It also draws attention to the location of the Knop polymorphism. In essence, the current project set out to explain how polymorphisms displayed so remotely with respect to the established binding sites could nonetheless modulate the latter's activities.

### 1.2.8 Locations of binding sites for C1q and MBL on CR1

Complement components C1q, and MBL, as well ficolins, function to eliminate invading microorganism, either by activating the classical pathway (C1q) or the lectin pathway (MBL and ficolin) (Ma *et al*., 2004). C1q, mannose binding lectin (MBL), and other members of the collagen family of proteins are pattern recognition molecules, able to enhance the phagocytosis of pathogens, cellular debris, and apoptotic cells in vitro and in vivo (Korb and Ahearn, 1997). C1q binds to antibody-antigen complexes and therefore actively mediates protection against infectious disease as an effective follow-up sequel to antibody production and binding to antigens. However, C1q can also recognize and bind to apoptotic cells (Korb and Ahearn, 1997), and may

participate in the clearance of autoantigens. Interestingly, a few reports have purported to show that C1q binds directly to CR1 (Klickstein *et al*., 1997), and suggested that C1q may participate directly in clearance of infected cells. If C1q does indeed assist in clearance of infected cell via binding to a site on CR1, then this might be releavant to the current polymorphism study. Importantly the putative C1q binding site was reported to lie in LHR-D **(**Tas *et al*., 1999**)**. MBL has also been found to bind CR1 (Ghiran *et al*., 2000, Sander *et al*., 1999) although as with C1q, the literature on this topic is thin and its physiological purpose unestablished. The presence of both the Knops blood-group antigens and binding sites for C1q (and possibly MBL) in LHR-D is clearly potentially pertinent to the hypothesis that the Knops blood group variants modulate some disease-related aspect of CR1 function and this requires investigation.

## 1.3 Malaria

Malaria is still responsible for millions of deaths worldwide, especially in sub-Saharan Africa, in spite of technological and economical progress in vector eradication, disease prevention and treatments (Breman *et al*., 2001 ; Holding and Snow, 2001 ; Sachs and Malaney, 2002). It is a parasitic infection caused by a protozoan of the Apicomplexan group of the genus *Plasmodia*. Among the species that infect man (*Plasmodium malariae, P. ovale, P. vivax* and *P. falciparum*), *P. falciparum* is the most devastating. Several factors account for the severity of the problem and the difficulty of eradicating this disease (Hviid, 1998). One of these factors is that the early signs of malaria (Birmingham and Hebert, 2001) look much like the symptoms of other diseases

(Warrell, 1993) so they are not always taken seriously. Second, diagnosis is quite difficult since it requires trained technicians to conduct the microscopy (Hanscheid, 1999 ; Warhurst and Williams, 1996 ; Weatherall *et al*., 2002). Third, the parasites responsible are mainly intracellular, hence effective treatment must be targeted at parasites within cells, which is challenging. In any case, the ability of the parasite to switch or change surface proteins leads to resistance to drugs (Noedl *et al*., 2008) and so frustrates drug designers. Meanwhile, the mechanism leading to severe anaemia is poorly understood (Stoute *et al*., 2003 ; Waitumbi *et al*., 2000) Moreover, the parasitic life cycle is complex**,** involving both man and mosquito Ross, 1897 ; Gilles, 1993); this implies that a really effective control strategy would be targeted at the relevant events in both man and the vector.

Studies have been made in different areas of malaria such as pregnancy associated and var2csa towards vaccine production (Dahlbäck *et al*., 2011; Sander *et al*., 2008 ; Arnot *et al*., 2001 ; Salanti *et al*., 2011 ; Pinto *et al*., 2011 ; Joergensen *et al*., 2009 ; Sander *et al*., 2011 ; Fried and Duffy, 1996)

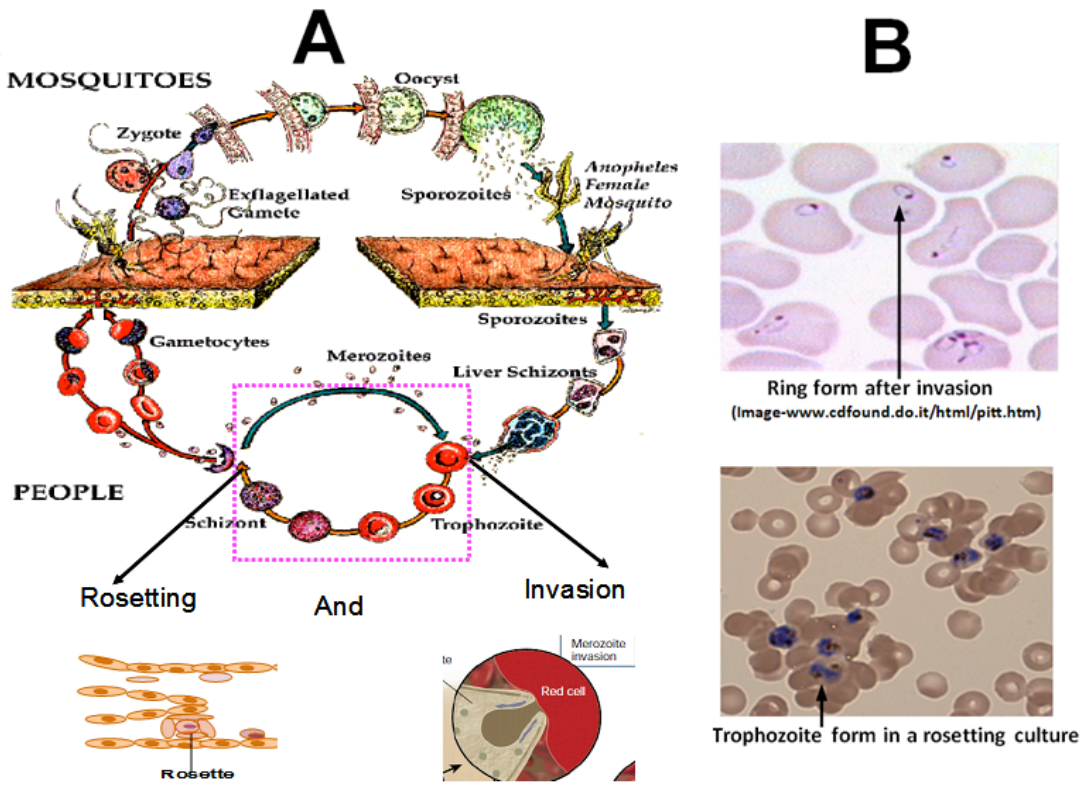



***Figure 1.5 Life cycle and life forms of P. falciparum***

**(A)** The pink dashed rectangular region and the black arrows indicate the part of the life cycle that will be of importance to the current study. A lot of activities take place in this part of the cycle, but two main phenomena are of particular interest: Rosette formation (as explained further in the text below) and invasion of erythrocyres**. (B)** Stained slide showing microscopic ring form and trophozoites of *P. falciparum*. (Image-www.cdfound.do.it/html/pitt.htm). The form of the parasite observed in the red blood cells after invasion is the ring form seen in the top slide; the lower slide on the other hand shows the trophozoite forms of the parasite that is associated with rosetting. (Picture taken during the author's parasite culture work).

### 1.3.1 Life cycle of *Plasmodium falciparum*

As the mosquito feeds, sporozoites, resident in the salivary ducts of the mosquito, are injected into the subcutaneous tissue and then into the bloodstream of the host (Miller *et al.,* 2002), before quite rapidly invading hepatocytes. The sporozoites then undergo nuclear divisions which results in the formation of liver schizonts over a period of 7-10

days. The parasites now in their merozoites incarnation are released into the bloodstream whereupon they very rapidly invade erythrocytes (Miller *et al.,* 2002, 2004).

Invasion of red blood cells in the human host by *P. falciparum* commences the erythrocytic cycle. Importantly, most pathological features of the disease are as a result of this erythrocytic cycle (Ramasamy, 1998). These include fever and chill which are a consequence of the waste products released from liver schizonts along with the merozoites (Mackintosh *et al.,* 2004). *P. falciparum* has the capability of adhering to venular endothelium and thus being involved in cyto-adherance. This is also the stage in the life cycle at which rosetting occurs during which infected and non-infected cells associate with one another in the blood stream *(Section 1.3.2)* with pathogenic consequences (Newbold *et al.,* 1997). Whilst this cycle lasts 48-72 hours in the case of most *Plasmodium spp*., for *P. falciparum* it usually lasts 48 hours (Biggs and Brown, 2001). After multiplication through several cycles of invasion, release and reinvasion, *P. falciparum* then enters the sexual stage of the cycle (Biggs and Brown, 2001).

Some merozoites in red blood cells differentiate into gametocytes (Biggs and Brown, 2001). These are picked up by the mosquito during a blood meal. In the stomach of the mosquito, fusion of gametocytes and meiosis occur. This leads to the production of a zygote, which has the ability to penetrating the midgut forming an oocyst. Asexual division of the mature oocyst then occurs to form sporozoites. The sporozoites migrate to the salivary glands thus completing a full cycle (Gilles, 1993).

To help reduce the morbidity and mortality of malaria, researchers over the years have put much effort into understanding various parts of its life-cycle and the

potential for interventions. Two aspects of the life-cycle that are of particular interest here are parasite invasion and the mechanism of rosette formation or "rosetting". Notably, severe malaria is often associated with rosetting of erythrocytes (Newbold *et al*., 1999).

### 1.3.2 Rosetting

Rosetting can be described as the spontaneous attachment of more than two uninfected erythrocytes to one or more *Plasmodium*-infected ones. This might lead to cell agglutination, disturbance in the free flow of blood and sequestration of the *Plasmodium* parasites in the blood stream. Many of the malaria parasites in endemic areas have been shown to undergo rosetting and rosette appears to be detrimental in various ways that might explain the connection with severe forms and often life-threatening forms of malaria (Newbold *et al.,* 1999).

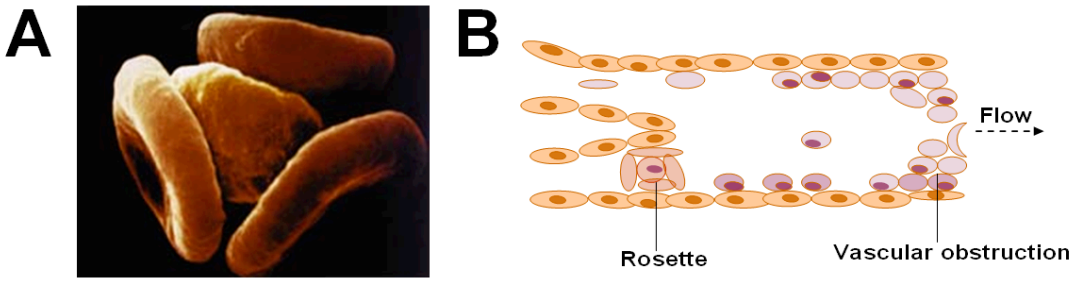



***Figure 1.6 Rosette formation***

**(A)** An infected RBC (centre) binding healthy RBCs causing clumping or agglutination called "rosetting" (image source: Texas Medical Center website). **(B)** Depicts formation of a rosette in the microvasculature. Rosetting might be a strategy used by the parasite to remain sequestered in the microvasculature so as to avoid destruction in the spleen and liver. Erythrocyte rosetting causes obstruction of the blood flow in microcapillaries. This may help to explain the observation that formation of rosettes correlates with severity of malaria (Rowe, Obeiro *et al.* 1995).

# CHAPTER 1 INTRODUCTION

Rosetting should be confused neither with cytoadherance nor with clumping, even though all these phenomena are associated with severe malaria. Cytoadherance is the attachment of infected red blood cells to the walls of blood vessels (Ho and White, 1999). Clumping, in the case of malaria, usually refers to the attachment of platelets to infected red blood cells (Kaul *et al*., 1991). Cytoadherance seem to occur in all forms of severe malaria, while rosetting is associated primarily with selected strains of *P. falciparum.* Rosetting, which normally occurs alongside cytoadherance, has been demonstrated to contribute to the obstruction of blood flow (Miller *et al*., 1994 ; Kual *et al*., 1991).

For rosetting to occur, the appropriate parasite factors and host receptors are clearly a requirement. Complement receptor type 1 (CR1) is among the major host receptors found to be involved in rosetting. *Plasmodium falciparum* erythrocyte membrane proteins 1 (PfEMP1) has been identified as the parasite ligand for CR1 (Rowe *et al.*, 1997). Some additional host factors that have been implicated in rosetting include the ABO antigen (Carlson and Wahlgren, 1992; Barragan *et al*., 2000), immunoglobulin M (Scholander *et al.*, 1996, Clough *et al*., 1998), and CD36 (Handunnetti *et al.*, 1992, Wahlgren *et al*., 1992); all are implicated in binding to parts of PfEMP1. The duffy binding-like 1α (DBL1α) domain of PfEMP1 has been mapped as the likely binding site of CR1 during the rosetting process (Rowe *et al.,* 1997; 2000).

Subversion of proteins that work within the host immune system is a strategy commonly used by pathogens (Lindahl *et al.*, 2000). Many of these pathogens bind complement receptors and complement regulatory proteins. This approach commonly facilitates entrance to host cells although in others, it modulates complement activation.

For instance, microorganisms as diverse as *Leishmania* (parasite), mycobacteria (bacteria) and HIV (virus) exploit complement receptor type 1 (CR1) for at least some aspects of cell entry (da Silva *et al*., 1989 ; Wyler *et al*., 1985 ; Schlesinger and Horwitz, 1990 ; Cooper 1998; Thieblemont *et al*., 1993 ; Munson *et al*., 1995). It is likely that a different purpose is served when PfEMP1 is expressed on the surfaces of infected erythrocytes. As briefly mentioned, rosette formation could be a strategy used by the parasite to remain sequestered in the microvasculature and to thereby avoid destruction in the spleen and liver (Wahlgren *et al*., 1989). An obstruction of the blood flow in microcapillaries is a well-established consequence of erythrocytes rossetting (Wahlgren *et al*., 1992).

### 1.3.3 PfEMP1 protein (family and binding partners)

The PfEMP1 protein is encoded by the *var* gene family. The number of *var* genes per parasite clone is estimated to be between 40 and 50. They are found on all chromosomes, usually at telomeres (Chen *et al.,* 2000). At any given time, the parasite expresses a single *var* gene but expression is subjected to sequential switches from one gene to another (Chen *et al.,* 2000, Miller *et al.,* 2002, Su *et al.,* 1995 ; Baruch *et al.,* 1995). This helps the parasite to avoid recognition by the host's adaptive immune system. Quite extensive work have been done on the var gene family, antigen switching and different parts of the variety of PfEMP have been expressed using different expression systems (Chen *et al.,* 1998; Wang *et al.,* 2009 ; Salanti *et al.,* 2002 ; Lavstsen *et al.,* 2003 ; Staalsoe *et al.,* 2002 ; Joergensen *et al.,* 2010 ; Victor *et al.,* 2010)

The PfEMP1 proteins range in size from 200 to 350 kDa. Their extracellular portions are organized into domains. An N-terminal segment is followed by a variable number of Duffy-binding-like (DBL) domains, cysteine–rich interdomain regions (CIDR), and sometimes, a C2 domain (Figure1.7). Based on sequence homology, DBL domains fall into five classes (α-γ) (Smith *et al.,* (2000)). In nearly all PfEMP1 proteins, the DBLα domain at the N-terminus is followed by a CIDRγ region (Figure 1.7). These N-terminal DBLα domains exhibit ~20% sequence identity across the PfEMP1 family (Su *et al.,* 1995). The remaining portion of PfEMP1 varies substantially.

The PfEMP1 protein is primarily responsible for the array of binding activities of parasitized erythrocytes (Chen *et al.,* 2000 and Miller *et al.,* 2002). This protein has been reported to interact with several types of soluble ligands and surface molecules, including intracellular adhesion molecule 1 (ICAM-1), type A and B blood group antigens, thrombospodin, E-selectin, chondroitin sulfate and CD36 as well as CR1 (Chen *et al.,* 2000) (see below).

The conserved sequence amongst N-terminal DBLα domains has been suggested to mediate binding to a common ligand, whereas variable sequences (in other domains) could be involved in recognizing a range of specific ligands. DBLα domains and CIDR regions both have adhesive properties. Both of these domains are implicated in rosetting, with DBLα being suspected of interacting with CR1 (Rowe, 1997; 2000).

To investigate the CR1-PfEMP1 interaction further, attention has focussed on the N-terminal DBLα but also on the adjacent CIDR. The sequences of DBLα-CIDR and DBLγ (as provided by Matt Higgens, University of Oxford) are shown in the

appendix (Appendix F). When over-expressed, the encoded DBLα had an expected size of 43 kDa. Since these portions of the protein are normally conserved, their structure would be expected to be similar to the structures of equivalent regions across the PfEMP1 family.

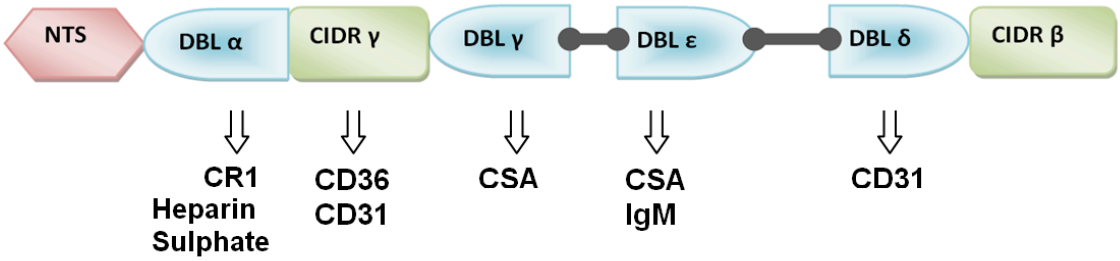



***Figure 1.7 Domains of PfEMP1***

Representation of *PfEMP1* (of R29) showing the suspected binding site for CR1 in DBLα as well as other putative binding sites.

### 1.3.4 The binding site on CR1 for PfEMP1

As discussed above, CR1 has been implicated in the rosetting of erythrocytes and thus could contribute to the development of severe malaria. In a previous study (Rowe, Moulds *et al.*, 1997; Rowe, Rogerson *et al.*, 2000) conducted to map the site of CR1 that binds to PfEMP1 in rosette formation, CR1 site 2 (specifically, modules 15-17) was implicated. Truncated fragments of CR1 that contained the proposed binding site brought about a reversal of rosetting in an experimental context (Rowe *et al.*, 1997 ; 2000). It is of particular interest that functional site 2 (see above) is also known to bind to both C3b (and C4b). As mentioned above there are two copies of Site 2 since modules 15-17 are virtually identical in primary sequence to modules 8-10.

As has been outlined in the earlier section that discussed the spatial relationship of C3b/C4b-binding sites and Knops blood group antigens, architectural features within the CR1 exodomain could allow module 24/25 to influence functionality at the proposed PfeMP1-binding site in CCPs 15-17.

### 1.3.5 Process of erythrocyte invasion

The survival of *P. falciparum* within the human host depends largely on the ability of that parasite to invade red blood cells, and to do so rapidly. This is a complex process that has not yet been fully understood. It has been suggested that the process begins with the specific binding of the merozoite to the erythrocyte through multiple parasite proteins. There follows a re-orientation of the merozoites, resulting in the apical end, which contains specialized secretory organelles, being adjacent to the erythrocyte membrane. There is then a fusion of the sub-cellular organelles with the merozoite membrane, leading to the organelle membrane proteins being placed on the surface of the merozoite. A junction between the two cells is then formed in which the surface of the merozoite interacts with the erythrocyte surface proteins using their ligands localized at the apical tip. Following engagement between parasite protein and host cell surface receptor, the parasitic organism actively enters the erythrocyte via a mechanism that has not been fully characterised. Invasion involves redistribution of the host membrane proteins away from the erythrocyte – parasite junction. A parasitophorous vacuolar membrane (PVC) forms around the junction, and the parasite, using actin-myosin generated force, moves into this organelle (Cowman *et al*., 2006). The multistep process can be summarised as involving initial contact of the merozoite with the erythrocyte, then an apical reorientation occurs so that there is a formation of a tight

junction that moves progressively towards the posterior end of the parasite until host cell membrane fusion is complete allowing cell entry (Cowman *et al*., 2006). When a merozoite invades an erythrocyte, it initiates the erythrocytic stage of the parasite's life cycle (Miller *et al*., 2002).

### 1.3.6 Invasion protein families and pathways

Two different gene families encode proteins that are important in *P. falciparum* invasion. One gene family, the erythrocyte-binding–like antigens (EBA), includes EBA-140/BAEBL, EBA-175, EBA-181/JESEBL and EBL-1. The second family is reticulocyte-binding–like homology proteins (RBPs or PfRhs) amongst which are PfRh1, PfRh2a, PfRh2b, PfRh4, and PfRh5 (Peterson and Wellems, 2000; Adams *et al*., 1992; Rayner *et al.,* 2000, Rayner, 2001). The EBA group are categorised as being involved in sialic acid–dependent invasion pathways on the basis that their interactions are sensitive to neuraminidase treatment of erythrocytes. On the other hand, the RBPs or PfRhs family are described as being in the sialic acid–independent invasion pathways for reasons contrary to those used to define the EBA group. Amongst the erythrocyte receptors that have been found to bind to *P. falciparum* invasion ligands in the EBA group are glycophorin A (GpA) interacting with EBA-175 (Sim *et al.,* 1994), glycophorinB with EBL-1 (Mayer *et al*., 2009), and glycophorin C with EBA-140 (Maier *et al*., 2003). GpA is a major glycoprotein found on human erythrocytes; it is heavily sialylated (Marchesi *et al*., 1972).

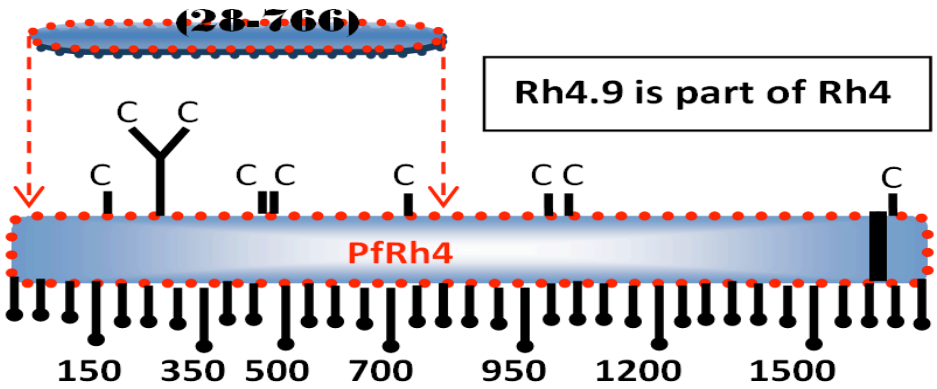



***Figure 1.8 A representation of PfRH4***

This is an example of a protein that is associated with the sialic acid-independent invasion pathway. The cartoon shows the region that corresponds to a recombinantly produced fragment of PfRH4.9 that presumably corresponds to one or more domains and has been recently shown to interact with CR1 in this invasion pathway. The Cs denotes cysteine residues, and the black lines with round heads shows the amino acid sequence numbers (Adapted from Tham *et al*., 2009).

Numerous studies suggested the interaction between EBA-175 and glycophorin A , (Adams *et al*., 1992; Camus and Hadley, 1985; Klotz *et al*., 1992; Orlandi *et al*., 1992; Sim *et al*., 1990), whose pathway (EBA-175/GpA) is chymotrypsin-resistant (Duraisingh *et al*., 2003a).

### 1.3.7 Possible connection between CR1 polymorphisms and Rh4-mediated invasion

While the majority of ligands for the EBA group of receptors had been identified prior to the current work commencing, the identification of CR1 as an important receptor for the sialic acid–independent invasion pathway, in multiple laboratory strains and wild isolates, came only recently. A still more recent paper (the result of a collaboration between the Cowman laboratory in Melbourne and our laboratory) presented strong

evidence that PfRh4 is the likely parasite binding partner (Spadafora *et al*., 2010, Awandare *et al*., 2011 ; Tham *et al*., 2010). Notably, when the gene for PfRh4 was disrupted, in the W2mef strain, the parasite lost its ability to switch invasion pathways *i.e.* was unable to invade neuraminidase-treated erythrocytes (Stubbs *et al*., 2005). Moreover, when anti-PfRh4 antibodies were included in a growth assay, it became even clearer that PfRh4 is the major ligand responsible for invasion through the sialic acid–independent pathways (in the region of 50–80%, depending on the parasite strain used) (Tham *et al*., 2009). In brief, when PfRh4 expression is activated, the parasite is able to switch receptor usage from sialic acid–dependent to sialic acid-independent pathways. This provides a mechanism for the parasite to invade by multiple alternative different pathways (Stubbs *et al*., 2005) and this facility presumably has evolved to assist evasion of the immune system.

Since invasion of erythrocytes is an essential component of the life cycle of *P. falciparum*, it has been crucial to identify all the parasite ligand, and their host erythrocyte receptors, that are used in invasion. This will help in appreciating the full repertoire of invasion pathways available to *P. falciparum* and in designing strategies for vaccine design or therapeutic intervention. Given that CR1 is an important receptor for PfRh4 and that this interaction is critical for invasion, it was hypothesised that the region of CR1 containing the Knops group polymorphisms might be able to influence binding to PfRH4. Were this to the case it would help explain the link between these polymorphism and the risk of developing severe malaria discussed above. Therefore, the current study set out to identify the critical regions of CR1 for binding to the parasite protein and also to measure affinities of the different variants for PfRh4.

### 1.3.8 Justification for CR1 polymorphism and structural study

The hypothesis to be addressed in this work clearly depends on the Knops blood group variants having functional consequences that must have a structural basis. With that in mind, the possible contribution to CR1 architecture of the long linker between LHRs-C and D was considered.

The two loci where the amino acid substitution occur (K1590E and R1601G) are only separated by ten residues and in homology-based models of CCP 25 (CCP-db website[1]) are surface exposed **(**Moulds, 2010**)**. Thus the side-chains in question (as outlined above) potentially interact with neighbouring CR1 modules, with host or parasite ligands or with other CR1 molecules. The middle cartoon of Figure 1.9A depicts a model of CCP 25 in which the surface-exposed region corresponding to the amino acid substitutions (R1601G) is highlighted in red. The flanking cartoons (Fig. 1.9A) demonstrate the loss of the large positively charged and relatively hydrophilic side chain of Arg1601 when it is substituted with small, neutral and relatively hydrophobic glycine (as in R1601G). (Leninger *et al.,* 2004; Moulds *et al.,* 2001). The Lys/Glu1590 variants are clearly different by two charges although both sidechains are relatively large and hydrophilic.

Thus if residues 1590 and/or 1601 occur within or close to a binding site of CR1, or in a region critical for key intramolecular (*e.g.* intermodular) contacts then quite drastic functional and/or structural consequences might result. Note, indeed, that the junction orientation between CCPs 24 and 25 could be affected by the K1590E

[1] http://www.bionmr.chem.ed.ac.uk/bionmr/public_html/ccp-db.html

change, since the substitution would have taken place just one amino acid before the first Cys (C1591) of CCP 25

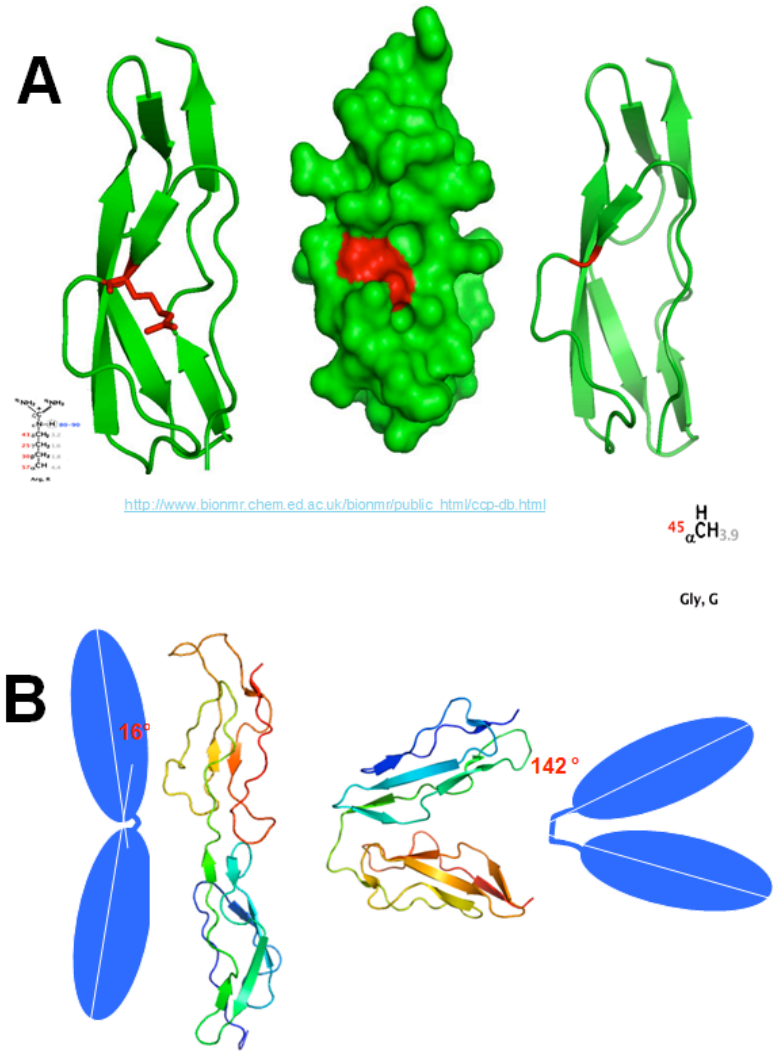



***Figure 1.9 Potential structural aspects of CR1 polymorphisms***

(**A**) 3D cartoons of polymorphic forms of CCPs 25 (**B**) Comparison of tilts between linker lengths. (Left) 4 amino acid residue linker of CR1 16-17, tilted at an angle of 16.3° whiles (Right) eight amino acid residue linker of CR2 1-2, titled at an angle of 142°

The linker between modules 21 and 22 stands out both for its unusual length and because of its strategic position between the LHR-C containing a C3b, C4b and

PfEMP1 binding site, and LHR-D which contains the McC and Sl Knops blood group polymorphisms. Although no study has so far implicated this linker in any known biological function, intuition suggests it has evolved for a purpose and that it could have a key role in the architecture of CR1. Another eight-amino acid residue linker is present between the N-terminal CCPs of CR2 and these modules were shown to adopt a V-shape with an intermodular angle of 142° (Figure 1.9B). Other module pairs, linked by three or four residues, are much more extended with a tendency towards being almost linear (see Fig. 1.9B). But to complicate matters, CCPs 12 and 13 of FH linked by eight residues form a bend but not a V-shape structure. Clearly, the orientation of modules 21 and 22 requires experimental investigation.

## 1.4 Hypothesis

To summarise the hypothesis is that the Knops blood group-polymorphism will have an effect on the way CR1 interacts with one or more of its ligands. This is easily testable in the case of known ligands such as C3b, C4b, C1q, PfEMP1 and Rh4.9. We also hypothesized that the CR1 molecule adopts a bent-back conformation that apposes CCPs 15-17 and CCPs 24-25. We finally conjectured that modules 24 and 25 might participate directly in interactions with ligands that have yet been identified (perhaps relevant to host-parasite interactions)

CR1 Knop polymorphism will make a difference in the way CCP 15-25 interacts with its binding partners

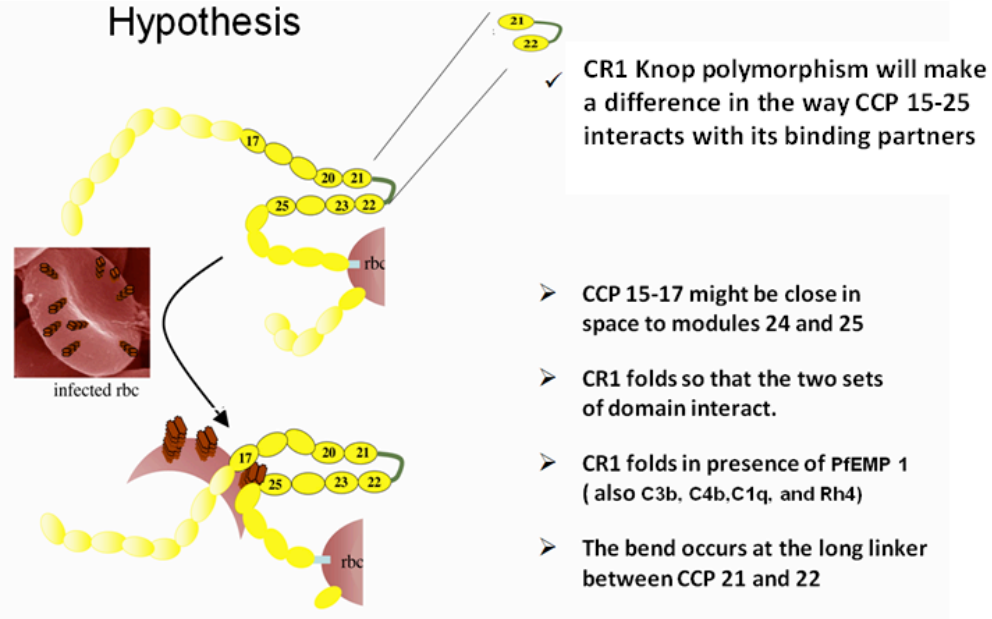



***Figure 1.10 Illustrates portions of the Hypothesis***

## 1.5 Aims and objectives

The aim of the study was to investigate whether polymorphisms, displayed at a remote part of the protein from previously mapped binding sites, can affect the function of the protein when interacting with its ligands. This will be done by: (1) Recombinantly producing relevant constructs of CR1; (2) Structurally testing whether the CR1 molecule in the LHR-C,D region has a bent-back conformation; (3) Biologically testing whether the allotypic variants bind differentially to ligands or exhibit differences in functional assays.

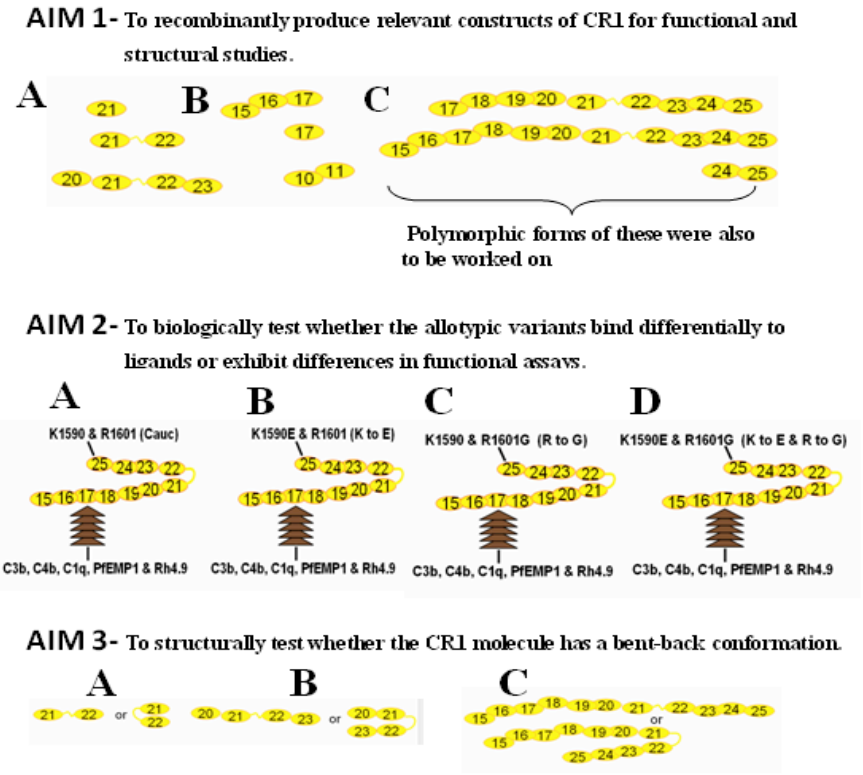



***Figure 1.11 Illustrate the steps involved in attaining the specific objectives***

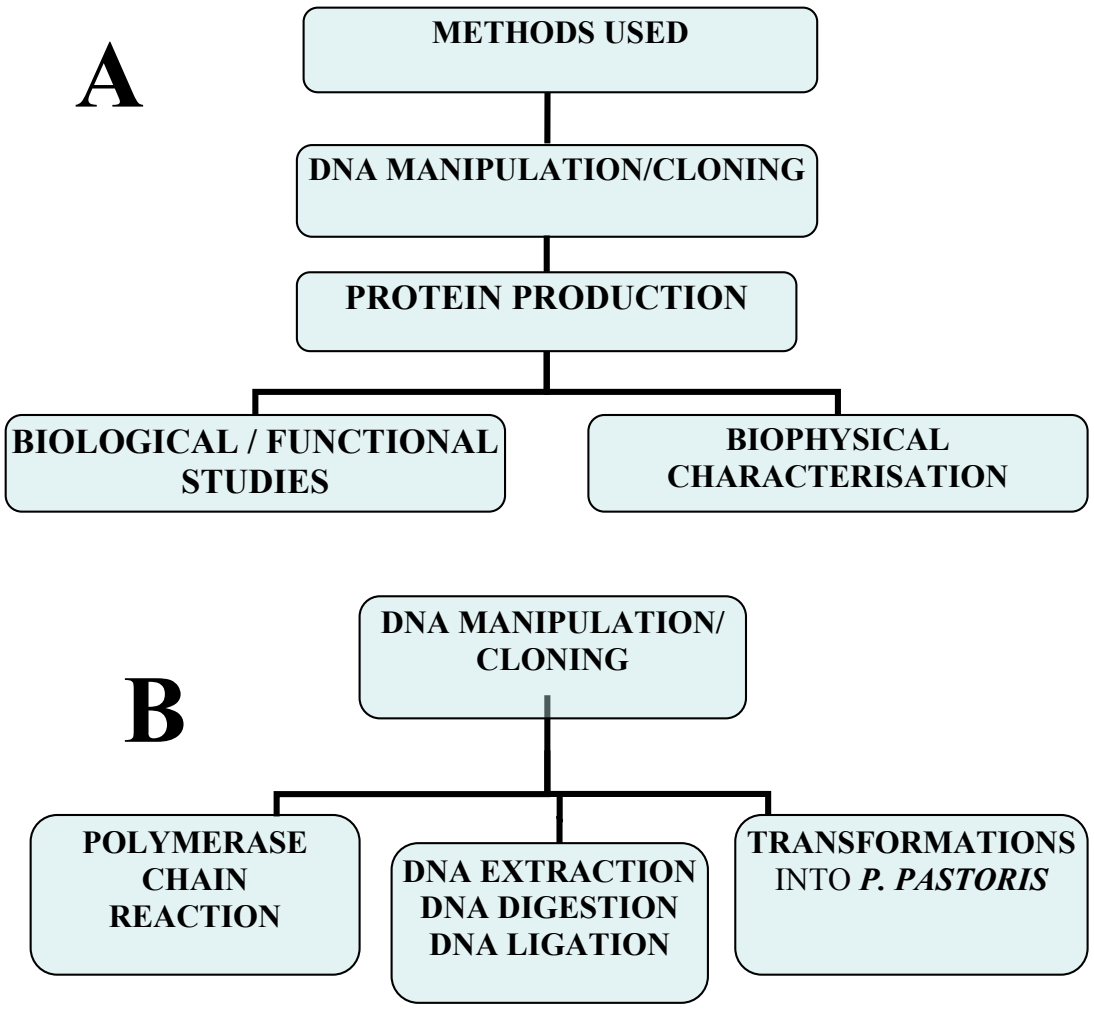
In all (Aim 1, 2 & 3), Yellow oval shapes represent CCP modules. Aim1 - Protein production and purification. Aim 2 – Biologically studies – Shape represent a bend back conformation. Whiles brown triangles represent their binding partners (C3b, C4b, C1q, PfEMP1 & Rh4.9). Aim 3 – Biophysical studies. Differents shapes in A, B and C represent extended (straight) and globular (curved).

# CHAPTER TWO
# MATERIALS AND METHODOLOGY

# CHAPTER 2 MATERIALS AND METHODOLOGY

## 2.1 Overview of methods

A range of methods were used in order to accomplish the goals of this project. The following flow chart (Flow chart 2.1) summarises the main activities.




***Flow chart 2.1 Main activities in the current project***
(A) Overview. (B) Methods employed in the molecular biology/DNA manipulation component of the work.

## 2.2 DNA manipulation

The techniques used in this component of the work include the polymerase chain reaction (PCR), DNA amplifications and extractions, "TOPO cloning" (Invitrogen),

DNA digests using restriction enzymes, DNA ligations, and transformations into *E. coli* and *P. pastoris* cells (Flow chart 2.1B). A summary of the steps involved, from creation of the insert encoding the protein of interest to the transformation of the expression plasmid into *P. pastoris,* is presented in Figure 2.1 and outlined below. Each step will be elaborated upon subsequently.

Note that human CR1 cDNA was provided by John Atkinson (Washington University Medical School) and oligonucleotide primers for PCR were ordered from Sigma or Invitrogen. The PCR-amplified (details in Section 2.2.5) DNA segments of choice were subjected to blunt-ended topoisomerase-mediated cloning into the vector pCR®4Blunt-TOPO (Invitrogen) followed by transformation into One Shot Top10 chemically competent *E. coli* cells (Invitrogen). Successfully transformed colonies were selected on ampicillin-containing Luria-Bertani (LB) media (agar initially and, later, broth). Following amplification of the TOPO vector in *E. coli*, it was extracted and double digested using restriction enzymes (purchased from New England Biolabs) (see Section 2.2.8 for details). The insert was then ethanol precipitated or gel extracted and ligated into a cut pPicZ αB expression vector (Invitrogen). After amplification in *E. coli*, the linearised (using *Sac*I, New England Biolabs) plasmid was transformed into the wild-type KM7I H strain of *P. pastoris* (Invitrogen) for protein production. Transformed *P. pastoris* colonies were selected on high-Zeocin containing media (yeast-extract peptone dextrose –YPD) to favour selection of those with high copy numbers. Mutations were introduced using a QuickChange site mutagenesis kit (Stratagene) in some cases to make it possible to produce polymorphic forms of the protein.

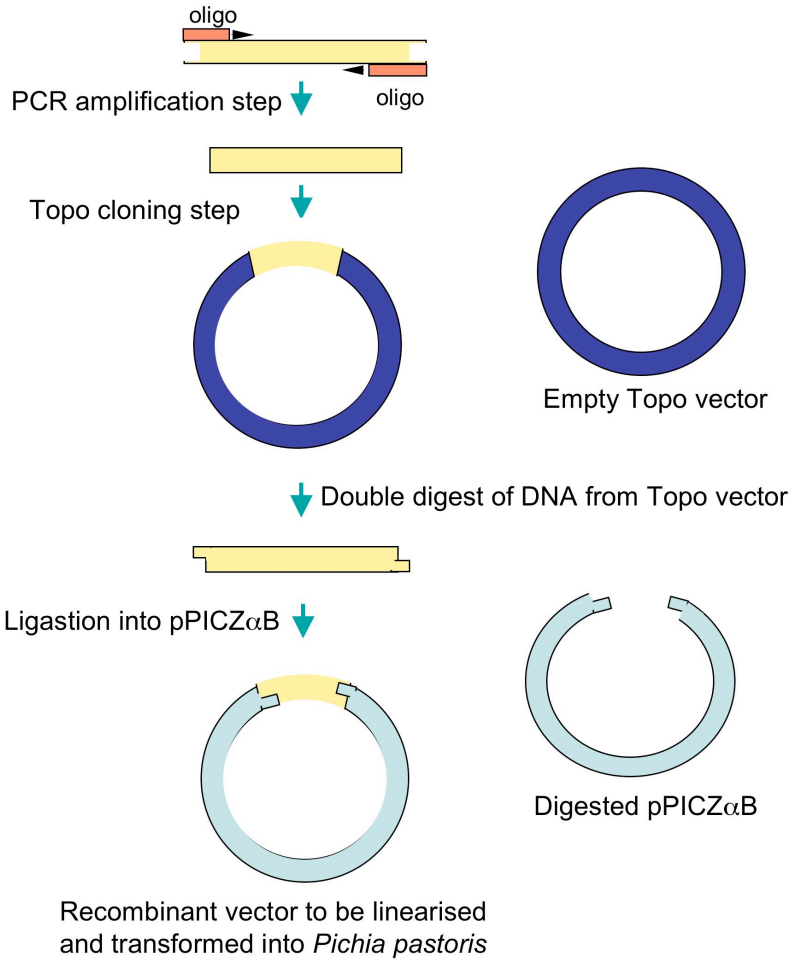



***Figure 2.1 Summary of cloning and transformations of DNA encoding the CR1 fragments***

Oligonucleotides/primers are shown in orange, the DNA insert is in yellow, the pCR®4Blunt-TOPO vector is dark blue, pPICZ αB vector is in light blue. From top: The amplification step involves the use of forward and reverse primers (orange, see Table 2.1 for sequences). The TOPO cloning step involved the amplified PCR product (yellow blunt-ended bar) and the pCR®4Blunt-TOPO ("empty Topo vector" – blue circle). After transformation into, and DNA extraction from, *E. coli* cells the "sticky-ended" DNA insert (yellow) is a result of the restriction enzyme digestion of the TOPO-cloned product (restriction enzymes used for the double digestion in this study were *Pst*I, *Xba*I and *Not*I). Ligation of the DNA fragment into the *P. pastoris* expression vector pPICZ αB required double digestion of the plasmid (light-blue incomplete circle), using the same restriction enzymes as used to create the insert. The resultant vector (light blue and yellow) is amplified (with transformed *E. coli* cells), extracted, linearised (using *Sac*I), and transformed into *P. pastoris* (KM71 H strain).

### 2.2.1 Primer sequences

Table 2.1 shows the sequences of primers used in this study. The restriction-enzyme sites are shown in lower-case letters, while the annealing region are in upper-case letters. The primer-pairs numbered 1 – 10 were designed (with the aid of the Sigma Genosys website, which has a DNA calculator used for cross-checking the parameters and characteristics of the primer) and purchased by the author, while Dr. Christoph Schmidt (University of Edinburgh) kindly provided the primer-pairs numbered 11 and 12. The latter two pairs of primers were mainly used for screening and sequencing after DNA fragments had been cloned into the TOPO vector or the *P. pastoris* cloning vector (pPicZ αB). The bold black line in Table 2.1 separates the primers without restriction-enzyme site (lower ones) from the primers that were designed with restriction sites (upper ones).

# CHAPTER 2 MATERIALS AND METHODOLOGY

| NO | PROTEIN CONSTRUCT | OLIGONUCLEOTIDE DESCRIPTION AND SEQUENCE | | RESTRIC-TION SITE |
|---|---|---|---|---|
| 1 | CR1 21 | 5’ | aactgcaggcGAACATATCTTTTGTCCAAATCC | Pst1 |
| | | 3’ | aactgcaggcGAACATATCTTTTGTCCAAATCC | Xba1 |
| 2 | CR1 21- 22 | 5’ | aactgcaggcGAACATATCTTTTGTCCAAATCC | Pst1 |
| | | 3’ | gctctagactaTCTACAGTTGTCTTCAACACTTGAC | Xba1 |
| 3 | CR1 20 -23 | 5’ | aactgcaggcTCCTGTGATGACTTCTTGGG | Pst1 |
| | | 3’ | gctctagactaGATCTCACAAATAGGTGCCTTC | Xba1 |
| 4 | CR1 17 | 5’ | aactgcaggcCGAATTCCTTGTGGGC | Pst1 |
| | | 3’ | gctctagactaGTTAGGTATAATGCACTGAGGG | Xba1 |
| 5 | CR1 10-11/17-18 | 5’ | aactgcaggcCGAATTCCTTGTGGGC | Pst1 |
| | | 3’ | gctctagactaCCTGGAGCAGCTTGG | Xba1 |
| 6 | CR1 17-25 | 5’ | aactgcaggcCGAATTCCTTGTGGGC | Pst1 |
| | | 3’ | gctctagactaGGAGCAGTGTGGCAGC | Xba1 |
| 7 | CR115-25 | 5’ | aactgcaggcCTGGGTCACTGTCAAGCC | Pst1 |
| | | 3’ | gcgcggccgcctaGGAGCAGTGTGGCAGCTTG | Not1 |
| 8 | CR1 24-25 | 5’ | aactgcaggcATATCTTGTGAGCCACCTCC | Pst1 |
| | | 3’ | gctctagactaGGAGCAGTGTGGCAGC | Xba1 |
| 9 | CR1 15-25 K1590E | 5’ | CCTCGGTGTATTTCTACTAAT**G**AATGCACAGCTCCAGAAGTTG | |
| | | 3’ | CAACTTCTGGAGCTGTGCATT**C**ATTAGTAGAAATACACCGAGG | |
| 10 | R1601G | 5’ | CAGAAGTTGAAAATGCAATT**G**GAGTACCAGGAAACAGGAG | |
| | | 3’ | CTCCTGTTTCCTGGTACTC**C**AATTGCATTTTCAACTTCTG | |
| 11 | Alpha-Factor & AOX1 | 5’ | GGGGATTTCGATGTTGCTGTTTTG | |
| | | 3’ | CCGGTCTTCTCGTAAGTGCC | |
| 12 | pUB/Bsd-TOPO | 5’ | GCAGCTTATAATGGTTACAAATAAAGCAATAGC | |
| | | 3’ | GGTAACGCCAGGGTTTTCCC | |

*Table 2.1 Primer sequences*

In primer-pairs numbered 9 and 10, the base pair to be changed using site-directed mutagenesis, is highlighted in red. The names of primers used in the QuickChange (Stratagene) site-directed mutagenesis kits are written in such a way that the amino acid substitution is indicated (*e.g.* K 1590 E). For the purposes of this study, these two pairs

of primers (K1590E and R1601G) were used to generate the McC and Sl Knops blood-group polymorphic forms of specific CR1 constructs such as CR1 15-25, CR1 17-25 and CR1 24-25. In all these cases, the K1590E and the R1601G amino acid changes were achieved by introducing A4795G and A4828G base-pair changes, respectively.

### 2.2.2 Amplification of coding sequences using PCR

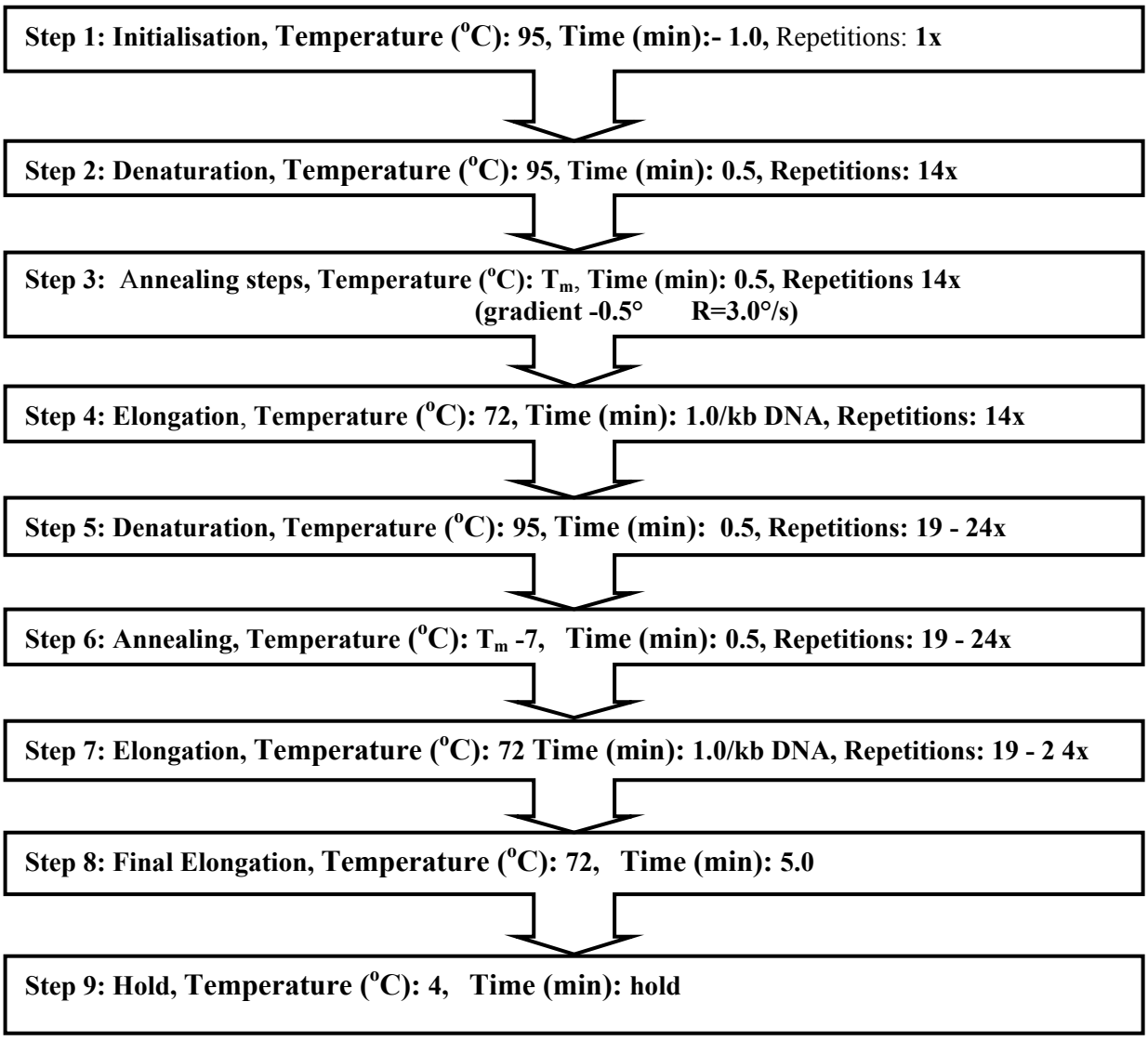
Amplification of the coding sequences is one of four applications of the PCR technique used in this study. The other applications are: introducing mutations needed to effect desired amino acid substitutions; screening of transformed cells; and sequencing of extracted DNA.

Materials required for this section included: oligonucleotide primers (see Table 2.1, Sigma or Invitrogen), human CR1 cDNA (provided by John Atkinson, Washington University Medical School), Herculase Pfu reaction buffer, Herculase HotStart and Turbo Pfu polymerase (all from Strategene), dimethylsulfoxide (DMSO), molecular biology-grade $H_2O$, and deoxynucleotide triphosphates (dNTPs) (from Roche).

The reaction mixture was usually made up of 1 µl each of forward and reverse primers (equivalent to 10 µmol), variable amount of template DNA (in the region of 40 ng), 5 µl Herculase or Pfu reaction buffer (the "10x concentrate" supplied by Stratagene), 2.5 µl DMSO (in the case of Herculase being used), 1 µl dNTPs (from the 10 mM solution supplied by Roche) and 0.5 µl of either Herculase HotStart or Turbo Pfu polymerase solutions (as supplied at 5 U/µl by Stratagene) using molecular biology-grade $H_2O$, the total volume was adjusted to 50 µl. If the yield of DNA

following PCR amplification was low, the procedure was repeated using this initial product as the template but the annealing temperature (see below) was adjusted.

A summary of steps in PCR-amplification of the coding sequences is shown in Flow chart 2.2, indicating the appropriate PCR-cycling parameters. The elongation time was adjusted to around one minute per thousand base pairs of DNA as per the protocol (Strategene manual). Non-binding regions, restriction-enzyme sites or mutation substitutions of the primers were all taken into consideration when the melting temperatures ($T_m$s) were estimated for use in annealing steps. When the two primers (forward and reverse) had dissimilar $T_m$s, the lower of the two $T_m$ values was used as an annealing temperature. Thus the annealing temperature and time were adjusted for each construct, according to the $T_m$ of the primers and the length of the construct.




***Flow chart 2.2 Program used for PCR for amplification of DNA***

### 2.2.3 Running of agarose gel electrophoresis

Agarose gel electrophoresis was performed after the amplification process to quantify the yield of the reaction, assess product purity, and check the product was running at a position commensurate with the expected size/molecular weight. This was needed to ascertain whether the PCR itself had worked and ensure the amplified DNA produce is not a contaminant.

# CHAPTER 2 MATERIALS AND METHODOLOGY

The materials used in running agarose gel electrophoresis included agarose obtained from Qiagen, tris-acetate EDTA (TAE) buffer (Tris base, glacial acetic acid, 0.5 M EDTA, pH 8.0, and dd$H_2O$), ethidium bromide/CYBR Safe **(**Invitrogen), a gel-running tank and its associated apparatus and a microwave oven or water bath, DNA molecular-weight markers (New England Biolabs), DNA (*e.g.* a PCR product, see previous section), loading buffer (Invitrogen) and a UV trans-illuminator (Ultraviolet Products). An appropriate amount of agarose powder was mixed with an appropriate volume of TAE (which was usually diluted from the 50X stock supplied) to make up a 1% (w/v) agarose gel. This was heated with gentle mixing at short intervals until the agarose was completely dissolved. An appropriate amount of ethidium bromide/CYBR Safe was then mixed gently with the dissolved agarose after it had been allowed to cool, but before it set (at about 50 °C). Generally, a final concentration of about 50 μg/l for ethidium bromide was sufficient to visualise the DNA.

The running tank was filled with TAE buffer, and the DNA samples (8 μl), mixed with 4 μl loading buffer, and were placed in the wells. Two DNA molecular-weight markers (New England Biolabs) were run simultaneously alongside the samples. The "100-base pair" marker was useful for DNA products in the range of 250 to 1500 base pairs while the "1-kb marker" was essential when the molecular weight of the DNA products were within the range of 500 – 10,000 basepairs. The gel was run at 80-100 V until good separation of the markers was obtained, but ensuring that the DNA did not run out of the gel. Using UV radiation from a trans-illuminator, the gels were visualised, and images captured for archiving with a digital camera.

### 2.2.4 Gel extraction and ethanol precipitation

These two procedures were deployed to clean up the DNA (*e.g.* the PCR product) and to separate it from other materials such as enzymes, primers, other unwanted amplified DNA fragments, proteins *etc.* Materials included; Qiagen gel-extraction kit, centrifuge (Sorvall legend RT, with SH-3000 swinging bucket rotor), sodium acetate, 100% ethanol and 70% (v/v) ethanol.

Gel extraction was performed before cloning using the protocol found in the Qiagen manual. A DNA agarose gel was run (as previously described) and after separating the desired DNA band from all other materials, the DNA fragment was excised from the gel with a clean, sharp scalpel. This gel slice was weighed, put into a colourless tube and then dissolved in three times as much volume of the buffer as the weight of the gel (*e.g.* 100 mg of gel dissolved in 300 $\mu$l of buffer) and incubated at 50 °C for 10 minutes, mixing at 2-3 minute intervals until the gel slice had completely dissolved. After checking to make sure that the colour of the mixture was yellow, one gel-volume of isopropanol was added to the sample with mixing. The sample was then applied to the QIAquick (Qiagen) column that had been placed in a 2-ml collection tube, and centrifuged for 60 s. The flow-through was discarded and the QIAquick column placed back into the same collection tube. After the QIAquick column has been washed, an additional centrifugation was done at 13,000 rpm for 60 s and then the QIAquick column was placed into a clean 1.5-ml microfuge tube for elution. This was achieved by adding 50 µl or 30 µl of the elution buffer to the centre of the QIAquick column, and centrifuging (13,000 rpm) for 60 s, after it has been allowed to stand for 60s.

Ethanol precipitation was used to concentrate the DNA and further clean it up. Before commencing, it was important that the volume of the DNA sample was noted, because that determined how much salt as well as ethanol was to be added. A 1/10th volume of 3 M sodium acetate, pH 5.2, was then added, followed by 2 to 2.5 volumes of cold 100% ethanol. After vigorously mixing by vortexing and placing on ice, or in the -20 °C freezer, for about 30-60 minutes, the mixture was centrifuged ( using a benchtop Eppendorf centrifuge 5415R, from Hamburg, Germany) at maximum speed for 15-30 minutes, and then the supernatant was carefully removed and discarded. Sufficient 70% (v/v) ethanol was then added to wash the residual salt from the pellet, after which it was centrifuged (13,000 rpm) for about 15 minutes and the supernatant was discarded. The pellet was re-suspended in water after it has been left to air dry.

### 2.2.5 TOPO® cloning of PCR product

The materials used in blunt-ended topoisomerase cloning included**:** the DNA product from the PCR, and the Invitrogen TOPO® cloning kit. Cloning was carried out according to the Invitrogen manual. To set the reaction up, a 2 µl aliquot of the PCR product was mixed with 0.5 µl of salt solution, 0.5 µl sterile $H_2O$ and 0.5 µl of the TOPO® vector as supplied. The solutions were gently mixed and incubated for 5-10 minutes at room temperature. A part of this 3.5 µl mixture containing plasmid DNA was then transformed into Top10 chemically competent *E. coli* cells as described below.

### 2.2.6 Transformation of Top10 E. coli cells

The materials used for this section were: Plasmid DNA (*e.g.* the product from the TOPO® cloning procedure described in Section 2.2.4), super optimal broth with

catabolite repression (SOC) medium, LB agar plates containing 100 μg/ml ampicillin or LB Lennox containing 25 μg/ml Zeocin, or LB broth containing 100 μg/ml ampicillin, shaking incubator (250 rpm, Stuart Scientific Orbital Incubator SI50).

The chemically competent Top10 cells (50 μl) were thawed on ice and 1-2 μl of the plasmid DNA was added and mixed by tapping the tube. After incubation on ice for 30 minutes, the cells were placed in a water bath at 42 $^{o}$C for 45 s and then quickly returned to the ice for 120 s. About 250 μL SOC medium was pre-warmed in the 42 $^{o}$C water bath and this pre-warmed SOC medium was added to the transformed cells and the mixture incubated at 37 $^{o}$C, for one hour shaking at 200 rpm. LB agar plates (pre-warmed to 37 $^{o}$C) were placed in the incubator and used for the spreading/plating of the transformed cells. When TOPO® vector was used for the cloning, LB agar plates containing 100 μg/ml of ampicillin were used for the spreading and incubation at 37 $^{o}$C overnight. Note that LB Lennox agar containing 25 μg/ml Zeocin plates were utilised when the vector involved was pPicZ αB.

### 2.2.7 Cultures and plasmid DNA extraction ("Minipreps" and "Maxipreps")

Materials for this part of the work included: QIAprep miniprep or QIAprep maxiprep kits for plasmid extraction and the QIAquick spin gel kit, all purchased from Qiagen.

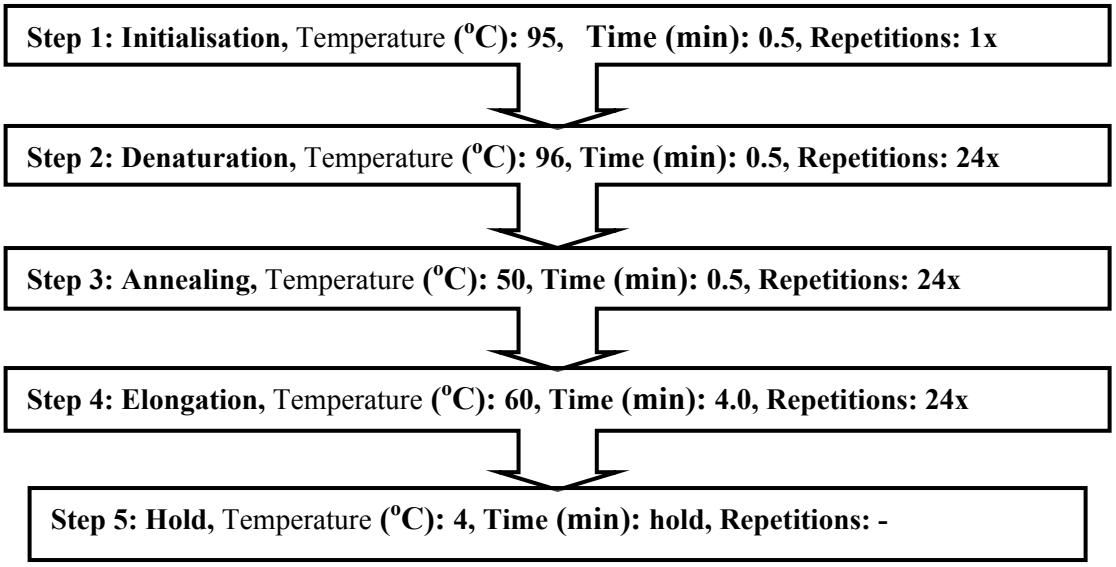
The culture and DNA extraction were carried out according to the protocol provided with the Qiagen QIAprep spin gel kits. After transformation and following overnight incubation on LB agar plates with 100 μg/ml of ampicillin in the case of TOPO® vector (but 25 μg/ml of Zeocin for the pPicZ αB vector), single colonies were picked and inoculated into 5 ml LB broth containing 100 μg/ml of ampicillin (for

TOPO vector and 25 µg/ml of Zeocin for pPicZ αB vector). For a small preparation ("miniprep"), an overnight incubation at 37 $^{o}$C, shaking at 225 rpm, was sufficient. For larger culture volumes (maxiprep) 100 $\mu$l of the 5-ml starter culture was re-inoculated into 50-250 ml of appropriate media (LB broth containing ampicillin or Zeocin for TOPO or pPicZ αB vectors respectively). The overnight culture of 5 ml (miniprep), or 50-250 ml (maxiprep), was pelleted by centrifugation (4,000x g for 5-10 minutes) and DNA extraction was performed as per the above protocol. The eluted DNA concentrations were checked by spectroscopy (Eppendorf BioSpectrmeter) at a wavelength of 260 nm, DNA was stored at -20 $^{o}$C.

### 2.2.8 Sequencing

DNA sequencing was performed at various junctures, for example after cloning into a different vector or site-directed mutagenesis. The materials needed for sequencing were primers (see above) and the ABI Prism dGTP BigDye Terminator Version 3.1 sequencing kit (Applied Biosystems).

The total reaction volume was kept at 20 µL using sterile water and the mixture was made of about 1 to 2 µl of the template DNA (150-300 ng), 1 µl primer (3.2 pmol), and finally 4 µl of the BigDye Terminator V3.1 solution as supplied. After mixing well (for 5 s), the program below was used for the PCR sequencing reaction.




***Flow chart 2.3 Program use for PCR prior to sequencing***

After the PCR reaction has been carried out, samples were submitted for automated sequencing to the service in the School of Biological Sciences, University of Edinburgh. The results were analysed using the BioEdit software downloaded from the (http://www.mbio.ncsu.edu/BioEdit/bioedit.html).

### 2.2.9 Restriction enzyme digests (double digestions)

Materials for this step included: Restriction endonuclease (*Pst*I, *Xba*I, *Not*I and *Sac*I), NEB buffer and bovine serum albumin (BSA) ( all from New England Biolabs).

The restriction enzymes were used to cut out the DNA insert from the TOPO vector at the appropriate restriction sites for re-cloning into the pPicZ αB. For a final reaction volume of 20 µl, 2 µl of the appropriate 10x concentrate NEB buffer was mixed with 16 µl of eluted plasmid DNA. About 0.2 µl of BSA was then added, and finally 1 µl of each of the restriction enzyme (20 U/µl) was added. The mixture was

quickly vortexed, centrifuged (using Eppendorf centrifuge 5415R, from Hamburg, Germany) and finally incubated at 37 $^{o}$C for 1-4 hours. Following electrophoresis of products on an agarose gel, extraction of cut DNA fragments was carried out using the QIAquick Gel Extraction Kit as previously described.

### 2.2.10 Ligation into pPicZ αB

Materials needed for this step included: pPicZ αB (Invitrogen), and 2x Quick Ligation Reaction Buffer and Quick T4 DNA Ligase, both purchased from New England Biolabs, and LB Lennox containing Zeocin (Autogen Bioclear)

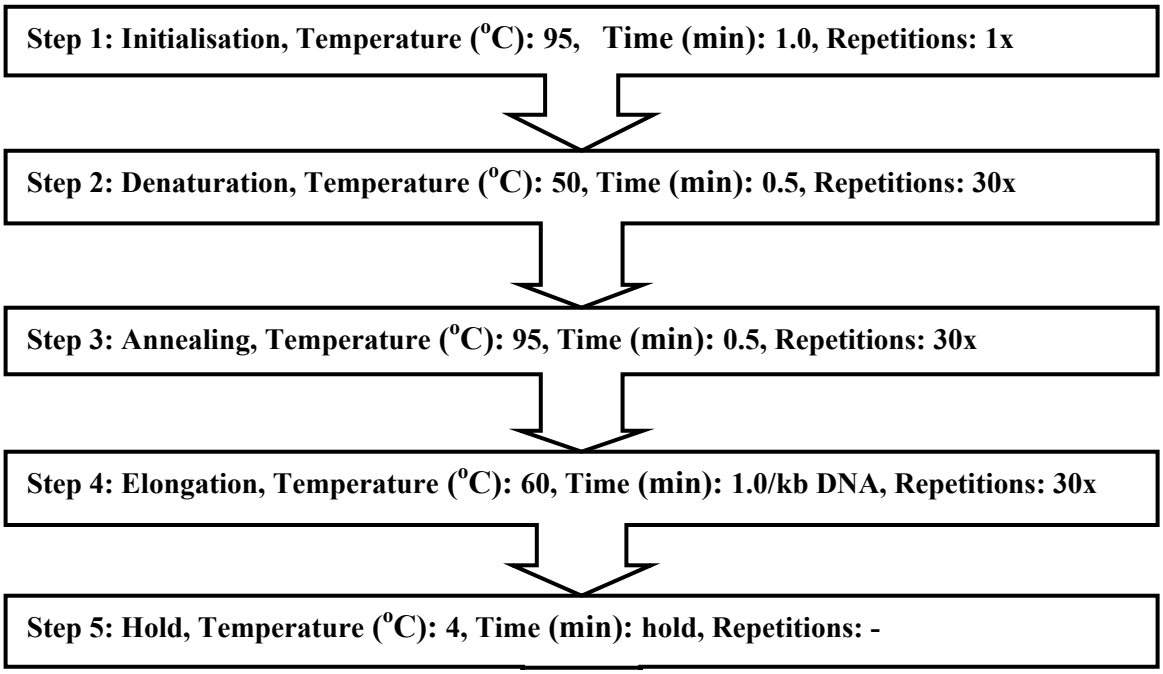
The ligation was performed according to the New England Biolabs manual where 50 ng of the cut vector was mixed with a three-fold molar excess of insert and adjusted with sterile water to a volume of 10 µl. Then 10 µl of 2X Quick Ligation Buffer and 1 µl of Quick T4 DNA ligase were added and mixed thoroughly. After centrifugation, the reaction mixture was incubated at room temperature for five minutes prior to Top10 *E. coli* transformation. Transformed cells were spread onto LB Lennox plus Zeocin (25 µg/ml) plates and incubated overnight at 37 °C. Transformation was performed as described earlier.

### 2.2.11 PCR Mastermix screening of transformed colonies

This PCR-based screening step was locally termed 'Mastermix screening' because it was performed with Mastermix (Promega), which contained the polymerase, dNTPs and the polymerase buffer. It was used in this study mainly to ensure that the DNA of the correct size had been inserted before sequencing. Apart from the Mastermix,

materials needed for this step were the oligonucleotide primers (Invitrogen or Sigma) and EB buffer (10 mM TrisHCl, pH 8.5, from Qiagen)

This screening was done both for the TOPO-cloned transformants and after cloning into pPicZ αB vector. Between four and six single colonies per cloned construct were picked from an overnight culture. Some were transferred onto a new plate while the rest were mixed with 20 µl of EB buffer in a labelled PCR tube. To lyse the cells and to release the plasmid DNA, the mixture was heated to between 90-100 °C for 5 minutes, and a 1-2 µl aliquot of this was taken for the screening reaction. The reaction mixture was made of 2 µl each of appropriate primers (M13 forward and reverse primers for the TOPO vector, and alpha-factor forward and AOX reverse primers for the pPicZ αB vector – see Table 2.1 for sequence), 1-2 µl of plasmid DNA and, finally, 5 µl of the PCR Mastermix solution. After mixing well, the screening was started using the PCR program below. This screening would not have been necessary if gel extraction after double digestion have been successful before re-cloning into the pPicZ αB vector.




***Flow chart 2.4 PCR screening cycle programme***

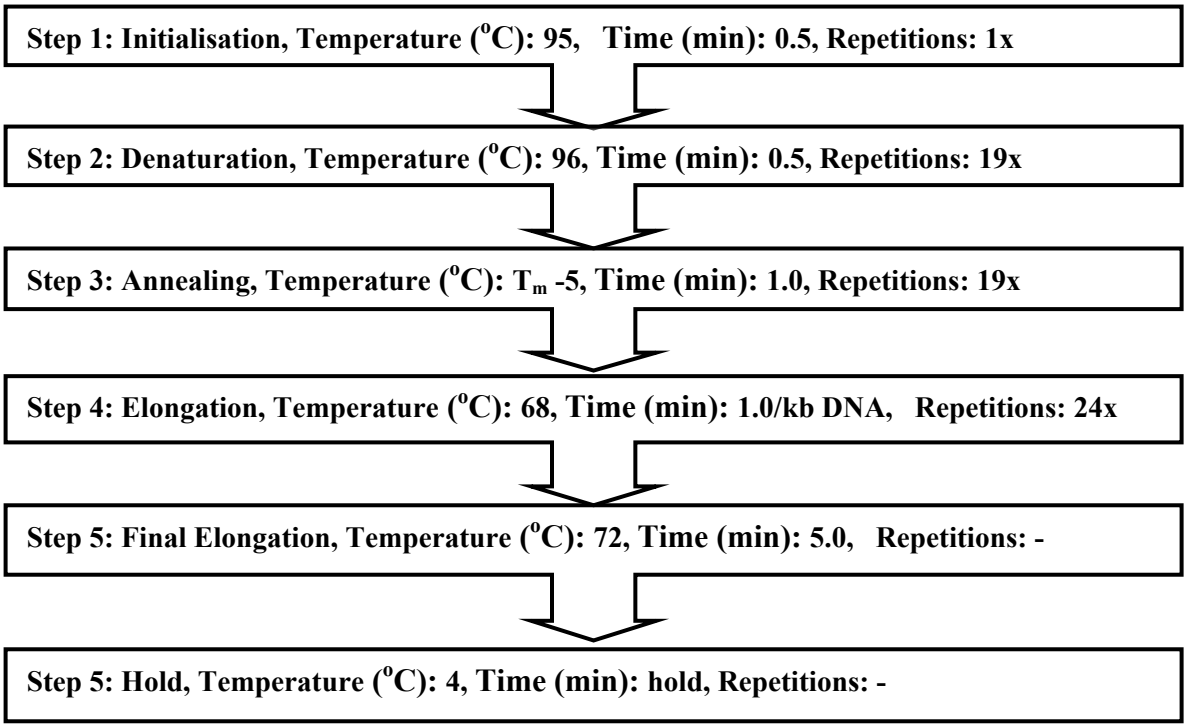
The screening products were analysed by running an agarose gel, after which samples that screened positive (*i.e.* showing a band of the expected size for the DNA fragment plus the additional amplified parts of the vector) were sent for sequencing. If the desired sequence were confirmed, it was selected for further work. For example some samples were used as templates for site-directed mutagenesis.

### 2.2.12 QuickChange site-directed mutagenesis

In this work, the QuickChange site-directed mutagenesis kit was used to introduce the nucleotide substitution needed to express the Knops blood-group polymorphic variations that occur in CCPs 24-25 of CR1. The template was therefore normally the DNA coding for the CR1 construct already cloned into the pPicZ αB vector. The

materials needed for this work included the QuickChange Kit (Stratagene), oligonucleotide primers (see Table 2.1), the template DNA (from the previously cloned CR1 construct), and *Dpn*I (10 U/µl) (New England Biolabs). The QuickChange kit mentioned contains the Pfu reaction buffer, Pfu polymerase, molecular biology-grade $H_2O$ and dNTPs.

The reaction mixture consisted of 1 µl of template plasmid DNA (normally about 40 ng), 1 µl each of the forward and the reverse primers (about 10 µmol each), 1 µl of dNTPs (from a 10 mM stock as supplied) and 5 µl of the 10x concentrated Pfu buffer. With the exception of the 1 µl Pfu Turbo polymerase (2.5 U/µl), which was the last reagent to be added, the volume was adjusted to 50 µl using molecular biology-grade $H_2O$. Immediately after the polymerase was added, a brief vortexing was performed, the sample was centrifuged and PCR carried out. The generalised form of the PCR program is shown in Flowchart 2.6.



***Flow chart 2.5 PCR cycling parameters for QuickChange site-directed mutagenesis***

The product was mixed with *Dpn*I (10 U/µl) and incubated for between one and two hours at 37 ºC to allow for digestion of the methylated template DNA plasmid prior to transformation into (XL blue) super-competent bacterial cells. Selected colonies were picked and cultured, and then DNA extraction and sequencing were performed as described above. DNA from the colonies that yielded good sequencing results was extracted using the QIAprep Spin maxiprep kit, linearised by *Sac*I-digestion and cleaned up using phenol-chloroform extraction and ethanol precipitation. (The procedure for the *Sac*I enzymatic digestion is similar to the one previously described above under enzymatic digestion).

### 2.2.13 Phenol-chloroform extraction

Materials for this step included: phenol, chloroform and isoamyl alcohol purchased from Sigma-Aldrich, chloroform and the linearised maxiprep product described above. The total volume of the linearised maxiprep product was noted and an equal volume added of commercially prepared phenol: chloroform (normally kept in the 4 ºC fridge). This was vigorously mixed by vortexing and spun down by centrifuging at maximum speed on the bench-top microcentrifuge (~13,000 rpm) for 120 s at room temperature. The aqueous phase (*i.e.* upper part) was transferred into a fresh centrifuge tube and another equal volume of commercially prepared phenol:chloroform was added. The centrifugation procedure and transfer of the aqueous phase into a fresh tube were repeated. Finally, an equal volume of chloroform was added, mixed, spun down and the aqueous phase was transferred into a fresh tube for ethanol precipitation. Ethanol

precipitation was performed as has been previously described. The purified linearised product was then transformed into *P. pastoris* KM71 H cells.

### 2.2.14 Pischia pastoris expression system of choice

The *P. pastoris* KM71 H strain (from Invitrogen) uses the catalytic ability of alcohol oxidase (AOX) to metabolise methanol and it belongs to the $MUT^S$ phenotype. $MUT^S$ has advantages over the other *P. pastoris* methanol metabolising phenotype, $MUT^+$ and was chosen as the expression system for this study. The *P. pastoris* system was picked for expression work in this case because as a eukaryote, yeast is able to form disulfide bonds (this can be problematic with *E. coli*-based systems). The chosen system facilitates the secretion of the desired protein into the supernatant, facilitating harvesting and purification. Secretion requires the α-factor signal-peptide. Prior to secretion this signal peptide is naturally cut off by endogenous enzymes that recognize the last amino acid residues of the signal peptide (EAEA) ahead of the desired protein sequence (Cereghino *et al.,* 2000, Cregg *et al.,* 1993, Clare *et al*., 1991). The cells can utilise methanol as a carbon and energy source and are induced to express the target gene (and produce and secrete the protein) by addition of methanol to the medium.

### 2.2.15 Transformation of P. pastoris KM71 H cells

The transformation was carried out according to the Invitrogen manual. The process started by inoculating 5 ml of YPD with the *P. pastoris* strain (KM71 H wild-type) and culturing overnight at 30 °C. Between 0.1 and 0.5 ml of the overnight culture was then put into a two-litre flask containing 500 ml of fresh medium and grown overnight to an $OD^{600}$ of 1.3-1.5. The cells were centrifuged at 1500 x g for five minutes at 4 °C and the pellet re-suspended with 500 ml of ice-cold, sterile water. The cells were

centrifuged as described earlier, the pellet re-suspended with 250 ml of ice-cold, sterile water, centrifuged again, re-suspended with 20 ml of ice-cold 1 M sorbitol, centrifuged again and then finally re-suspended in 1 ml of ice-cold 1 M sorbitol for a final volume of approximately 1.5 ml. The volumes were proportionally reduced depending on how much of the KM71 H wild-type cells were needed.

An aliquot of 80 µl of the competent cells of KM71 H prepared above was mixed with 5-20 µg of linearized DNA (in 5-10 µl tris-acetate EDTA buffer, TAE) and transferred to an ice-cold 0.2-cm electroporation cuvette. The cuvette containing the cells was incubated on ice for five minutes while making sure that the BioRad GenePulser II was set at 1500 V, 25 µF and 200 milliohms. Pulsing was done for about 6 s, and 1 ml of ice-cold 1 M sorbitol was immediately added to the cuvette. The contents of the cuvette was transferred to a sterile microcentrifuge tube and incubated at 30 °C for about 1-3 hours. Different amounts of this were pipetted and distributed onto pre-warmed YPDS plus Zeocin agar plates, spreading aliquots ranging from 100 to 400 µl over each plate. These were incubated at 30 °C until colonies had appeared. For plates containing between 100 and 300 µg/ml of Zeocin, successfully transformed colonies were usually observed by the third day. By the fourth day, isolated colonies were picked for protein-production trials and were also re-plated on higher Zeocin-containing YPDS agar plates.

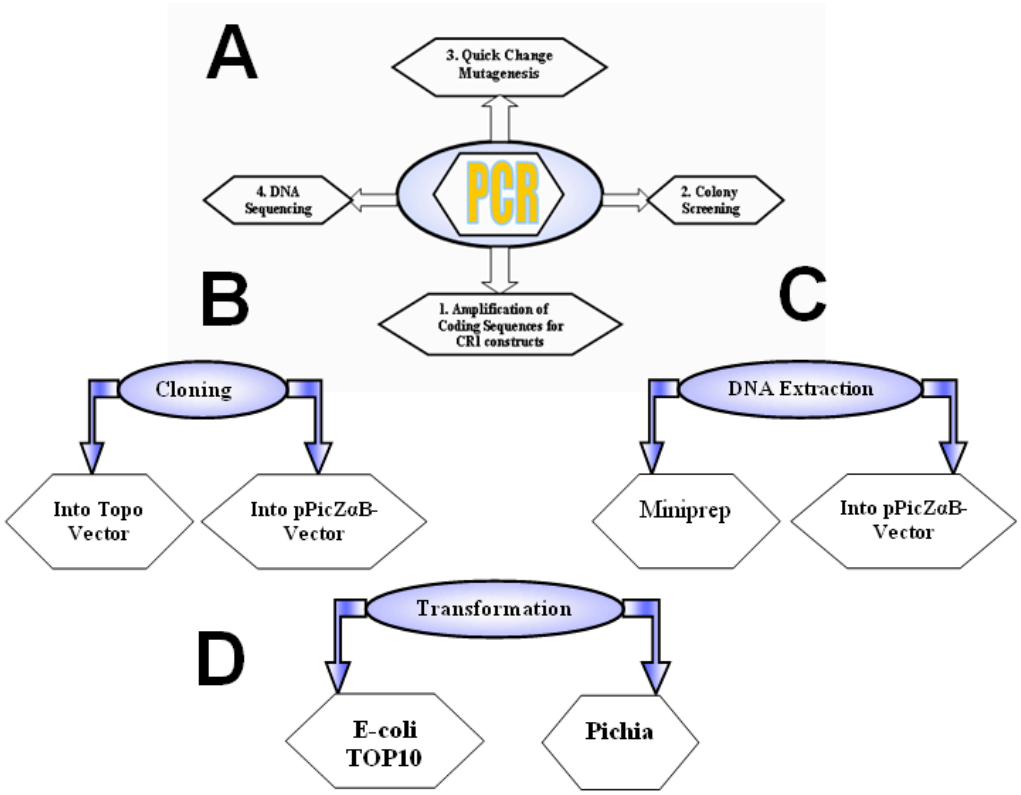


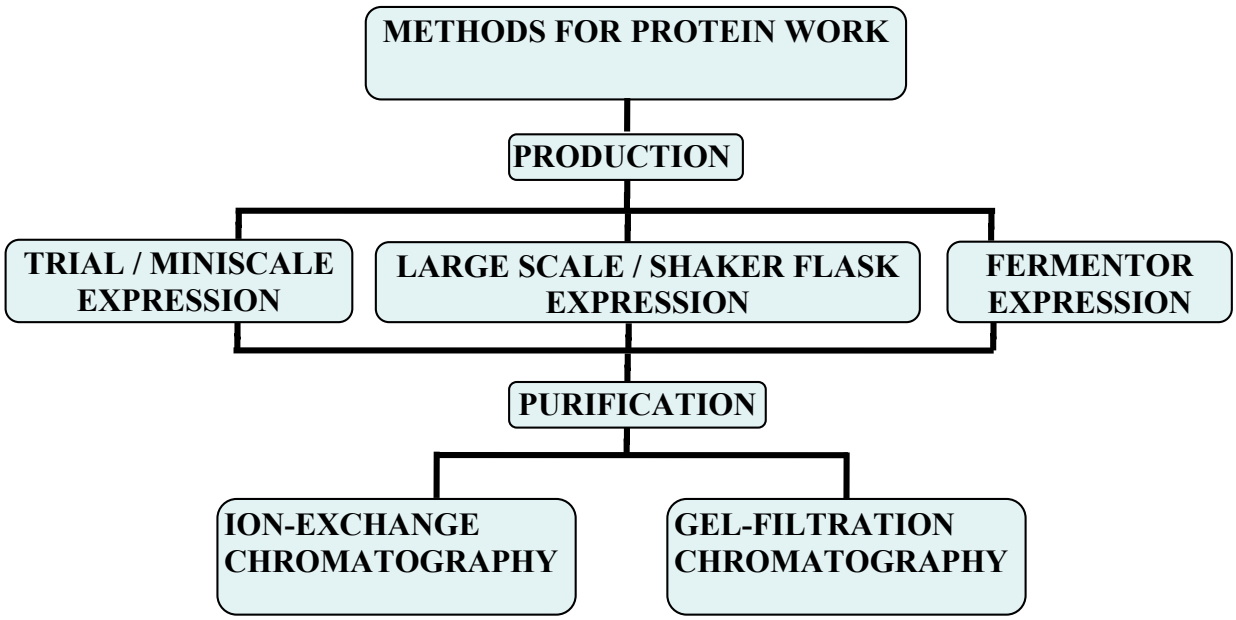

***Figure 2.2 Summary of some techniques used in the molecular work***

(A) Summery of different forms of PCR employed during the current study. These summaries are written and numbered 1, 2, 3 and 4 in hexagonal shapes around the central encircled PCR. (B) Two broad areas of cloning used in the study- TOPO cloning and cloning into the *P. pastoris* expression vector, pPicZ αB. (C) Forms of DNA extractions, employed in the current study – miniprep and miniprep. (D) Types of transformations used in this study – *E. coli* or *P. pastoris.*

# CHAPTER 2 MATERIALS AND METHODOLOGY

## 2.3 Protein production and purification

Methods employed for production of proteins include small-scale production trials ("miniscale"), as well as larger-scale production in shaker flasks and, ultimately, the fermentor. After harvesting, purification was carried out using ion-exchange as well as gel-filtration chromatography (see Flow chart 2.7). Materials needed for this section included; BMG (buffered minimal glycerol), BMM (buffered minimal methanol), baffled flask, methanol (100%), centrifuge, shaking incubator and fermentor.




***Flow chart 2.6 Overview of methods for protein production and purification***

### 2.3.1 Small –scale protein production trials

Protein production trials were carried out according to the guidelines in the manual provided by Invitrogen (with their EasySelect™ Pichia Expression Kit**)** with the need

for some minor modifications. Fifty-ml Falcon tubes containing 5-10 ml of BMG, or - in the case of low-yielding strains - small (250-mL) baffled flasks containing 80-100 ml of buffered minimal glycerol (BMG), were inoculated with an isolated colony (or with a small starter culture in the case of baffled flasks). They were then incubated for two days at 30 °C in a shaking incubator. Since *P. pastoris* cells have a high demand for oxygen for metabolism it was important that vessels were shaken at a speed of at least 300 rpm. Once cells had attained an $OD_{600}$ of between 2 and 6 the culture was spun down at 1500 x g for 5 minutes. The supernatant was then discarded and the pellet re-suspended in BMM (to 50-60% of original volume of BMG) for further incubation in the shaking incubator at 30 °C. The 0.5% methanol (v/v) in this replacement media induces expression of the inserted gene that is under the control (as explained above) of the AOX promoter. On each of the three subsequent days the cells were fed with methanol to give a maximum final concentration of 1% (assuming the methanol from a previous feed had been exhausted). On the day of harvesting the culture was spun down at 4000-6000 x g and the supernatant was filtered through a 0.2-µm filter. Generally after a trial protein production, some form of concentration is needed to help visualise (with Coomassie staining) the resulting recombinant protein on a polyacrylamide gel. Tricholoroacetic precipitation and spin concentration were used for this purpose.

### 2.3.2 Trichloroacetic acid (TCA) precipitation and centrifugal concentration

For TCA precipitation, an aliquot of about 1 ml of the supernatant from the "miniscale" production trial was pipetted into a microcentrufuge tube and then an equal

volume of 20% (v/v) trichloroacetic acid (TCA) was added, shaking gently to mix. The mixture was kept on ice for approximately 30 minutes prior to being centrifuged (4 °C, 12,000 rpm in Eppendorf centrifuge 5415R) for between 15 and 30 minutes. After the supernatant has been carefully removed and discarded, 300 µl of cold acetone was added to the precipitate and the sample was centrifuged for a further five minutes. After careful removal and discarding of the supernatant the pellet was air-dried. Using SDS-PAGE loading buffer the pellet was re-suspended, and after heating at 80 °C for three minutes, SDS PAGE was run with protein markers running alongside.

For concentration (as an alternative to TCA precipitation) the supernatant was applied to a Vivaspin concentrator and centrifuged according to manufacturer's instructions, being careful not to exceed the maximum g-force. This work made use of 0.5-ml, 6.0-ml and 20-ml concentrators with either 3-kD, 5-kD or 10-kD molecular-weight cut-off membranes depending on the size of the protein. The flow through (filtrate) from the first spin was discarded and then more supernatant was used to top up as necessary until all of the supernatant had been concentrated to the desired volume. The concentrate was then pipetted up and down to mix, transferred into a new microfuge tube and, following addition of SDS-PAGE loading buffer, subjected to SDS-PAGE.

### 2.3.3 Running sodium dodecyl sulphate-polyacrylamide gel elctrophoresis (SDS - PAGE)

The following materials were used: Pre-stained protein molecular-weight markers (BioRad), pre-cast polyacrylamide gradient gels (normally 4-12% NuPage Bis-Tris gels from Invitrogen but 4–20% Criterion gels from Bio-Rad were also used), BioSafe

Coomassie stain (from Bio-Rad), gel-running buffer (NuPAGE MES-buffer from Invitrogen and TGS-buffer from BioRad). The gel apparatus (consisting up of the gel tank, the gel itself, the running buffer and hardware supplied by Invitrogen) was assembled and the protein samples were mixed with the loading buffer that contained 100 mM mercaptoethanol, 2% (w/v) SDS, 0.1% (w/v) bromophenol blue, 50 mM Tri-HCl and 10% (v/v) glycerol.

These samples were then heated at 65 - 70 ˚C for three minutes, and then 20 μL was loaded into each well and the electrophoresis run at 180-200 V. Samples were run alongside pre-stained protein molecular-weight markers. The gel was washed (three times in water, changing the water every 10 minutes) and then Coomassie (from Bio-Rad) stained (by adding just enough stain to cover the gel and rocking the container) to visualise bands corresponding to resolved proteins.

Knowing the molecular weight of the target protein, SDS-PAGE provided insight into whether the desired protein(s) had been produced and in approximately what yield(s). On occasions where it still was not clear at this stage whether detectable protein production had been achieved or where there was doubt over the identity of a band, Western blotting was performed. For all miniscale trials, between 18 – 24 μl was loaded into wells depending on the type of NuPAGE gel. When the harvest was from a fermentor run, samples were not concentrated before loading on the gel, but a maximum volume was loaded. However, a concentration step was always employed after “miniscale” production trials or shaker-flask protein production trials. In the current study, therefore, between 14 and 18 μl of protein solution was mixed with a 4X loading buffer thereby filling wells that could accommodate between 20-25 μl. However, after purification (whether by ion-exchange or gel-filtration

chromatography), sample loading depending on the concentration of protein, and ranged from as low as 3 μl (variants) to as high as 14 μl.

### 2.3.4 Western blot

Primary murine antibodies against human CR1, specifically, 7G9 which recognizes CCPs 19-21, were provided by Professor Alex Rowe (University of Edinburgh). The choice of that antibody was based on its ability to recognise a common component of the longer constructs used in the current study (*i.e.* CR1 17-25, CR1 15-25, and their variants) that proved more difficult to produce in high yields than their shorter counterparts. Since CR1 20-23 had already been successfully produced in high yield (prior to making the longer constructs), and was recognised by the same antibody, it was selected as a positive control in a Western blot that screens for production of the longer constructs. Secondary antibody (anti-mouse, also gifted by Professor Rowe) was conjugated with horseradish peroxidase. Other materials used included: nitrocellulose membranes from BioRad, Towbin/transfer buffer (made from 3.03 g/l tris base, 14.4 g/l glycine , 100 ml-200 ml/l methanol), a Mini Trans-Blot cell (from BioRad), PBS, PBS plus 5% (w/v) non-fat dried milk (blocking buffer, locally prepared by author), PBS plus 0.05% (w/v) TWEEN20 from BioRad, SuperSignal West PICO-Chemiluminescent Substrate from Pierce, and Ponceau stain from Sigma-Aldrich.

In addition to sensitive detection of proteins produced at low levels, and authentication of candidate bands (particular those running differently to what might be expected from their molecular weight) western blots were used to search for the products of degraded target proteins. First, samples were run on SDS-PAGE as

described above, alongside pre-stained molecular-weight markers. The gel was quickly rinsed in water and briefly washed in Towbin buffer before preparing to transfer resolved proteins onto a membrane. At this stage, the nitrocellulose membrane, filter or blotting paper, and sponge or fibre pad, were all soaked in Towbin buffer.

A transfer "pack" was then assembled by sandwiching the gel and the membrane between two blotting papers and (on the outside) two sponges. Using a Mini Trans-Blot cell (from BioRad) and in Towbin/transfer buffer, proteins were transferred from the gel onto the nitrocellulose membrane either by applying a constant current of 150 mA for 2-3 hours or running it overnight at 30 mA in the cold room. To assess the success of the transfer, Ponceua staining of the membrane was used, although the transfer of the pre-stained molecular-weight markers had in any case suggested that the procedure had worked.

The membrane bearing the transferred protein was then immersed into about 100 ml of PBS containing 5% (w/v) non-fat dried milk for up to 2 hours and at 37 °C for blocking of non-specific interactions. The membrane was subsequently incubated (rocking platform) for about 2-3 hours in 5 ml of 1-in-500 to 1-in-1000 diluted primary antibody. The dilutions of both primary and secondary antibodies were done in PBS with 5% (w/v) non-fat dried milk. A multiple washing step of 10 minutes with PBS, 20 minutes with PBS plus 0.05% (w/v) TWEEN20, and 10 minutes with PBS was carried out following incubation with the primary antibody. The membrane was then incubated in 5 ml of a 1-in-1000 to 1-in-3000 diluted secondary antibody for two hours and washed as above.

Using the developing reagents from Amersham Biosciences, the secondary antibody conjugated to horse radish peroxidase was detected and the film developed in

the dark room by X-ray for a time ranging from a few seconds to a minute. The pre-stained molecular-weight markers were then transcribed or re-drawn onto the developed film for comparison.

### 2.3.5 Larger-scale (shaker flask) protein production

The procedure adopted for larger-scale shaker flask protein production runs or trials was similar to that of the "miniscale" expression described above. The main exceptions are that the 5-10 ml BMG culture in the "miniscale" protocol served as a starter culture for a larger volume (typically 500 ml in each of four baffled two-litre flasks). Because of the lag phase likely to occur following this seeding process, the number of days needed for the cells to be cultured in BMG was increased by one day. Subsequently it was found that the larger-volume cultures required longer spinning times prior to the induction step (10 minutes at 1500 x g). The required amount of spinning became even higher during harvesting of the target protein-containing supernatant (typically 8000 x g for 45-60 minutes).

### 2.3.6 The use of fermentors for protein production

Fermentation was carried out in an effort to increase the protein yield on the grounds that the better aeration and control of nutrients that is possible in the fermentor should result in higher cell densities. Equipment and materials needed included: a cylindrical Bioflow 3000 fermentor vessel and associated hardware purchased from New Brunswick Scientific, Antifoam 206 from Sigma-Aldrich, purity-grade fermentation trace mineral salts (PTM1 salt) from Amresco, materials for basal salt, glycerol, methanol and - in some cases - isotopic labels (see Sections 2.3.6.2 to 2.3.6.5). The fermentor had the advantage that in addition to being able to control the rate of

agitation and deliver oxygen or air from a compressed source, a range of probes were available that were used to monitor and control parameters such as dissolved oxygen, pH and temperature.

The fermentation process begins with a starter culture similar to what has been previously described under shaker-flask protein-production trials where a 5-10 ml culture was used to inoculate 200-400 ml of BMG in a two-litre baffled flask. The volume of this starter culture used was at least 1/10th of the final fermentation culture volume. While the starter culture was growing, the initial medium (basal salt) was prepared and poured into the assembled fermentor vessel, with all the probes attached for autoclaving. Although there were similarities between the various growths that were performed in terms of the preparation of the initial medium/basal salt mixtures, there were variations between fermentation runs in the actual contents and the details of subsequent feeding protocols. In particular where fermentation was being carried out to produce isotopically labelled protein the differences in the composition of media are detailed below (Sections 2.3.6.1 to 2.3.6.5).

After the autoclaved vessel, with its appropriate basal salt, and attached probes had cooled, the probes were connected to the electrical power source of the fermentor, enabling the disolved oxygen ($DO_2$) probe to be charged over night. While charging the $DO_2$ probe, air (from a compressed source) was introduced into the vessel through a sterile filter and agitation was set to 200 rpm. At this stage the growth medium was acidic and this retards growth of any contaminating organisms. On the following day, after the probe had been fully charged, the pH was adjusted to 5.0 by the addition of varying amount of base, through the base-feeding line and then the $DO_2$ probe calibrated. Then, 2.5 ml/l of high-purity grade fermentation trace mineral salts

(PTM1 salt, Amresco) and 0.5 ml/l of Antifoam 206 (from Sigma-Aldrich) were added to the medium and the temperature set to 30 ºC.

The fermentor was programmed to link the $DO_2$ level to the agitation rate, such that the $DO_2$ was maintained at 40% of saturation. Thus the instrument was set up so that the agitation rate automatically increased to a maximum of 1000 rpm, from the normal 200 rpm, when dissolved oxygen fell below 40%. This procedure helped in the monitoring of activity and growth of the cells because the agitation rate increased as nutrients were metabolised and oxygen levels fell, but decreased when nutrients were exhausted and no oxygen was needed.

Once all of the aforementioned had been set up and checked, the starter culture was spun down at 1500 x g and the pellet re-suspended in 10-40 ml of 100 mM potassium phosphate buffer, pH 6.0. The starter culture was then injected into the fermentor, following which a rapid reduction in the dissolved oxygen level occurs and a corresponding increase in the agitation rate was expected.

There were slight variations between fermentor runs in the times of induction and the subsequent feeding programme, depending on whether there was a need for isotopic labelling or not, as well as the type of labelling required. With the exception of $^{13}C$ labelling, which called for introduction of $^{13}C$-glucose, the protocol was as follows. Induction normally took place after two days of cell growth on 1% (v/v) glycerol, mixed with PTM1 salt (4.35 ml/l). For induction, temperature was reduced to 15 °C and expression induced by addition of methanol mixed with (4.35ml/l) PTM1 salts, to a final concentration of 0.5% (v/v) of the culture volume. For three subsequent days, methanol mixed with 4.35ml/l of PTM1 salts was added to a final concentration of 1.0% (v/v) of culture volume and to about 1.2% (v/v) (in some cases) on the last day.

Sometimes, feeding was carried out twice a day depending on the metabolic activity within the culture, as indicated by the agitation rate and the inferred $DO_2$ level.

Finally, at the time of harvesting, the contents of the fermentor were retrieved and cells removed by centrifugation in a similar way to that described above for the larger-scale shaking flask production runs, except that before the 45-minute 8000 x g spin, a preliminary spin at about 4000 x g for 15 minutes was performed. This extra step helped avoid the contamination of the supernatant by cells that would otherwise frequently block the 0.2-µm filter used to clarify the supernatant after centrifugation.

#### 2.3.6.1 Fermentation without isotopic labels (production of non-labelled protein)

For producing non-labelled protein by fermentation, 1 l of initial medium was made from; 27 ml of 85% (w/v) phosphoric acid, 0.95 g of $CaSO_4$ x7.$H_2O$, 15.0 g of $MgSO_4$ x 7.$H_2O$, 18.2 g of $K_2SO_4$, 4.2 g of KOH, 25 ml of glycerol; water was added to make up to 1 l and the pH was adjusted to pH 5 by addition of 4% $NH_4OH$ solution. Note that ammonium hydroxide was used in the non-labelling fermentations as both the base for regulating the pH and as a nitrogen source, while glycerol was the carbon source.

#### 2.3.6.2 Fermentation with isotopic labels

For fermentations intended to produce isotopically labelled proteins, the initial medium was made of: 0.95 g $CaSO_4$ x 7.$H_2O$, 12.0 g $MgSO_4$ x 7.$H_2O$, 6.0 g $K_2SO_4$, 60 ml 1 M potassium phosphate buffer, pH 6.0, and water to make up to a total of 600 ml of medium. A 2 M KOH solution was used as a base for pH adjustment during the fermentation.

### 2.3.6.3 $^{15}$N flabelling (single labelling) of proteins in the fermentor

For $^{15}$N labelling of the target protein, about 8 g of $^{15}$N-$(NH_4)_2SO_4$ (Sigma-Aldrich) was dissolved in about 10 ml $H_2O$ and sterile-filtered through 0.2-µm filter then (normally) added to the culture at the same time as inoculation. Note that in this case glycerol served as the carbon and energy source.

### 2.3.6.4 $^{15}$N and $^{13}$C labelling (double labelling) of proteins

For $^{15}$N and $^{13}$C double labelling of the target protein, about 7 g $^{15}$N-$(NH_4)_2SO_4$ was added to 15 g of $^{13}$C-glucose (Sigma-Aldrich), dissolved in about 40 ml of $H_2O$ and the solution was filtered, as previously described, to sterilise it. The culture was fed with the $^{13}$C-glucose for two days after which a 1g feed of $^{13}$C-glycerol (Sigma-Aldrich) was providedn. Subsequently, the cells were starved for about 3-4 hours in preparation for induction by addition of $^{13}$C-methanol (Sigma-Aldrich) to a final concentration of 0.5% (v/v). Thereafter $^{13}$C-methanol feeds were delivered as previously described for the non-labelling fermentations.

### 2.3.6.5 $^{15}$N and $^{2}$H labelling (deuteration) of CR1 20-23

The procedure employed for producing a deuterated sample was similar to that used for $^{15}$N labelling**,** except that "light" water ($H_2O$) was replaced by "heavy" water ($D_2O$). Thus, to the greatest extent possible all solutions of reagents and nutrients were constituted in 98% $D_2O$ (Sigma-Aldrich). The starter culture was also grown up in 98% $D_2O$. Cells in the fermentor grew more slowly in $D_2O$ due to the "isotope effect" as

expected, and therefore an additional day for cell growth prior to induction was allowed.

### 2.3.7 Protein purification

Prior to protein purification, phenylmethyl sulfonyl fluoride (PMSF) and ethylenediaminetetraacetic acid (EDTA) were routinely added to final concentrations of 0.5 mM and 5 mM, respectively, to deter proteolysis. Some chromatography columns and resins were purchased from Amersham while, unless stated otherwise, all pre-packed columns and affinity resins were purchased from GE Healthcare.

#### 2.3.7.1 Chromatography on bench-top self-poured columns

The first purification step entailed low-resolution (step-gradient) ion-exchange chromatography on a bench-top column (Econo-Pac, volume = 5-10 ml, BioRad) packed with SP Sepharose. The main goal of this step was to extract the target protein semi-selectively out of the large volume of the growth medium and concentrate it into a much smaller and more manageable volume. The eluant was monitored at 280 nm using the Eppendorf BioSpectrometer. Prior to loading on the column, the crude supernatants from fermentors or shaker flasks were diluted 1-in-5 (to reduce ionic strength) and the pH adjusted to one at which the protein would be expected (basd on theoretical pI values) to bind to the resin. After washing at a typical flow rate of 5 ml/min (using a pump to deal with volumes between 1 and 4 litres), two or three column-volumes of low-salt buffer (normally 20 mM sodium acetate, pH 4 but this

depended on the calculated theoretical pI). Elution was done using the same buffer but containing 1 M NaCl.

#### 2.3.7.2 Desalting

After bench-top step-gradient ion-exchange chromatography, the proteins had been eluted in a concentrated form but in 1 M NaCl solution, and therefore a desalting step was required prior to the subsequent higher-resolution ion-exchange purification step. Desalting was achieved when the eluted samples from the previous purification were applied to a 5-ml HiTrap (Amersham) desalting column. A 20 mM sodium acetate buffer was used (pH adjusted according to pI, but pH 4.0 unless stated otherwise). The flow rate was 3 ml/min and absorbance was recorded at 280 nm.

#### 2.3.7.3 Cation-exchange chromatography

After desalting, a more refined form of cation-exchange chromatography was carried out, on the BioCad 700E system, as a second purification step. After loading onto (normally) a Tricorn MonoS column (4.6 mm x 100 mm, column volume = 1.7 ml) (Amersham) and then washing with (typically) 20 mM sodium acetate buffers, pH 4-4.5 (depending on the calculated pI of the protein), samples were eluted by applying a linear gradient to 1 M NaCl in the same buffer and absorbance of eluate monitored at 280 nm. Protein-containing fractions were analysed by SDS-PAGE and then appropriate fractions pooled for the third and final purification step.

#### 2.3.7.4 Gel-filtration (size-exclusion) chromatography

Gel-filtration or size-exclusion chromatography, performed on an ÄKTAdesign™-FPLC system (having a pump P-920, UV-detector unit UPC-900, from GE-Healthcare), was used as a "polishing step" and to separate impurities from the protein of interest based on size. It was carried out on a HiLoad 16/60 Superdex 75 (preparation grade) column (GE Healthcare) using (normally) phosphate buffered saline as an elution buffer. Note that in the cases of proteins intended for NMR-based studies, 20 mM potassium phosphate buffer, pH 7.4, with 500 mM NaCl was used as an elution buffer to make the subsequent buffer-exchange into an NMR-compatible buffer (low salt) easier.

### 2.3.8 Enzymatic deglycosylation

$Endo_H{}^f$, $Endo_H$, or PNgases (from New England Biolabs) were used to remove or "trim" N-glycans (that are added at N-glycosylation cognate sites by *P. pastoris* but, being hypermannosylated, do not correspond to mammalian equivalents and are in any case often highly heterogeous). These enzymes cleave the bond between the innermost N-acetylglucosamine residue and the remainder of the glycan. Enzymatic deglycosylation was mostly done between the initial step-gradient ion-exchange protein-capture stage (after which diffuse bands on SDS-PAGE furnished evidence of glycosylation) and the subsequent higher-resolution ion-exchange chromatographic step; but in the case of small-scale trial protein productions it was performed on crude cell-free supernatant. Between 10-15 µl of the enzyme was used for deglycosylating a range of 0.2-0.3 mg protein eluted from SP-Sepharose resin. The enzyme were added to the protein sample, mixed and incubated for 2-3 hours at 37 °C. Subsequently the glycosidase-treated

samples were run alongside untreated ones on SDS-PAGE in the hope of detecting more defined protein bands of lower molecular weight corresponding to the expected (calculated) value.

### 2.3.9 Determination of Protein concentrations

Protein concentrations were estimated based on UV absorbance (280 nm), calculated molecular weights and theoretical extinction coefficients (obtained using the on-line ExPASy ProtParam tool). Absorbances were measured on a BioSpectrometer (Eppendorf).

### 2.3.10 Buffer exchange by spin concentration or dialysis

Buffer exchange was essential in the final preparation of the sample for both structural work and functional studies. Buffer exchange by spin concentration was mainly used when isotopically labelled (or non-labelled for a 1D experiment) proteins were being prepared for NMR. Dialysis, on the other hand, was typically used in preparing the samples for the biological assays (*e.g.* in SPR, all samples were dialysed into the working buffer at the same time).

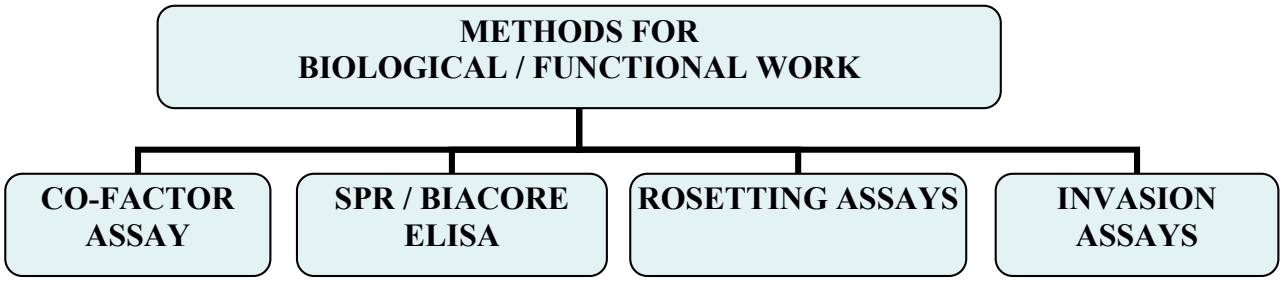
In order to achieve buffer exchange using centrifugal concentrators, the protein was loaded into one or more Vivaspin devices (Sartorius Mechatronics, note that it was important to thoroughly wash the filtration membrane to remove the preservatives), with suitable capacities and molecular-weight cut-offs, and concentrated as described above to a smaller defined volume but being mindful of the danger of protein aggregation or precipitation. Then a known quantity of cold replacement buffer was added (with thorough mixing using the pipette tip) and the protein sample was re-

concentrated. This procedure was repeated until it was calculated that the required extent of buffer exchange had been attained.

Buffer exchange by dialysis involves loading a suitable amount (*e.g.* 250 µl) of the protein solution into a thoroughly rinsed dialysis cassette (Bio-Rad), gently covering it with the lid (in which the membrane is located in this apparatus) and floating it over the replacement buffer. The buffer was constantly stirred using a magnetic stir bar and changed at intervals until the desired extent of buffer exchange had been attained.

## 2.4 Biological Studies

A range of methods was used to assess various aspects of the biological functions of the purified recombinant proteins. These include measurement of fluid-phase factor I co-factor activity against C3b and C4b, surface plasmon resonance (SPR)-based binding studies and measurements of affinity for host and parasite ligands, and erythrocyte rosette-disruption assays (in *P. falciparum*-infected blood samples). Carried out elsewhere by collaborators were *P. falciparum* merozoite invasion-competition assays and Enzyme-Linked Immunosorbent Assays (ELISAs) (see Flow chart 2.7**).**




***Flow chart 2.7 Overview of methods for Biological studies***

### 2.4.1 Co-factor Assay

The factor I co-factor assay was employed as a simple measure of native complement regulatory activity of the recombinant proteins, and was employed in comparisons of the four allotypic variants of CR1 15-25 corresponding to the Knops blood-group antigens McC(a/b) and Sl(1/2). This assay was based on the principle that native full-length CR1 is a cofactor for cofactor I-mediated cleavage of C3b and C4b, producing iC3b and iC4b, and indeed (uniquely to CR1) further cleavages to C3dg or C4dg (Figure 2.3A).

In the current study, four sets of experiments were set up. In two of these identical amounts of the putative cofactor were used (namely 1 μg) but the incubation times varied (15 minutes and one hour) while in the other experiments the amount of co-factor was varied (0.25 μg and 1 μg) and the incubation time maintained at one hour. Positive controls included soluble full-length CR1 (sCR1) and the three-module fragment CR1 15-17 corresponding to functional site 2; CR1 21-22 and buffer (PBS) alone were used as negative controls. All samples contained 0.1 μg of factor I, 2.5 μg of C3b and either 0.25 μg or 1 μg of the appropriate CR1 construct in a total volume of 20 μl. In a related study by our collaborators (in the Atkinson lab, Washington University Medical School), the cofactor activity of the CR1 15-25 polymorphic forms were assessed for both both C3b and C4b claevage.

The proteolytic reaction was initiated by addition of factor I (0.1 μg) and stopped by adding 6 μl NuPAGE reducing buffer containing lithium dodecyl sulfate (pH = 8.4, from Invitrogen). Figure 2.3B illustrates what would be expected from SDS-PAGE of the products of a cofactor assay. Note that in the presence of cofactor (*e.g.*

CR1 or factor H), factor I (initially) cleaves sequentially and very specifically the C3b α' -chain (*i.e.* the remains of the α-chain after cleavage of C3 to C3b) in two places to release the small C3f peptide and create iC3b. The β-chain is not cleaved by factor I. Thus upon reduction of the inter-chain disulphide linkage, two bands corresponding to the 43-kD and 67-kD products of the C3b α'–chain can be resolved by SDS-PAGE along with the intact β-chain. Further cleavage of iC3b by factor I occurs in the presence of CR1, yielding the large C3c fragment and the smaller but important C3dg fragment (cleaved subsequently by other proteases to ~30 kD C3d). Thus, in this assay, activity is inferred from the amount of cleavage products observed following SDS-PAGE analysis of the reaction mixture.

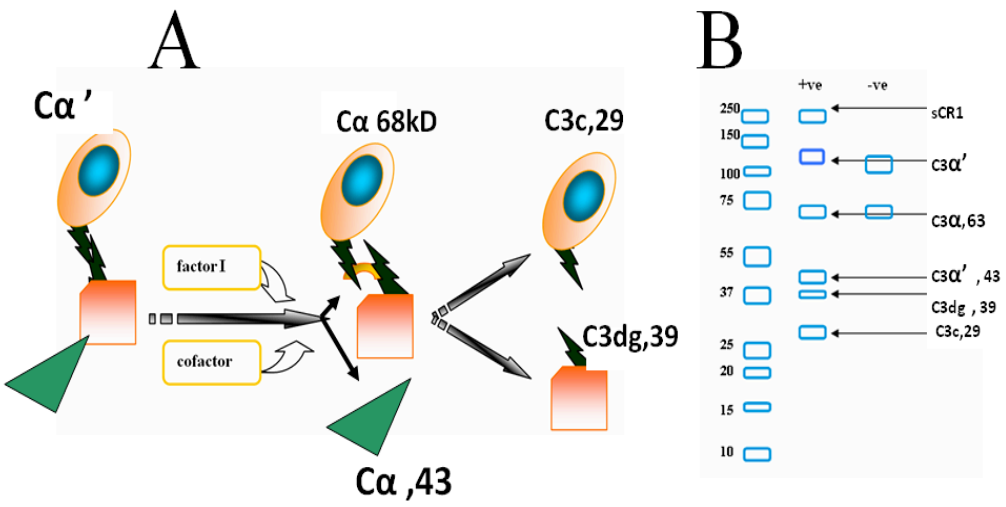



***Figure 2.3 Illustration of the principle and the expected outcome of the fluid-phase cofactor assay***

(**A**) CR1 is a cofactor for factor I-mediated cleavage of C3b α'-chain. Cleaved fragments are represented with cartoons and approrpriately labelled.(**B**) A schematic to show the expected outcome of SDS-PAGE performed to analyse the reaction products. Only CR1 is cofactor for third cleavage (to C3dg, not C3d (C3dg gets cleaved by other proteases to C3d in vivo).For details of the molecular weights and chains of origin of the cleaved fragments, see section 5.2 (chapter 5).

### 2.4.2 Surface plasmon resonance (SPR)

The phenemonon of surface plasmon resonance (SPR) was utilised to follow binding of the recombinant CR1 constructs, and approriate controls, to C3b, C4b, the classical pathway complement protein, C1q, and the parasite protein domains, DBLα and Rh4.9. A Biacore T100 (GE Healthcare) instrument was used for these studies. The theory of SPR is beyond the scope of this thesis but, in practice, a protein is chemically immobilised onto the polymer coated surface of a sensor chip. The chip "senses" the increase in surface density that arises from the binding of these protein molecules. The magnitude of this response is recoded accurately in real time, and expressed in resonance units (RU). For most proteins one RU corresponds to a change in the surface density of 1 pg/mm$^2$ on the sensor surface (Fig. 2.4B). Next a solution of a second protein is flowed over the protein-bearing sensor-chip surface. Binding of the protein in solution ("analyte") to the one on the sensor chip ("ligand") is likewise sensed by the chip surface leading to to a further response. The strength of the recorded response (in RU) is thus a function both of the number of analyte protein molecules that bind to the ligand protein but also of their molecular weights.

In the current work, C3b, C4b, C1q, or the parasite protein domains, were immobilized as ligands while the various CR1 fragments produced in this study served as analytes. Amine coupling was used for covalently attaching the ligands onto the surface of Biacore series-S carboxymethylated dextran (CM5) sensor chips (Biocore/GE Healthcare) in the current study. Note that each chip had four distinct "flow cells" (numbered 1 to 4) such that a solution of analyte could be flowed simultaneously over four different ligand-bearing surfaces and data recorded from all

four channels. Immobilization was carried out according to manufacturer's instructions (Fig. 2.4C).

Following a number of trials (described in the Results chapter), two separate CM5 sensor chips were used for the definitive set of experiments that aimed to achieve reproducible and publishable data. The first was termed the "complement chip". It was loaded with three different human proteins - C3b, C4b, and C1q (from Complement Technology). These were individually immobilized, by amine coupling (Schmidt *et al*., 2008), on three of the four flow-cell surfaces, with the fourth flow cell being employed as a blank or reference surface. The second CM5 sensor chip was termed the "malaria chip**"** and was loaded with two recombinant *P. falciparum* protein fragments (desribed in more detail in Section 3.11 and in the Results), namely NTS-DBLa-CIDR (donated by Matt Higgins, Oxford) and Rh4.9 (from Alan Cowman, Melbourne). Details of loadings (in terms of numbers of RUs) are given in the Results chapter.

To prevent interference from any non-specific binding of the proteins to the reference surface of the chip, a phenemonon that usually translates into negative curves, a "blank immobilisation" was performed as described below. This was done twice. To ensure reproducibility, duplicate injections of all the samples were made. The flow rate ws set to of 30 µl/min, the temperature maintained at 25 °C, contact times of 90 and 45 seconds (for the 'complement' and the 'malaria chips' respectively) were used and the dissociation time allowed was 200 s. The running buffer was HBS-EP+ consisting of 10 mM HEPES-buffered 150 mM saline, 3 mM EDTA and 0.05% (v/v) surfactant polysorbate 20 at pH 7.4 (HBS-EP+)**.** As the experiment proceeded, two injections of 1 M NaCl with a contact time of 45 s were used between sample-injections to regenerate the surfaces of the chips. These conditions emerged following trial runs (performed

with the kind assistance of Dr Christoph Schmidt**)** that aimed to achieve optimised conditions.

Data were processed using the Biacore T100 evaluation software version 2.0, and the points of reporting (for $K_D$ calculations) were set to be 2 seconds before the beginning and after the end of the injections. The software derived the dissociation constants in this case by fitting steady-state binding levels. Note that the response recorded on the blank surface was subtracted from the response unit recorded on the surface bearing the ligand.

#### 2.4.2.1 Sensor chips and amine coupling

The sensor chip comprises a glass surface that has been coated with a thin layer of gold. In this work the CM5 chip was utilized, in which the gold surface is modified with a carboxymethylated dextran layer. This dextran hydrogel layer forms a hydrophilic environment for attached biomolecules, preserving them in a non-denatured state. Other derivatized surfaces were available but not used in this study (Fig. 2.4D).

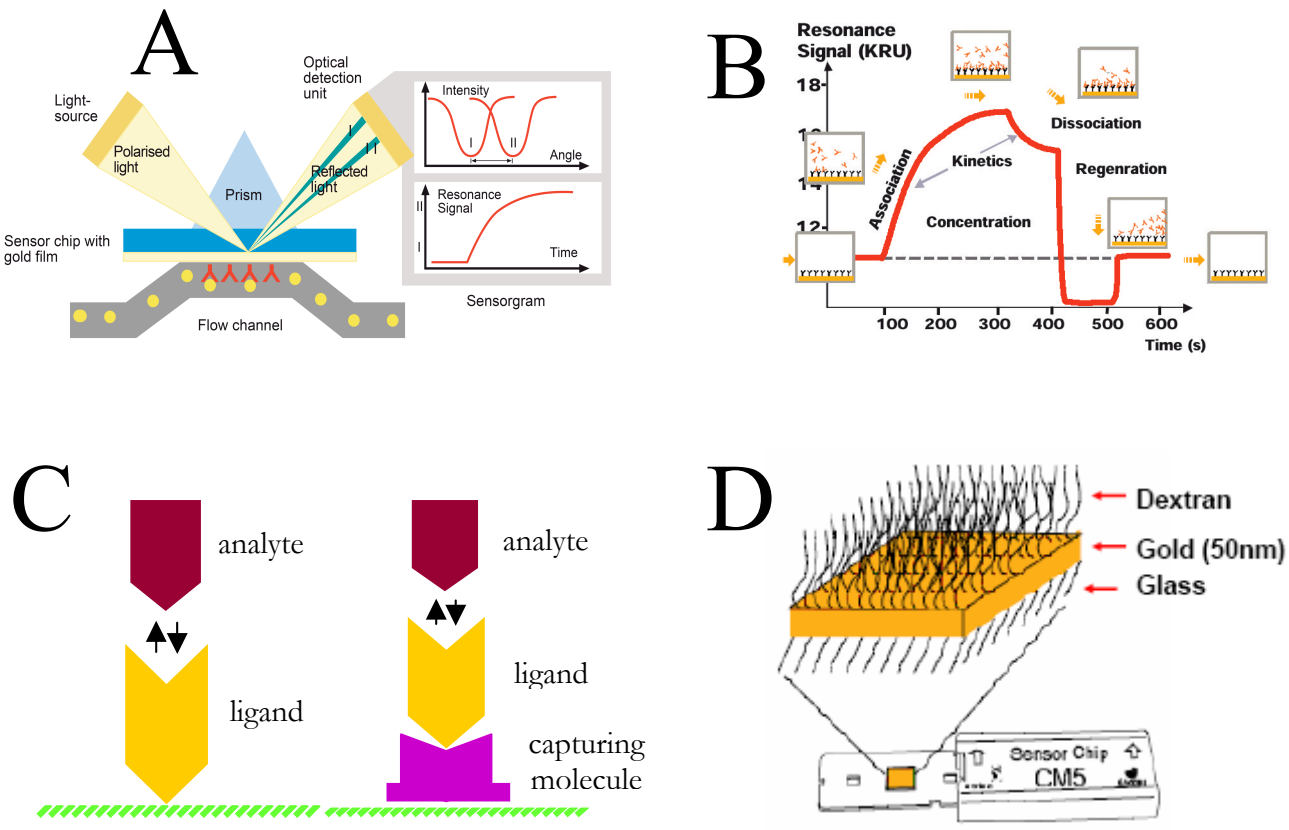



***Figure 2.4 Surface plasmon resonance***

(A) The principle of SPR. (B) Sensorgram generation. (C) The terminology of SPR, illustrating analyte and ligand, as well as the types of immobilization that may be employed. (D) A schematic to represent the composition of the chip surface (Diagrams adapted from www.biacore.com)

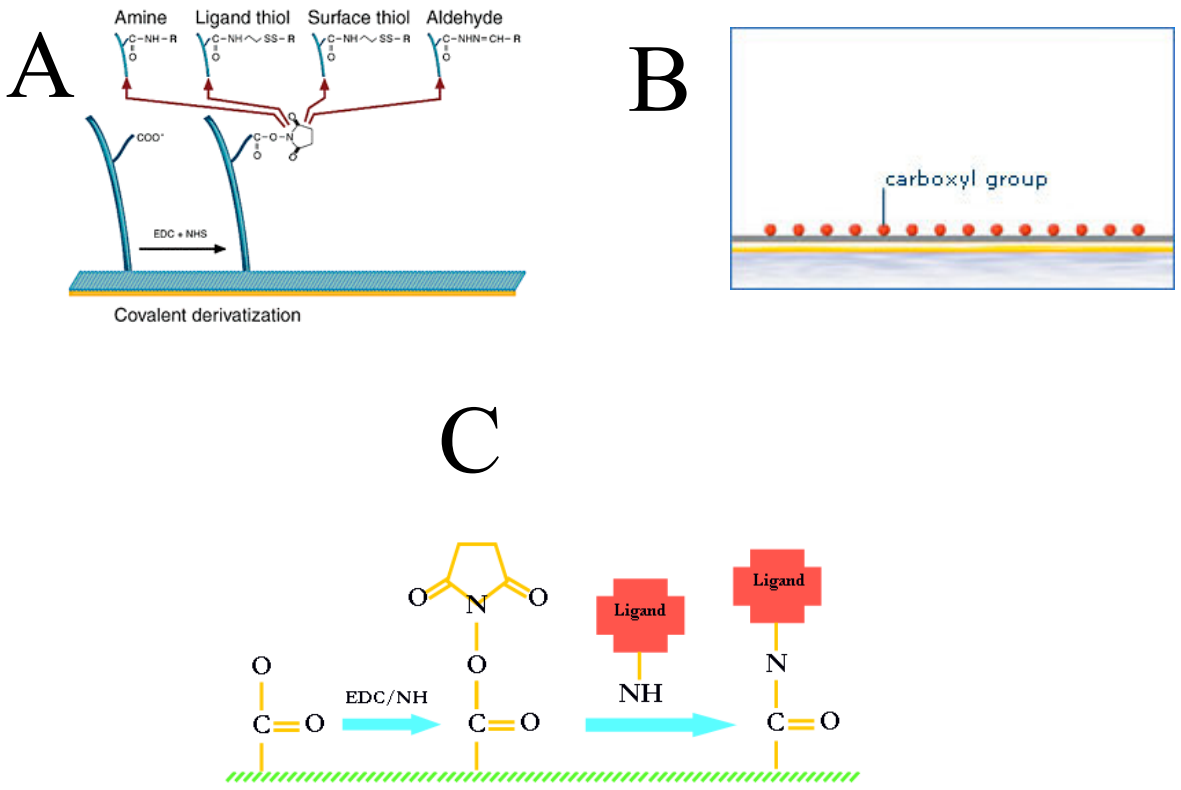



***Figure 2.5 Sensor chips and coupling***

(A) A schematic representation of the CM5 sensor chip used in the current study. (B) For comparison, a C1 sensor chip. (C) Amine coupling was used in the current study to immobilise ligands (e.g. C3b, C4b Rh4.9 *etc.* to the CM5 chip sensor surface). (Diagrams adapted from www.biacore.com).

### 2.4.3 ELISA

The principle of this tool is based on the detection of antigen-antibody complexes. Enzyme-Linked Immunosorbent Assay (ELISA) used in this study were mainly carried out in John Atkinson's laboratory in St Loius, MO, USA. The following describes the binding experiments of sCR1 and CCPs 15-25 variants to C3b, C4b, C1q and MBL.

To a C3b-coated plate, solutions of 1 mg/ml CR1 variants (25 mM NaCl) were added. The primary antibody used for detection was CR1 polyclonal antibody, and the secondary antibody was anti-rabbit IgG conjugated with horseradish peroxidase (HRP).

The results of three separate experiments were averaged. Similar experiments were performed on plates coated with C4b.

For detection of C1q binding to CR1, however, sCR1 or CR1 15-25 variants were coated on plates. The concentration of C1q added was 140 ng/ml (in 75 mM salt), the primary antibody used was rabbit anti-C1q, and the secondary antibody was anti-rabbit IgG conjugated with HRP. The results of three separate experiments were averaged

Finally, results for MBL were obtained when 5 μg/ml MBL (25 mM salt) was added to the CR1-coated plate and binding was detected with anti-MBL antibody made in mouse, and anti-mouse IgG conjugated with HRP. The results of two experiments were averaged.

### 2.4.4 Rosette inhibition /Disruptive assay

Figure 2.6 illustrates in simple terms the steps taken on the day of performing a rosette-disruption assay. The materials needed were a rosetting *P. falciparum* culture (at about 2% haematocrit in malaria culture medium), ethidium bromide, "binding medium" (RPMI 1640 with additives and 10% serum, but no bicarbonate, approx pH = 7.3 (Sigma-Aldrich) microscope slides, cover slips and Vaseline, a fluorescence microscope (Olympus) and a water bath set at 37 $^{o}$C.

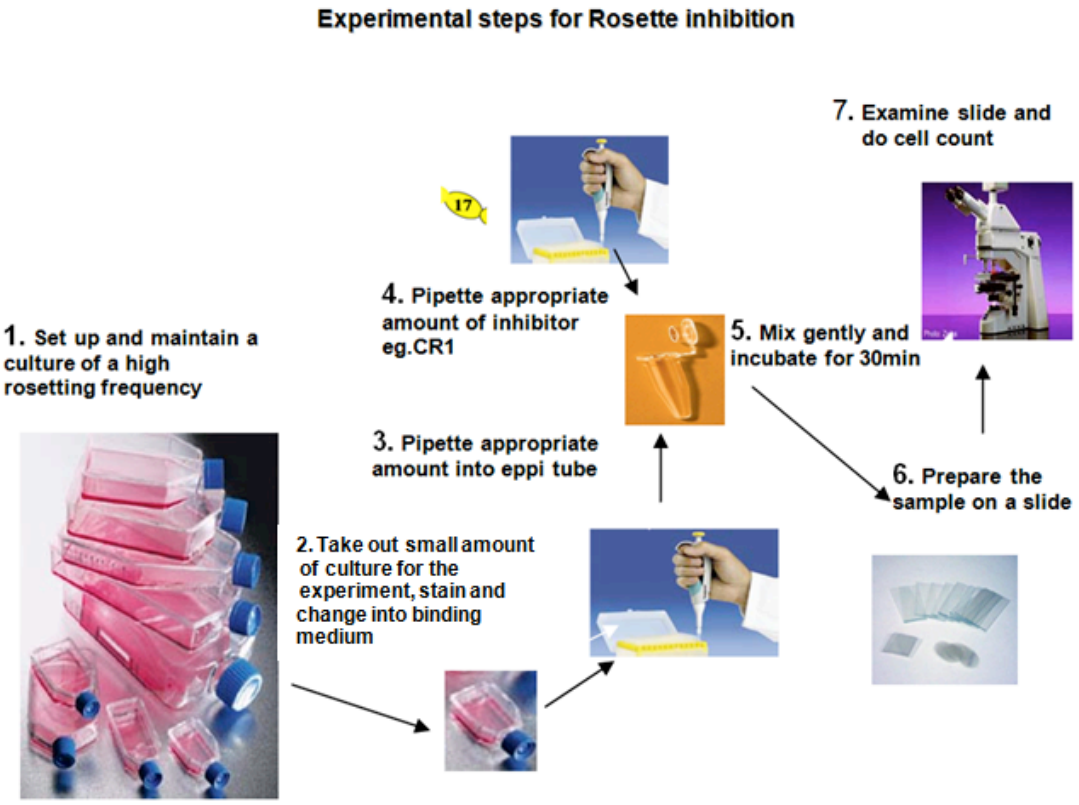



***Figure 2.6 Experimental steps involved in rosette-disruption assay***

1- Cell culture flask containing rosetting parasite. 2- Separate culture to be stained and changed into binding medium. 3 - Aliqout of culture added into microfuge tube 4-Addition of potential rosette disrupter. 5- Mixing and incubating . 6- Sample prepared for microscopic examination.

The first step involves setting up a culture and maintaining it to a high rosetting frequency. This has been detailed later in a separate section, since it is a complex procedure. On the day of the experiment, an aliquot (between 10 – 15 ml) of parasite culture suspension, made up of mainly the trophozoic form (in step 2, see Fig. 2.6) was pre-stained with 25 μg/ml ethidium bromide for about 5 minutes at 37 $^{o}$C. The cell suspension was centrifuged (2000 rpm for 2 minutes) and the ethidium-bromide containing supernatant was discarded (via a proper disposal procedure due to its carcinogenic potential). The cell pellet was re-suspended at about 2% haematocrit in a

"binding medium" containing 10% (v/v) serum. It differed from the regular culture medium (see section 2.4.4. for details).

In steps 3 and 4 (Fig. 2.6) aliquots (usually about 25 or 50 μl) of the pre-stained culture suspension were placed into a microfuge tube. Potential inhibitors or disruptors of rosetting were added at various concentrations. It was important to include positive and negative controls here (see below for details). The mixture was incubated (step 5) at 37 $^{o}$C for 30 mins. During this incubation period, cells were re-suspended every 10 minutess by flicking the tube gently.

Subsequently (step 6), a drop (about 10 μl) of the pre-stained culture suspension was placed on a microscope slide. The edges of the cover slip (22 mm x 22 mm) were lightly coated with Vaseline before it was lowered gently over the drop. At this stage, control and experimental slides were labelled in code and then shuffled by a colleague so that ensuing steps were performed "blind" by the experimenter. Only four slides were prepared for counting at any one time to avoid delays between slide preparation and the counting step.

Each slide was viewed (as quickly as possible) under the fluorescence microscope using a 40x magnification objective lens – illumination was adjusted so that both infected erythrocytes (that fluoresce orange) as well as the non-infected cells (not fluorescent as they have no nucleic acid) can be seen. A total of 200 infected red blood cells were counted and scored as either being in a rosette, or not, where a rosette is defined as "an infected cell with two or more uninfected red cells sticking to it". (Note that when the slide or microscope was knocked accidentally, counting was aborted since this seemed to lead to erroneous readings due to clumping of cells because of the culture touching the cover slip). Only mature (that is, pigmented trophozoite or

schizont) infected cells were counted; these were easily differentiated from ring-stages because rings only give a "pin-prick" of fluorescence. This allowed calculation of "rosette frequency" as the number of infected cells in rosettes expressed as a percentage of the total number of infected cells counted. (Note that slides, tubes and tips were disposed off in a sharps bin for incineration).

To ensure a successful experimental day, much preparation was necessary in the preceding weeks. This is described in the following sub sections. These included media preparation, washing of cells (erythrocytes), thawing of the parasites, setting up the parasite culture, synchronizing it at intervals, and selecting and maintaining the culture at high rosette frequency.

#### 2.4.4.1 Preparation of culture media for the rosette-disruption assay

RPMI-1640 (developed at Roswell Park Memorial Institute) is a bicarbonate-buffered medium originally developed for culturing leucocytes. For the current studies, "incomplete RPMI-1640" was prepared as follows: to 500 ml RMPI-1640 (Sigma-Aldrich) (note this contains 26.7 ml/L of 7.5% $NaHCO_3$ solution or 2.0 g/L of $NaHCO_3$ powder bicarbonate) was added 12.5 ml of 1 M Hepes, 5 ml of 200 mM glutamine, 5 ml of 20% (w/v) glucose solution, 1.25 ml of 10 mg/ml gentamicin, and ~1 ml of 1 M NaOH (used to adjust the pH to approx 7.2 to 7.4 as judged by observing the colour. It was noted that a red colour indicated the right pH, not orange or pink. Complete RPMI-1640 was prepared using the same protocol but by adding, to a final concentration of 10% (v/v), pooled (from at least five donors) normal human serum (*i.e.* not heat inactivated and stored at -20 °C prior to use). Therefore an aliquot of 40 ml of serum

was added to 360 ml of incomplete RPMI. The "binding medium" was the same as complete RPMI-1640, but contains no bicarbonate – this maintained a stable pH in the absence of high $CO_2$ more effectively than medium with bicarbonate.

#### 2.4.4.2 Washing of blood

Group O red blood cells (Scotish National Blood Transfusion Service) were routinely used to avoid problems with ABO incompatibility. Some blood transfusion service blood packs came with a leucodepletion filter that was used to remove white cells before aliquoting. When this had been done, samples (labelled as "white cell depleted") were just washed two times, by centrifugation and re-suspension, in incomplete RPMI-1640 and stored at 50% haematocrit in incomplete RPMI. For whole blood, the procedure started with diluting the fresh blood with an equal volume of incomplete RPMI, and then lymphocytes and monocytes were removed by layering 10 ml of diluted blood over 5 ml of Lymphoprep in a 15 ml tube. It was spun for 15 minutes ( at top speed ) on a bench-top centrifuge (at room temperature), and the supernatant were aspirated. The red blood cells were then washed twice, by centrifugation and re-suspension, with 13 ml of incomplete RPMI. These steps were repeated and finally, the cells were re-suspended at 50% haematocrit in incomplete RPMI, to be stored at **4** °C. Ideally, the washed red blood cells should be used within a week and must be used within two weeks of being drawn.

#### 2.4.4.3 Thawing parasites and setting up a culture

To ensure thawing of the R29 strain of *P. falciparum* parasite ( a laboratory adapted strain cloned from IT/FCR3 strain), a frozen vial of this parasite was taken from liquid

nitrogen, placed in 37°C water bath and monitored by visualizing it. Once thawed, it was transferred into a 50-ml Falcon tube. Then, 200 µl of sterile 12 % (w/v) NaCl (at 37°C) was added dropwise with a sterile plastic Pasteur pipette (with continuous agitation over two minutes) to the 1 ml thawed suspension. The tube was allowed to stand for 5 minutes and then 10 ml of 1.8% (w/v) NaCl was added, again, in drops. Following this, 10 ml of 0.9% (w/v) NaCl and 0.2% (w/v) glucose were added slowly and the tube was centrifuged for 4 minutes at 2000 rpm. The supernatant was removed and cells were washed twice in a previously warmed (37°C) 20 ml of incomplete RPMI-1640, (re-suspending and spinning at each washing step) and finally re-suspended in complete RPMI-1640. Before incubation (at 37°C), the final suspension in the falcon tube was transfered into a T150 culture flask (Sigma-Aldrich®.), leaving a small aliquot for Giemsa staining. Gas was bubbled through the contents of the culture flask (a mixture of 94% nitrogen, 5% CO2, and 1% oxygen) prior to incubating (3 days) at 37 °C.

#### 2.4.4.4 Growing parasites

The growth medium (complete RPMI-1640) was warmed to 37°C (water bath) before use since the parasites would not tolerate cold shock. To the parasite cultures, prepared as described above, fresh medium was supplied daily. The medium was changed by transferring the culture suspension to a centrifuge tube, spinning out the cells (as above) and re-suspending in fresh medium (amount decided as described below). It was then transferred to culture flask, and "gassed" (see above) for approximately 30 seconds, before returning to the incubator at 37 °C.

### 2.4.4.5 Giemsa smears

In order to monitor the culture and maintain an optimal feeding regime, smears were prepared daily by transferring a small aliquot of the suspension into a microfuge tube which was then spun down, and cells re-suspended at about 50% haematocrit after which they were smeared onto a clean glass slide. The smear was dried and fixed with methanol from sealed glass bottle. This is important because if the methanol was stored in an unsealed plastic bottle, it could have absorbed water from the atmosphere, especially when it is warm. It then no longer fixes the cells properly, so the cells would look fuzzy and have holes in them. The dry – fixed smear is then stained with freshly made 10% Giemsa, and parasitaemia was estimated by counting 500 cells. This provided a rough indication of how much fresh medium and how many red blood cells were required for maintenance of the culture since most parasite lines exhibited approximately a five-fold extent of reinvasion of cells for each. The amount of red blood cells supplied per cycle depended on how many parasites were needed, and what level of parasitaemia was desired. The amount of medium that was required depended on the level of parasitaemia and the stage of the lifecycle. Parasites at ring-stage needed less medium than mature pigmented trophozoites/schizonts, while the higher the level of parasitaemia, the more medium was required. The equation for estimating the amount of medium required was; Medium (ml) = 5 x parasitaemia (%) x packed cell volume (ml). Care was taken not to overfill the culture flasks e.g. a packed cell volume of 2 ml (10% parsitaemia) was maximal given the maximum volume of medium of 100 ml in large flasks (T150).

### 2.4.4.6 Selection for rosetting

Parasites in culture had a tendency to lose the ability to rosette over time, through *var* gene switching. Two methods were therefore used to select for rosetting parasites: "Percoll" and "Plasmagel"**(** Corrigan and Rowe, 2010 **;** Ghumra *et al.,* 2011**).**

In the Percoll method (details below), which was done approximately once a week involved centrifuging parasite cultures through Percoll to select rosettes. A 90 % (v/v) stock solution was made up from 90 ml Percoll (GE Healthcare) and 10 ml of 10x concentrated incomplete RPMI-1640 (pH close to 7). Further solutions were made by dilution *e.g.* a 60% (w/v) Percoll solution, was made by mixing 33.3 ml of the 90% Percoll stock and 16.7 ml of incomplete RPMI-1640. Then 5 ml of 60% (w/v) Percoll was placed in a 15-ml centrifuge tube while the cultured parasite cells were spun down and re-suspended in about 5 ml of complete-RPMI. This suspension was gently layered over the Percoll solution and this was spun at top speed for 10 minutes at room temperature. This gave rise to a packed pellet of non-infected red blood cells, ring-stage parasites and rosetting mature-infected erythrocytes at the bottom of the tube, and a (resolved) band of non-rosetting mature trophozoites and schizonts at the percoll/RPMI-1640 interface. Thus most of supernatant could be removed by gentle suction and the parasite layer containing non-rosetting trophozoites and schizonts was carefully removed leaving behind the pellet (Fig. 2.7A). This pellet was washed twice in incomplete RPMI-1640 and then returned to culture.

Selecting of rosettes by Plasmagel (which is 3% gelatine in normal saline, supplied as a sterile solution by Rhone-Poulenc) flotation is based on the principle that when suspended in Plasmagel, non-rosetting mature pigmented trophozoite-infected cells float as a layer on the top, whereas the denser uninfected RBC and infected cells

with ring stages or rosetting pigmented trophozoites sink to the bottom layer (see Fig. 2.7A). Therefore to enrich for rosetting parasites, the bottom layer was kept and the top layer was discarded. In this procedure, the gelatine was pre-warmed to 37$^{o}$C while, separately, the cells from the culture suspension were spun down and supernatant was removed. The pellet volume was estimated and incomplete RPMI was supplied to afford a haematocrit of about 40-50%. This was transferred to a 15-ml Falcon tube and an equal volume of gelatine was then added. The tube was place upright in a 37$^{o}$C incubator for 10-15 minutes, until two separate layers became clearly visible. The top layer containing the non-rosetting trophozoites was removed while the bottom layer, containing RBCs, ring-stages and rosetting mature-pigmented trophozoites was washed once in incomplete RPMI and again in complete RPMI. It was re-suspended in complete medium, and then fresh blood was added to give a culture with overall ~2% parasitaemia that was gassed (see above) and returned to the 37 $^{o}$C incubator to continue the culture.

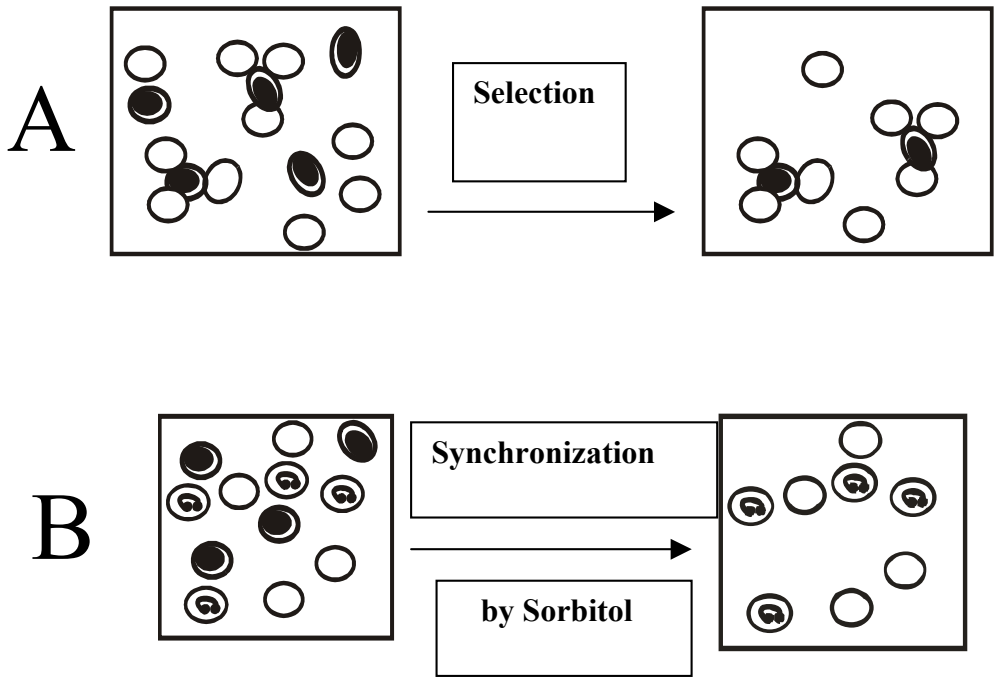



***Figure 2.7 Culture selection and synchronization***

(A) Selecting of rosetting using 60% Percoll or Plasmagel flotation. (B) Synchronization by sorbitol. Deep black represents- matured trophozoites and the squiggles represent – ring forms (in both panels A and B).

### 2.4.4.7 Synchronising parasites by Sorbitol lyses

The goal of this step was to achieve synchronicity of the paraite invasion/lysis cycle. The principle used here was that only parasitized red blood cells more than approximately 20 hours post-invasion are permeable to sorbitol. They can therefore be lysed by osmosis, leaving the younger ring forms and non-infected erythrocytes (Lambros *et al.,* 1979) (Fig. 2.7B). Cultured cells were spun down and supernatant removed. Approximately 5-10 ml of a pre-warmed 5% (w/v) sorbitol solution (at 37 $^{o}$C, from Sigma) was used to re-suspend the pellet. The suspension was incubated for 15 minutes (37 $^{o}$C) prior to two washes (by spinning down and re-suspending cells) using 13 ml incomplete RPMI-1640, to remove lysed cells. The non-lysed cells were then re-

suspended in complete medium, and a Giemsa-stained smear was examined to check whether mature pigmented trophozoites and schizonts had been removed. Finally the culture was "gassed" and returned to the 37 $^{o}$C incubator.

### 2.4.5 Invasion assays

The intention of these was to ascertain whether the CR1 fragments were able to compete with CR1 on red blood cells as a receptor for the silaic acid independent invasion pathway. These experiments were performed by our collaborator Dr W.H Tham, according to methods described in our publication (Tham *et al.,* 2010). In essence, parasite growth was monitored over two cycles of invasion and lysis (Persson *et al.,* 2008). Two strains of *P. falciparum* were used: W2mef-Rh4 and 3D7. W2mef-Rh4 is known to mainly use the sialic acid-dependent invasion pathway and therefore these strains served as a control. The 3D7 strain is able to switch its preferred invasion route when incubated with red blood cells that have been treated with neuraminidase to remove sialic acids from the cell surface. Thus in these experiments neuraminidase treated and untreated cells were used.

Neuraminidase (66.7 mU/ml)-treated or normal erythrocytes at 1% hematocrit in culture medium were inoculated with late-trophozoite stage parasites to give a parasitemia level of 0.2% and hematocrit of 1% in a volume of 50 µl. The parasites were subsequently cultured in 96-well round-bottom microtiter plates (Becton Dickinson). Solutions of the purified recombinant proteins to be tested were added, to a range of final concentrations, prior to the first reinvasion episode. After incubation with these proteins for two cycles of parasite growth, the parasitemia of each well was determined by flow cytometry of ethidium bromide -stained trophozoite stage parasites

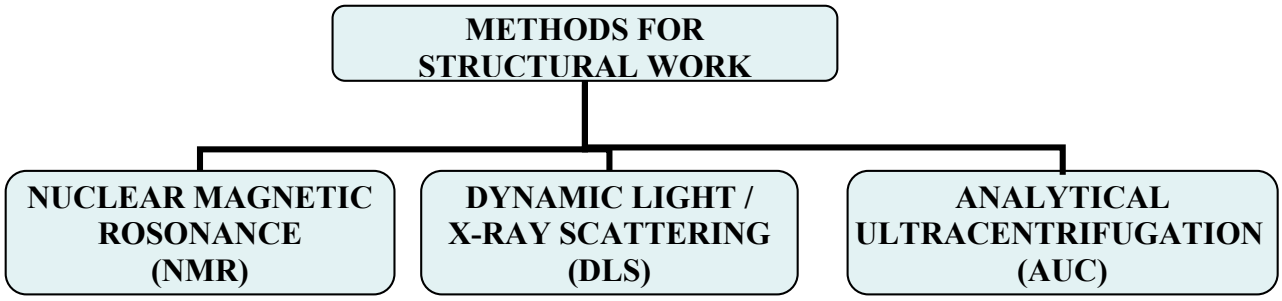
using Fluorescence-activated cell sorting (FACS) system with a plate reader (Becton Dickinson). For each well, 40,000 cells or more were counted. Growth was expressed as a percentage of parasitemia for the mean of two samples. Two independent assays were also performed.

Erythrocyte competitive binding assays were also carried out. These were performed by incubating the test proteins with matured culture supernatants prior to proceeding with standard erythrocyte binding assays as follows. An aliquot of culture supernatant (250 µl) was mixed with 50 µl of packed erythrocyte for more than 30 mins at room temperature. The mixtures were centrifuged at 12,000 rpm for 30 s through 400 µl of silicone oil (dibutyl phthalate; Sigma). This was done to remove unbound protein that has remained within the culture supernatant. Both the erythrocyte and bound protein were washed twice with 500 µl of phosphate buffered saline. Proteins bound to the erythrocyte were eluted by incubation with 10 µl of 1.5 M NaCl for 15 min at room temperature and then centrifuged for 30s at 12,000 rpm , and the elute was removed from the erythrocytes. An equal volume of 2x reducing sample buffer was then added to the eluted proteins. The eluted proteins were separated on SDS-PAGE and identified by immunoblotting.

## 2.5 Biophysical / Structural Studies

An overview of the methods used for structural studies is presented in Flow chart 2.9. As described in the Introduction, the aims of this component of the work were two-fold. It was important to assess any effect of the Knops blood-group polymorphisms in CCPs

24-25 on the architecture and self-association properties of CR1. It was also desirable to test the hypothesis that these variable residues are apposed to functionally important residues in functional site 2 *via* a hinge-like structure between LHR-C and LHR-D (*i.e.* modules 20-23).




***Flow chart 2.8 Overview of methods for structural studies***

### 2.5.1 Nuclear magnetic resonance spectrocopy (NMR)

**Theory and scope of experiments performed**

The phenomena recorded in nuclear magnetic resonance spectroscopy (NMR) derive from the responses of spins of certain nuclei (termed "spin=half" nuclei and including $^{1}H$, $^{15}N$ and $^{13}C$) to being placed in a strong applied magnetic field. Normally these nuclear spins are oriented randomly. But when a strong magnetic field is applied to a sample (*e.g.* of a protein) some of its nuclear spins become aligned either with the field or against it and these two states differ slightly in energy (at a level similar to the low energy of radio waves).

Thus when nuclei in a magnetic field are irradiated with radio waves of the appropriate frequency, interconversions between lower and higher energy states occur – a condition known as resonance. Different types of nuclei (*e.g.* $^{1}H$ *versus* $^{13}C$ or $^{15}N$) resonate at very different frequencies but the resonant frequencies of nuclei are also exquisitely sensitive to their immediate surroundings within the molecule. Hence a spectrum of $^{1}H$ NMR frequencies is observed for a protein, for example, that can be used to infer how compactly folded it is and ultimately (backed up by $^{13}C$ and $^{15}N$ frequencies and NMR experiments designed to explore and record structural and spatial relationships between these nuclei) to determine its structure (not performed in the current work). Moreover, interactions with other molecules (or other domains in the same protein) cause changes in resonant frequencies to varying extents depending on the size of the interacting interfaces and these can be used to delineate interacting sites in favourable circumstance.

The NMR data acquisition for this work was kindly carried out by the NMR manager of the Edinburgh Biomolecular NMR unit, Mr Juraj Bella. Dr Christoph Schmidt and Dr Mara Guariento also helped with data collection and/or processing of spectra with additional expert advice from Dr Dusan Uhrín. Data were acquired by the use of Topspin (the name of the relevant Bruker software) while AZARA (freely available from the Common Computing Protocol for NMR (CCPN), University of Cambridge) was used for spectral processing **(**Vranken *et al*., 2005**).** Most of the NMR data in this work were collected on the Bio800 NMR spectrometer but the Bio600 spectrometer was also used from time to time, *e.g.* for collecting $^{1}D$ spectra or to confirm that protein samples were still intact.

A range of one-dimensional ($^{1}$H-observe) and multidimensional (for details of sample conditions and pulse sequences, see below and in relevant sections of the Results chapter) spectra were recorded and processed, using standard techniques that have been described elsewhere (*e.g.* on the Edinburgh BioNMR Unit's website but also in textbooks and the literature), with greatest use being made of the $^{15}$N,$^{1}$H-HSQC experiment. Overlays of $^{15}$N,$^{1}$H-HSQC spectra of the various fragments were performed in order to compare chemical shifts of specific CCP modules in different contexts (*e.g.* in CR1 21-22 versus CR1 20-23) and thereby infer information regarding the extent of their interactions with other CCP modules.

The standard suite of three-dimensional NMR experiments (*e.g.* the CBCA(CO)NH and CBCANH "pair" as found in Bodenhausen and Ruben, 1980, Vuister and Bax, 1992, Grzesiek and Bax, 1992, 1993), using very well established pulse sequences (installed on our Bruker instruments) and required for backbone assignments, were performed on a $^{15}$N,$^{13}$C double-labelled sample of the CCP 21-22 pair that spans the abnormally long LHR-C to LHR-D linker. Finally a $^{15}$N,$^{1}$H-HSQC spectrum (and a TROSY spectrum that is particularly suitable for larger proteins) was recorded on a deuturated ($^{2}$H,$^{15}$N-labelled) sample of 20-23.

### 2.5.2 Backbone assignment of CR1 21-22

In backbone assignment (performed using ANSIG, from the CCPN, the $^{15}$N,$^{1}$H-HSQC experiment played a key role in the sense that it served as a central reference spectrum. Since each cross-peak in such a spectrum represents a distinct $^{15}$N with its attached proton, its coordinates on the y-axis (nitrogen) and x-axis (hydrogen) corresponds to the resonant frequencies (chemical shifts) of these nuclei.

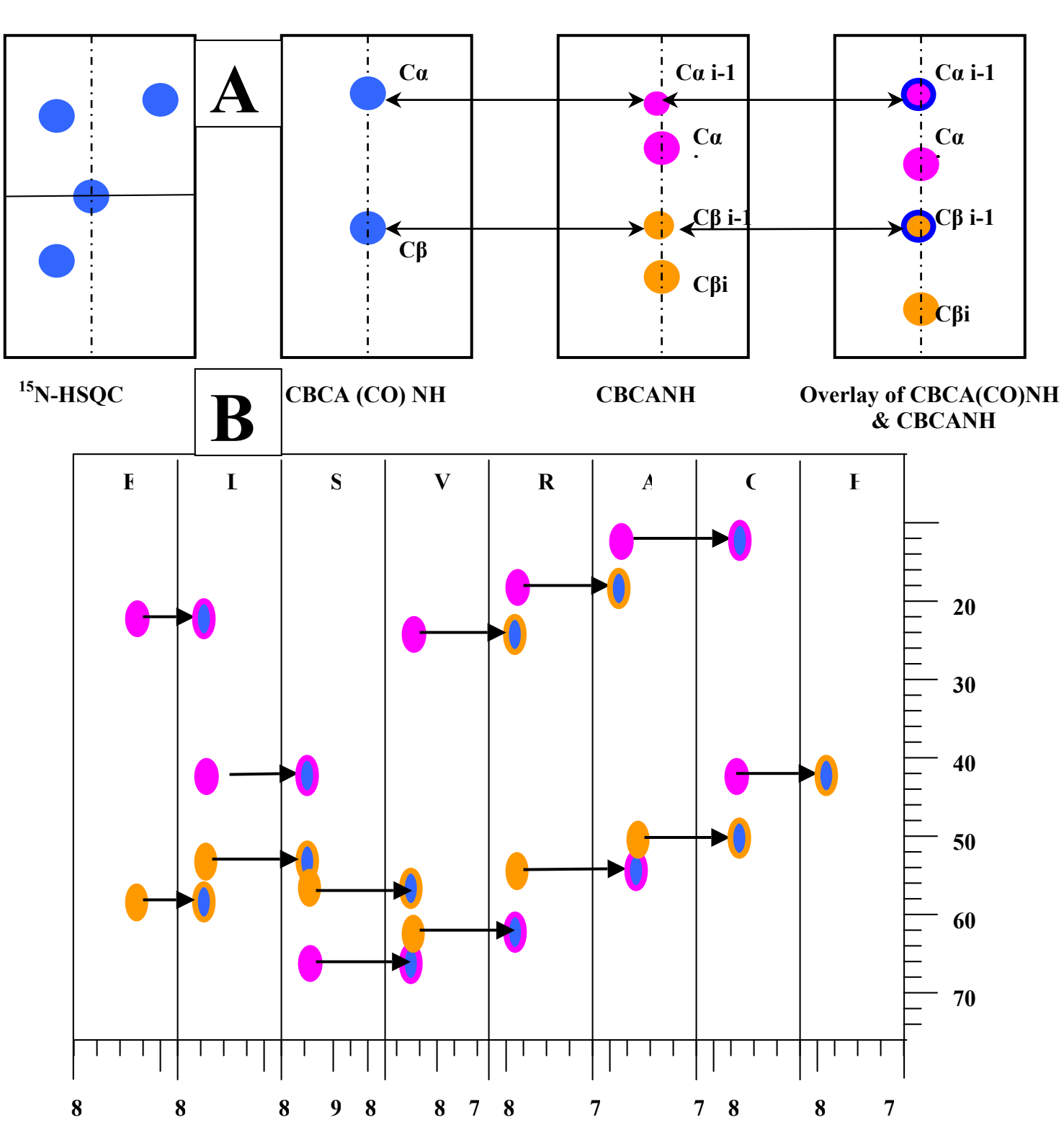



***Figure 2.8 Steps of Backbone assignment***

(A) Schematics to illustrate the $^{15}N,^{1}H$-HSQC, CBCA(CO)NH and CBCANH spectra needed for backbone assignment. (B) Illustration of a hypothetic sequential walk (through the long linker of CR1 21-22 – actual data appear in the Results chapter).

# CHAPTER 2 MATERIALS AND METHODOLOGY

In addition to the $^{15}N,^{1}H$-HSQC, the three dimensional CBCA(CO)NH and HNCACB triple resonance experiments were employed to sequentially assign the chemical shift of the Cα, Cβ, N and (N)H nuclei of the amino acids (Figure 2.8A). Before commencing resonance assignment, all the peaks in the $^{15}N,^{1}H$-HSQC, CBCA(CO)NH and HNCACB spectra were "picked" (in the ANSIG program installed on a desktop computer) by putting a cross in the centre of each peak that was considered to be valid (*i.e.* not to be noise nor to be an artefact nor a "diagonal" peak) . Since the three experiment were linked *via* having effectively common $^{1}H$ and $^{15}N$ axes, each cross-peak in the $^{15}N,^{1}H$-HSQC spectrum was easily traced to corresponding strips of cross-peaks (*i.e.* with connectivities to the "root" HSQC peak) in the CBCA(CO)NH and HNCACB spectra. These strips of peaks correspond to $^{13}C$ nuclei that are connected through bonds to the backbone amide $^{15}N$ of the root resonance. The pulse sequences used (well established and not discussed further here) ensure that in the HNCACB experiment the Cα and Cβ cross peaks of residue *i*, as well as those of residue *i-1* appear as cross-peaks. On the other hand in the CBCA(CO)NH experiment only the Cα and Cβ cross-peaks of residue *i-1* are detected. Thus an overlay of strips from the respective spectra allows identification of sequential pairs of strips (see Fig. 2.8). In Figure 2.8, colour coding (cross-peaks from CBCA(CO)NH are blue and represents the Cα and Cβ of residue *i-1*, while the HNCACB cross peaks are yellow for the Cβ (of residue *i*) and pink for the Cα (of residue *i*)) indicates this process. As seen in the last panel of Fig. 2.8, it became clear which of the Cβ and Cα cross peaks belonged to the $i^{th}$ residue or the $i\text{-}1^{th}$ residue, since the *i-1* cross-peaks are overlays of blue and yellow, or pink, in this colour scheme.

The figure also illustrates how, by extension, a sequential series of strips were identified. Thus to link an arbitrarily assigned residue (*i*) with the next amino acid residue in the protein sequence, horizontal lines were ruled from residue *i*'s directly attached Cα and Cβ cross peaks in that strip. Then by scrolling through the peak list database (generated in ANSIG when the peaks were picked), it was often possible to find a strip - corresponding to residue *i+1* - that exhibited the same pair of chemical shifts but in this case they corresponded not to its own Cα and Cβ nuclei but to the equivalent nuclei of the preceding residue, *i* (blue peaks for illustration purposes in Fig. 2.8). Then, in this new-found strip *(i+1)* the cross-peaks of directly attached Cα and Cβ were used to find the strip corresponding to *i+2*. This process could continue, both forwards and backwards along the protein sequence ("sequential walk") (see Fig. 4.8B) until a break is reached due to lack of data or the presence of a proline that lacks an amide group.

Most side chains have characteristic Cα and Cβ shifts that allow them to be identified or classified according to amino acid type. So given the identification of chains of sequential strips that could individually be tentatively assigned in many cases to amino acid types it was not surprising that comparison with the actual sequence allows unambiguous match between a series of strips and a run of amino acids - thus establishing sequence-specific assignments of $^{13}$Cα and $^{13}$Cβ as well as H($^{15}$N) and $^{1}$HN nuclei over much of the protein sequence. The availability of these assignments made comparisons of the HSQC spectra far more informative.

### 2.5.3 Dynamic light scattering (DLS) and small-angle X-ray scattering

Dynamic light scattering (DLS) is a relatively fast method of characterizing the size of biomolecules in solution, taking only minutes for a measurement. DLS may be used to distinguish between a homogenous monodisperse and an aggregated sample. This is important because quite frequently in nature, oligomeric states exist in equilibrium in solutions of biomolecules (or in a membrane) and sequence variations may affect this equilibrium with functional consequences. The apparent particle size measured by DLS is also of interest since it is modulated by molecular shape with extended macromolecules appearing to be bigger than compact macromolecules. Thus when used to compare a series of sequence variants, DLS data could reveal any gross changes in dimensions that might arise.

The principle of light scattering is based on what happens when light passes through a solution containing molecules. Depending on the optical parameters of the system, part of the light will be scattered. This scattered light may be analysed either in terms of its intensity or in terms of its fluctuations. The former is called static light scattering (see Harding, Sattelle and Bloomfield 1992). Dynamic light scattering, on the other hand (Berne, 2000; Pecora, 1985), detects the fluctuations of the scattering intensity due to the Brownian motion of molecules in solution. The statistics of the scattering signal are analysed with a correlator, and the resulting correlation function may be inverted to find a size distribution for the particles (molecules) in solution. This technique works without requiring knowledge of the exact sample concentration and has been used with success in structural biology (see, Bergfors, 1999; D'Arcy, 1994; Ferre D'Amare, 1994 for examples).

# CHAPTER 2 MATERIALS AND METHODOLOGY

In this work, the Zetasizer Nano-S system (Malvern Instruments) was used. The details of sample conditions used for collection of DLS data are given in the relevant Results section.

Small-angle X-ray scattering works on similar principles (although the theory is beyond the scope of this thesis) but is a more powerful technique with higher information content. Bead models may be constructed and scattering curves back-calculated in an attempt to define a narrow range of structures that fit with the experimental data. Samples were submitted to DESY in Hamburg for data collection on the X33 beamline of EMBL (Koch and Bordas, 1983 ; Roessle *et al.,* 2007). This experiment was kindly performed by our collaborators Dr Haydyn Mertens and Dr Dmitri Svergun, who also analysed the data using PRIMUS (Konarev *et al.,* 2003**)**, Guinier analysis (Guinier, 1939), and GNOM (Svergun, 1992), and performed the modelling as well as providing expert advice. The ab initio modelling program, DAMMIF was used for molecular weight determination whiles low resolution shape constructs were determined by ab initio beadmodelling in DAMMIF **(**Franke and Svergun, 2009**)**. To determine the most representative model from each of the ab initio methods, SUBCOMB was used **(**Kozin and Svergun, 2001**).** Also, averaged DAMMIF models were determined using DAMAVER **(**Konarev *et al.*, 2003**)** and then adjusted, so that they agree with the experimentally determined excluded volume,

using DAMFILT **(**Volkov and Svergun, 2003**).** Details of sample conditions and data analysis are found in the relevant Results section.

### 2.5.4 Analytical Ultracentrifugation (AUC)

Analytical centrifugation is similar to differential centrifugation in that both techniques apply the principles of centrifugal acceleration to separate components of a sample based on shape and mass differences. However, in analytical centrifugation analysis of the concentration of the sample during centrifugation can be performed since light detection devices are incorporated into the system; this is the key difference. Thus in analytical centrifugation two forms of hydrodynamic analysis are possible: (1) sedimentation velocity; and (2) sedimentation equilibrium. In the current work, the sedimentation velocity technique was used.

Sedimentation velocity is a powerful method because it allows both the mass and the shape of molecules to be determined – unlike DLS that conflates the two. Typical spin speeds are in the range of 40,000 - 60,000 rpm causing components to separate out in layers, forming boundaries or concentration gradients - in solution. A series of scans (based on detection of absorbance or refractive index) is performed on the sample during the spin enabling the movement of particle boundaries as a function of time (their velocity) to be recorded**.** Each resulting data set was analysed (SEDFIT, Schuck, (2000**)** This enables calculation of the sedimentation coefficient (*S*) that is a function of: molecular weight, density, molecular shape (proteins with a more elongated shape will experience more friction from solvent, so will tend to sediment more slowly), solute concentration, solvent viscosity and charge of the protein (since a charged particle will travel more quickly through a polar solvent). The theory involved and the corresponding equations are beyond the scope of this thesis. Yet, value of the partial specific volume for all four variants was computed using SEDNTERP **(**Laue *et al*., 1992**).**

In addition to determination of *S*, the diffusion coefficient (*D*) can be determined by measurement of the spreading of a boundary; this is helpful in determining, for a purified protein, the homogeneity of self-association. Thus a homogenous product (monomer, dimmer *etc.*) yields a sharper boundary than a mixture of self-associated forms in equilibrium.

Thus sedimentation velocity measurements performed on a set of variants as in the current work enabled comparisons of homogeneity and degree of self-association as well as overall shape of a protein (spherical or more extended). In the current project, samples were submitted to the University of Nottingham for measurement and expert analysis all of which was generously carried out by our collaborator, Professor Arthur Rowe. More precise details may be found in the relevant section of Results.

# CHAPTER THREE
# PROTEIN EXPRESSION AND PURIFICATION

# CHAPTER 3 PROTEIN EXPRESSION AND PURIFICATION

## 3.1 Introduction to overproduction of recombinant proteins

The production of a set of truncation mutants ("fragments") of CR1 containing various numbers of CCP modules was required for this project. A series of three fragments of increasing size (CR1 21, CR1 21-22 and CR1 20-23) was created for structural studies aimed at assessing whether the modules on either side of the exceptionally long linking sequence between LHR-C and LHR-D (CCPs 21 and 22) are folded back upon one another to create a "U-bend". These proteins also served as useful negative controls in various biological assays. Another set of smaller fragments (CR1 10-11, CR1 15-17, CR1 17 and CR1 24-25) was produced whose members correspond to parts of known or potential binding sites in CR1 and were intended for biological assays. Finally, polymorphic variants of each of two sets of longer fragments (CR1 15-25 and CR1 17-25 variants) were produced that incorporate modules contributing to structure and function, and these were used in both biophysical and functional studies.

For this project, involving functional and biophysical characterisation, multiple-milligram quantities of pure, properly folded and validated protein fragments were required. Only in the case of CR1 15-17 was an over-expressing clone (of *Pichia pastoris*) available at the start of the work – this was a generous gift from the Atkinson lab (Washington University School of Medicine, St Louis). This chapter describes the results obtained from the cloning and expression of DNA segments encoding the remaining protein fragments, along with the results from efforts at overproduction and purification of the resultant recombinant proteins. As highlighted in Materials and Methods (Chapter 2), in all cases DNA was amplified (directly or indirectly) from cDNA

encoding full length CR1, also kindly provided by the Atkinson lab; and the expression host used in all cases was the methylotropic yeast, *P. pastoris* (Schmidt *et al.,* 2011).The results presented will include images of agarose gels used to resolve and size DNA molecules, and of polyacrylamide gels following protein electrophoresis in the presence of SDS, as well as chromatograms showing protein elution profiles. Details of all these methods may be found in the Materials and Methods chapter (Chapter 2).

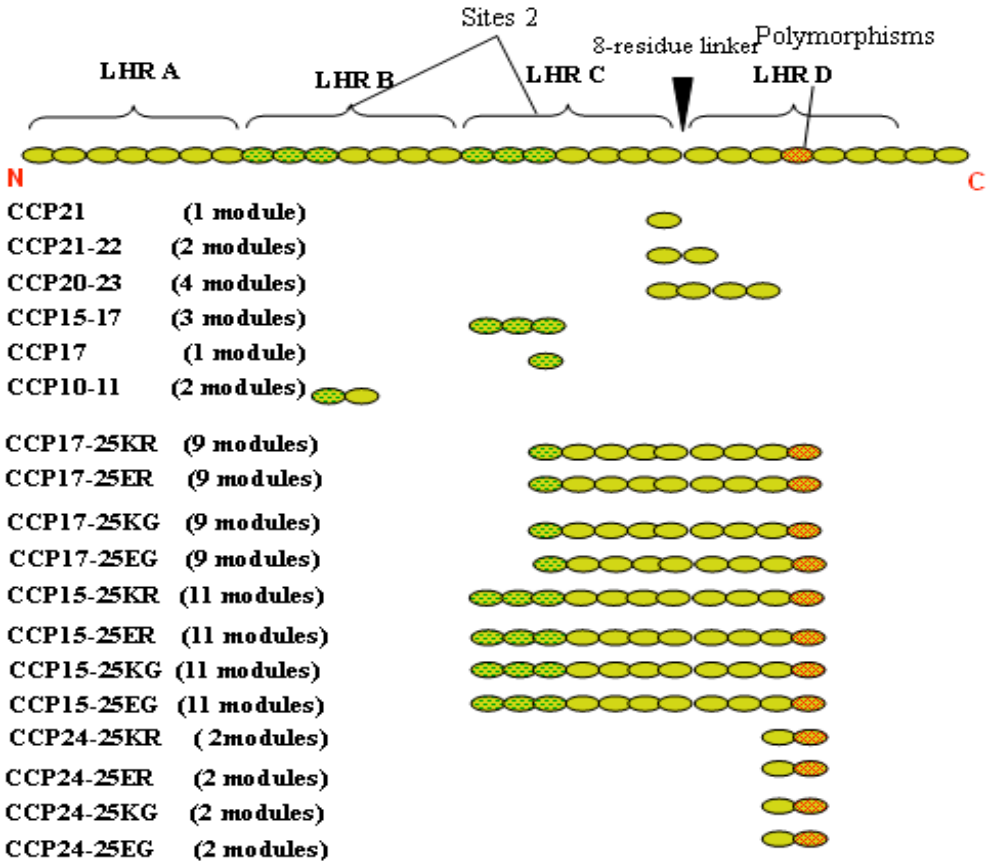



***Figure 3.1 Proteins whose DNA were cloned and expressed, and that were subsequently overproduced in P. pastoris, during the course of this project.***

A schematic representation of the 30 extracellular CCP modules (ovals) of CR1 (top) and (beneath) a summary of the recombinant CR1 protein fragments produced for the current project. The number of CCP modules in each fragment is summarised and also indicated are the extents of the long homologous repeats (LHRs A-D), the position of the two copies of functional site 2, the

locations of Knops blood group variations, and the location of the long intermodular "linker" of eight residues between LHRs C and D. The vertical arrangement (from top to bottom) of the constructs reflects the order in which the constructs were recombinantly made and purified.

Since cloning and overexpression of genes in *P. pastoris*, and the subsequent production and purification of recombinant protein, appeared to become more challenging as the target proteins incorporated more modules, initial work focussed on obtaining fragments consisting of single- and double-CCP modules. As this proved very successful, efforts extended incrementally to encompass longer constructs culminating in successful production of a set of eleven-CCP module fragments.

Some vital statistics for the set of recombinant CR1 fragments employed in the current study are summarised in Table 3.1.

| Name of Construct | Number of Amino Acid (AA No) | DNA Molecular Weight (bp) | Protein Molecular Weight (in KDa) | Theoretical PI | Extention Coefficient (ε) | N-Glycosylation Sites |
|---|---|---|---|---|---|---|
| CR1 21 | 82 | 246 | 8.790 | 5.38 | 8730 | No |
| CR1 21-22 | 143 | 429 | 15.65 | 5.09 | 17460 | No |
| CR1 20-23 | 264 | 792 | 28.76 | 5.90 | 30450 | Yes |
| CR1 15-17 | 197 | 591 | 21.70 | 8.46 | 27680 | Site removed |
| CR1 17 | 82 | 246 | 8.913 | 6.75 | 10220 | No |
| CR1 10-11 | 139 | 417 | 15.20 | 6.82 | 15970 | Yes |
| CR1 17-25KR | 591 | 1773 | 64.88 | 6.12 | 71120 | Yes |
| CR1 17-25ER | 591 | 1773 | 64.88 | 5.96 | 71120 | Yes |
| CR1 17-25KG | 591 | 1773 | 64.88 | 6.04 | 71120 | Yes |
| CR1 17-25EG | 591 | 1773 | 64.88 | 5.88 | 71120 | Yes |
| CR1 15-25KR | 711 | 2133 | 78.10 | 6.86 | 88580 | Yes |
| CR1 15-25ER | 711 | 2133 | 78.10 | 6.62 | 88580 | Yes |
| CR1 15-25KG | 711 | 2133 | 78.10 | 6.73 | 88580 | Yes |
| CR1 15-25EG | 711 | 2133 | 78.10 | 6.53 | 88580 | Yes |
| CR1 24-25KR | 137 | 411 | 15.0 | 6.07 | 15970 | Yes |
| CR1 24-25ER | 137 | 411 | 15.0 | 5.49 | 15970 | Yes |
| CR1 24-25KG | 137 | 411 | 15.0 | 5.75 | 15970 | Yes |
| CR1 24-25EG | 137 | 411 | 15.0 | 5.25 | 15970 | Yes |

*Table 3.1 Summary of "vital statitics" for the protein fragments made for the current study*

| Name of Construct | First residue name | Last residue name | No and location of N-glycosylation sites | |
|---|---|---|---|---|
| CR1 21 | E 1317 | R 1392 | 0 | **NONE** |
| CR1 21-22 | H 1317 | S 1456 | 0 | **NONE** |
| CR1 20-23 | S 1257 | I 1516 | 3 | **$N_{1310}$-$S_{1312}$ , $N_{1481}$- $S_{1483}$, $N_{1504}$-$S_{1506}$** |
| CR1 15-17 | L 940 | N 1136 | 2 | **$N_{959}$-$S_{961}$, $N_{1028}$-$S_{1030}$ ( But Knocked Out)** |
| CR1 17 | R1063 | N 1136 | 0 | **NONE** |
| CR1 10-11 | R 613 | R 745 | 1 | **$N_{702}$-$S_{704}$** |
| CR1 17-25KR | I 1064 | S 1647 | 7 | **$N_{1152}$-$S_{11454}$ , $N_{1215}$-$S_{1217}$, $N_{1310}$-$S_{1312}$ , $N_{1481}$- $S_{1483}$, $N_{1504}$-$S_{1506}$, $N_{1534}$-$T_{1536}$ $N_{1540}$-$T_{1542}$** |
| CR1 17-25ER | I 1064 | S 1647 | | |
| CR1 17-25KG | I 1064 | S 1647 | | |
| CR1 17-25EG | I 1064 | S 1647 | | |
| CR1 15-25KR | T 940 | S 1647 | 9 | **$N_{959}$-$S_{961}$, $N_{1028}$-$S_{1030}$ $N_{1152}$-$S_{11454}$ , $N_{1215}$-$S_{1217}$, $N_{1310}$-$S_{1312}$ , $N_{1481}$- $S_{1483}$, $N_{1504}$-$S_{1506}$, $N_{1534}$-$T_{1536}$ $N_{1540}$-$T_{1542}$** |
| CR1 15-25ER | T 940 | S 1647 | | |
| CR1 15-25KG | T 940 | S 1647 | | |
| CR1 15-25EG | T 940 | S 1647 | | |
| CR1 24-25KR | I 1517 | S 1647 | 3 | **$N_{1534}$-$T_{1536}$ $N_{1540}$-$T_{1542}$ $N_{1605}$-$S_{1607}$** |
| CR1 24-25ER | I 1517 | S 1647 | | |
| CR1 24-25KG | I 1517 | S 1647 | | |
| CR1 24-25EG | I 1517 | S 1647 | | |

*Table 3.2 Domain boundaries and N-glycosylation sites*

## 3.2 CR1 CCP 21 cloning, production and purification

The 21st CCP of CR1 (CR1 21) is the last module of LHR-C (see Fig. 3.1) and is joined to CCP 22 by an eight-residue linker. It was decided to prepare this single module (for domain boundaries used, see Table 3.2) with a view to a future NMR spectroscopy-based comparison of its chemical shifts with those of the recombinant double-module construct CR1 21-22.

DNA was amplified with the forward and reverse primers indicated in Table 2.1 (see Chapter 2, Materials and Methods). The amplification product (see Fig. 3.2A) migrates as a band of DNA on an agarose gel to a position consistent with its expected length of 246 base pairs. Subsequently, TOPO®isomerase (TOPO®) cloning (as described in the manual supplied with the TOPO® Cloning Kit, Invitrogen) was performed to place the PCR-derived insert into the TOPO® plasmid vector that was then used to transform Top10 *E. coli* competent cells (Invtrogen). The amplified and subsequently extracted DNA (using the QIAprep Miniprep Kit (Qiagen)) was digested using *Pst*I and *Xba*I. The digested insert was ligated into commercially available pPICZ αB vector (see Materials and Methods for more details). This was amplified in Top 10 *E. coli* competent cells and DNA extracted as before. Following linearization (*SacI*), the recombinant vector was transformed into *P. pastoris* KM71H cells for expression. To ensure that the transformation was successful, PCR screening was done on the *P. pastoris* colonies using, as forward and reverse primers, DNA encoding the α factor and the AOX

promoter, respectively; a band of 246 + 330 base pairs would therefore be expected for a successful insertion. In Figure 3.2B, therefore, the bands in lanes 1 and 2 are consistent with insert-positive colonies that were taken forward for protein-production trials.

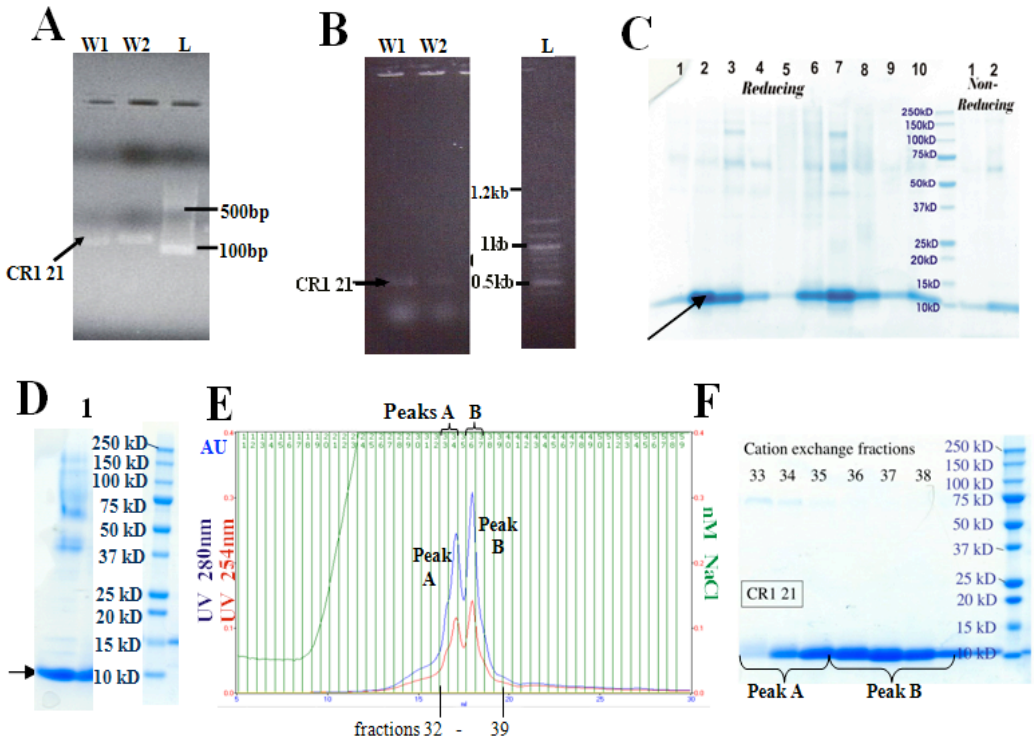



***Figure 3.2 Cloning, production and purification of CR1 21***

(A) PCR amplification of CR1 21-encoding DNA from CR1 cDNA; bands in well (W) 1 and W2 correspond to the expected 246-bp DNA; L is a 100-bp ladder (which ran poorly for reasons unknown, producing diffuse bands). (B) PCR screening of CR1 21-encoding DNA transformed *P. pastoris* colonies; the 576-bp bands in W1 and W2 corresponds to the expected size of the PCR prodict. (C) Gradient SDS-PAGE performed on TCA precipitated product derived from 1 ml of cell culture supernatant (resuspended in about 20 μL of 2X SDS loading buffer) from ten separate "mini-scale" (5 mls each) CR1 21 protein-production trials; an arrow indicates the position of protein bands at the expected mobility (molecular weight markers, in kD, are indicated) The right-hand lanes show that the protein (*e.g.* from colonies 1 and 2) still runs as a single band under reducing conditions (containing a final concentration of about 100 mM DTT). (D) SDS-PAGE of about 20 μL of a 10-fold concentrated aliquot (not purified) taken from growth media supernatant from a larger (one litre)-scale production run of CR1 21. (E) Chromatogram obtained after loading crude recombinant CR1 21 protein onto a cation-exchange (MonoS – details in Chapter 2) column in 20 mM sodium acetate buffer (pH 4.0) and eluting with a salt (NaCl) gradient from 0 to 1 M NaCl; (F) Commassie-stained SDS-PAGE of the fractions corresponding to the two major peaks, A and B, eluted from the cation-exchange chromatography column in panel E.

From the SDS PAGE gel (Figure 3.2C), it is apparent that all colonies screened in the mini-scale protien production trials expressed the CR1 21 gene well, with colonies numbers 2 and 7 subsequently selected for larger -scale protein production and the laying down of glycerol stocks. Note that CCP 21 has an expected MWt of 8-9 kDa and so as expected runs just ahead of the first band of the "Precision Plus" protein molecular weight markers (~10 kDa).

Using a shaker flask, a larger culture was set up leading to a one-litre harvest. Harvesting and preparation for purification were carried out as described in section 2.3.7 of Materials and Methods (Chapter 2; note similar methods were also used for all the other CR1 fragments). Before any purification commenced, the supernatant was analysed by SDS-PAGE to provide the basis upon which to very approximately estimate the level of protein production (see Fig. 3.2D). Lane 1 in Figure 3.2D contains a band running just below 10 kDa. This corresponds to the expected molecular weight of CR1 21 (see Table 3.2). Subsequently, a one-in-five dilution of the filtered supernatant was performed to decrease its ionic strength (and the conductivity checked), then the pH was adjusted to 4. (Note the predicted pI of the construct is 5.38 (Table 3.1)). This sample was subjected to cation-exchange chromatography performed on a self-poured SP-sepharose column with a step elution, using 1 M NaCl. The eluted product was de-salted and run through a second higher-resolution cation-exchange column (monoS, size and supplier). Two overlapping peaks (named A and B) were obtained (Fig. 3.2D) corresponding to fraction numbers 34 and 36, respectively. Peak B looked fairly pure according to SDS-PAGE and therefore fractions 36-37 were pooled for further

characterisation. The overall yield of purified protein from this one-liter growth was estimated to be 0.8 mM in a total volume of 1 ml.

## 3.3 CR1 21-22 cloning, production and purification

The 21st and 22nd CCP modules of FH form the boundary between LHRs-C and D, and, as was discussed above, are connected by the longest inter-modular linker amongst all RCA proteins. Using the appropriate primers (Table 2.1 in Chapter 2), the segment of the cDNA that corresponds to the DNA encoding CCP 21-22 of CR1 was amplified (See Table 3.2 for chosen domain boundaries) by PCR. The DNA product contains 429 bp and ran accordingly on an agarose gel (*i.e.* it was larger than the insert encoding CR1 21, see Fig. 3.3A). This PCR-amplified product was cleaned up, TOPO®-cloned as before, sequenced, and restriction-enzyme digested (using *Pst*1 and *Xba*1) using similar procedures to those applied to CR1 21. The insert was ligated into the *P. pastoris* vector pPICZ αB and transformed into top 10 *E. coli* competent cells for "maxi-prep" using the QIAprep®maxiprep Kit (Qiagen).

The resultant plasmids were linearised using *Sac*I and the products run out on a 1% (w/v) agarose gel alongside the non-linearised (uncut) plasmid that ran slightly faster, as expected. On this basis, the two highest expressing colonies were selected for further work (see Fig. 3.3B). Following clean up of these linearised DNA samples by phenol-chloroform extraction and ethanol precipitation (see Materials and Methods) they was used to transform *P. pastoris* strain KM71H. Screening by PCR, using Mastermix (Promega), of the resultant *P. pastoris* colonies was performed to ensure that the

recombinant gene had been successfully inserted prior to expression trials. Two colonies (see Fig. 3.3C) (in wells 1 and 2) screened positive for the presence of insert.

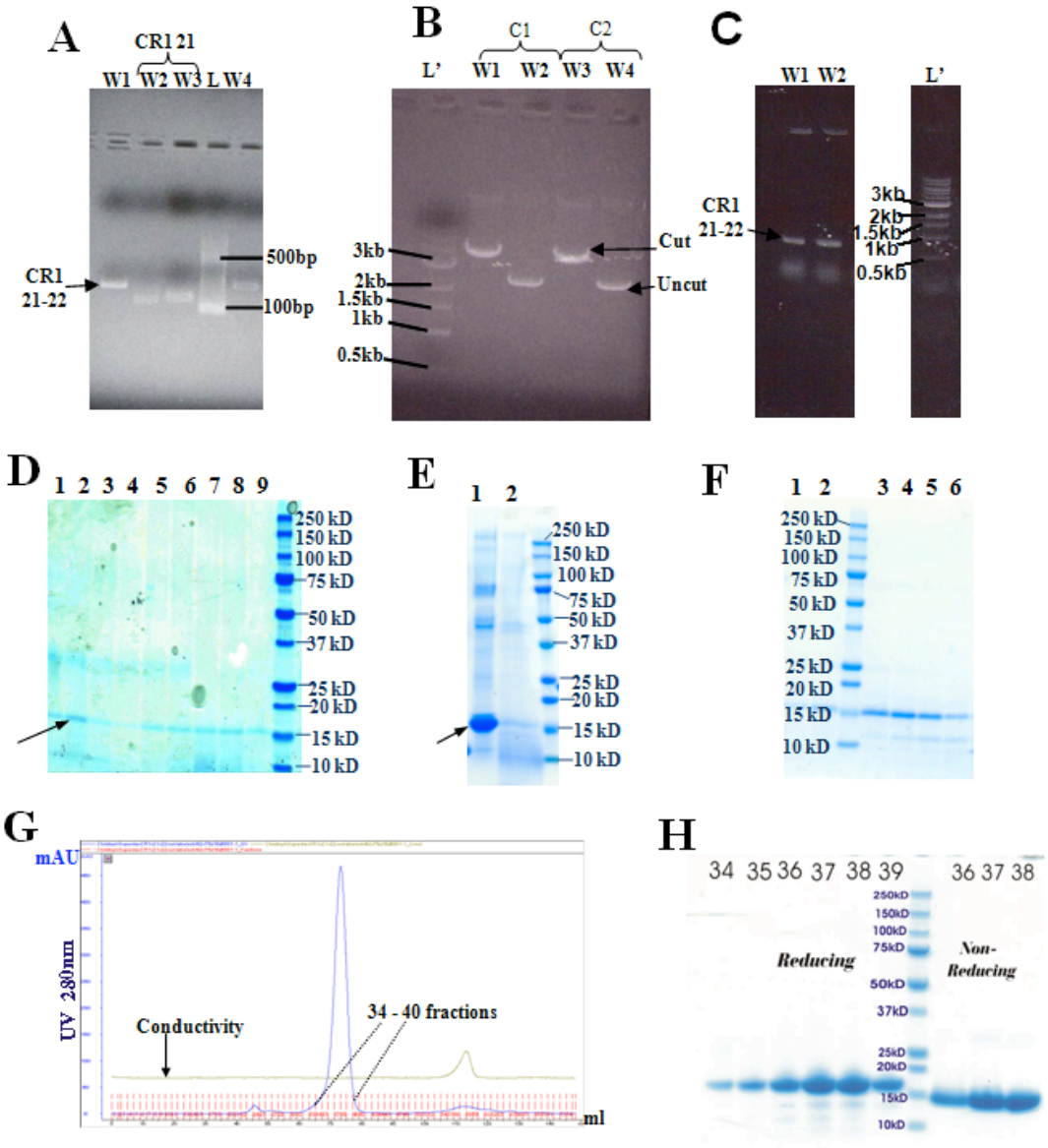



***Figure 3.3 Cloning, production and purification CR1 21-22***

(A) Amplification of CR1 21-22-encoding DNA from CR1 cDNA; bands in W1 and W4 correspond to the 429-bp of CR1 21-22. Arrow indicates a band running in the expected place and presumed to correspond to the insert coding for CR1 21-22. (B) *Sac*I digest (linearisation) of the CR1 21-22-containing plasmid; W1 and W2 correspond to cut and uncut plasmids from colony C1 whilst W3 and W4 correspond to cut and uncut copies of C2, respectively. L' is 1-kb ladder. (C) PCR screening of transformed *P. pastoris* colonies; the 759-bp bands in W1 and W2 corresponds to the expected size of insert plus flanking sequences; as before, L' is 1 kilo-bp ladder. (D) Gradient SDS-PAGE of mini-scale production trials of nine *P. pastoris* CR1 21-22

colonies; arrow indicate the faint bands running where the product would be expected. (E) SDS-PAGE, performed on a gradient gel, of concentrated but crude supernatant (after spinning out cells) from CR1 21-22 large-scale production trial with strong band (arrow; lane 1) running as expected for this size of protein and only a faint band from an unconcentrated supernatant (lane 2). (F) Results of SDS-PAGE perfomed on fractions obtained from cation-exchange purification of CR1 21-22; these fractions were eluted from a MonoS column with 20 mM sodium acetate buffer (pH 4.0), with 1 M NaCl. (G) Chromatogram obtained for the size-exclusion purification of CR1 21-22. (H) SDS PAGE following size-exclusion chromatography of (non-isotopically labelled) CR1 21-22.

Assuming successful insertion, the expected number of basepairs for the PCR product (using, as before, α-factor-derived forward primers and AOX promoter-derived reverse primers) is 759 bp, which is consistent with the band running just below the 1 kilo-bp marker in Figure 3.3C.

After a successful mini-scale protein production trial, however, it became clear that the yield of recombinant protein was relatively low, as evidenced by faint bands on the gel (see Fig. 3.3D). Colonies were subsequently re-streaked on YPD-agar plates containing higher amounts of Zeocin™ (200–300 μg/ml) in order to achieve more stringent selection of high-copy number clones. Selected colonies from the YPD-agar plate were then picked and cultured in shaker flasks; the cells were spun out and aliquots of the supernatants were concentrated and analysed by SDS-PAGE (see Fig. 3.3E). Yields improved following additional colony selection procedure. Similar purification steps to those used for CR1 21 were applied, since the theoretical pI of this construct is 5.1 (see Table 3.1). Thus cation-exchange chromatography was used for the first two purification steps employing, sequentially, a self poured SP-Sepharose column (between 10 ml of resin), and then a higher resolution monoS column. An SDS-polyacrylamde gel was run on the fractions eluted from the monoS column (Fig. 3.3F) revealing relatively pure protein with a degree of degredation and other impurities. Following buffer-

exchange into PBS, size-exclusion chromatography was performed as a polishing step (see Figs. 3.3G and 3.3H for the chromatogram and corresponding analysis of fraction using SDS-PAGE.) Note that purified proteins runs at the expected sites (about 15 kDa) under non-reducing conditions, while under reducing conditions bands run slightly higher, as expected, and there is little or no evidence of proteolytic clipping of CR1 21-22. The overall yield of purified protein was estimated to be 1 mM in 1.0 ml.

## 3.4 CR1 20-23 cloning, production and purification

To gain further knowledge of the architecture surrounding the long intermodular linker between LHR-C and LHR-D, the four-module construct CR1 20-23 was also needed. As before PCR was employed to amplify the appropriate segment of DNA (see Table 3.2 for the domain boundaries chosen) from the cDNA for full-length CR1. A product consistent with the expected 792 bp was obtained (see Fig. 3.4A). TOPO®-cloning and double digestion was used as before to produce an insert of appropriate size (see W2 in Fig. 3.4B) for ligation into the pPICZ αB vector and amplification in *E. coli*; note, a modified ligation protocol (see Materials and Methods) was used and PCR was employed to check for successful ligation. As shown in Figure 3.4C, vectors from *E. coli* colonies 1, 2, 7, 8, 10, and 11 all appeared to contain the insert on the grounds that the PCR products ran in the expected place (~1.2 kb). Several of these colonies were then checeked to verify that the sequences were indeed correct and that they were in frame. Plasmids were subsequently recovered, *Sac*I linearized, and run on a 1 % agarose gel alongside the uncut version (see Fig. 3.4D). After successful transformation into *P. pastoris,* another PCR-

based screen was carried out using forward and reverse primers precisely as described earlier - the band runs just above the 1 kbp marker as expected for a 1.2 kb segment of DNA.

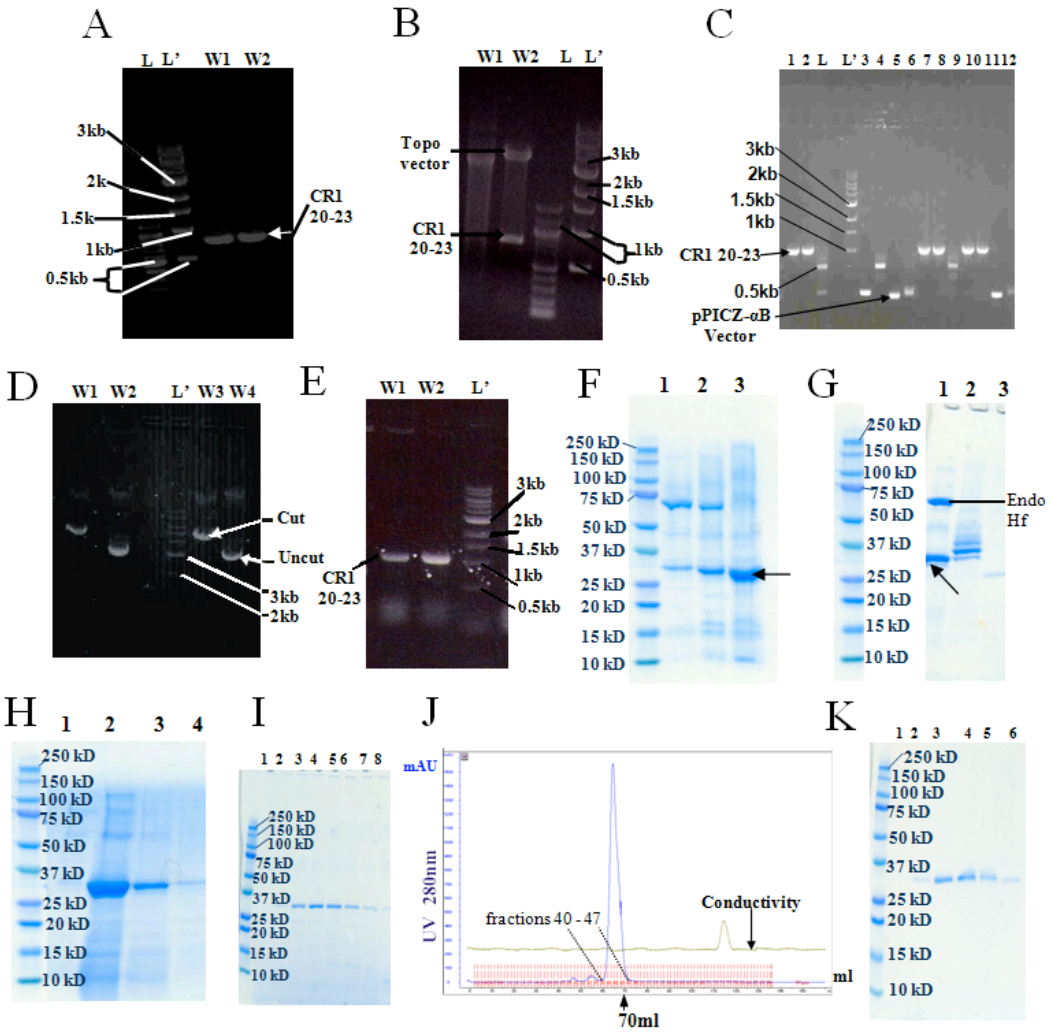



***Figure 3.4 Cloning, production and purification of CR1 20-23***

(A) PCR amplification of CR1 20-23-encoding DNA from cDNA. A band consistent with the expected size of 792 bp was obtained. (B) TOPO®-cloning and double digestion using *Pst*1 and *Xba*1 yields a band of the expected size. (C) Screening for successful ligation by PCR of plasmids produced in *E. coli*. (D) *Sac*1 linearisation of PicZ αB containing insert for CR1 20-23, (E) Screening by PCR of PicZ αB from *P. pastoris* that contain the DNA encoding CR1 20-23, L and L' denote size-marker ladders. (F) SDS PAGE following "mini-scale" production trial of CR1 20-23 (contents TCA-preicpited from 1 ml of culture superntant from a 5 mL culture). (G) SDS PAGE analysis following CR1 20-23 production in a one-litre shaking flask cell culture; lane 1 contains $EndoH_f$ deglycosylated protein while (for comparison) lane 2 contains the crude glycsosylated material. (H) Commassie-stained SDS PAGE following the first cation exchange chromatography step. (I) Commassie-stained SDS PAGE following second cation-exchange on

monoS column. (J) Chromatogram following size-exclusion chromatograph of CR1 20-23. (K) SDS PAGE confirming that the main peak obtained following size-chromatography contains pure protein with an expected mass of about 30 kDa.

Colony 3 (see Fig. 3.4F) was used for larger-scale (one-litre) protein production growth and the harvested supernatant (*i.e.* after spinning out cells) was diluted and subjected to a step-elution cation-exchange purification (pI =5.90, pH = 4.0) using a self-poured bench-top SP-sepharose column (10ml of resin), as before. The eluted protein (see Fig. 3.4G) was then deglycocylated using endo $H_f$ (two hours incubation at 37 °C). Note that this deglycosylation step had not proved necessary in the cases of CR1 21 and CR1 21-22 because they did not have any N-glycosylation sites. After de-salting, a second cation-exchange chromatography step was performed on a monoS column, in an identical fashion to that used for purification of CR1 21 and CR1 21-22. SDS-PAGE analysis confirmed the production of a protein band of the expected size for deglycosylated material (~28 kDa, see Fig. 3.4I and Table 3.1). Finally the pooled fractions, after buffer exchange into PBS, were subjected to a size-exclusion ”polishing” step on a Hiload Superdex 75 column (Fig. 3.4J ) and SDS-PAGE was performed on the resultant fractions (Fig. 3.4K). The yield of pure protein from this one-litre growth was 0.24 mM in 1 mL.

## 3.5 CR1 15-17 Cloning, Production and Purification

Modules 15-17 of CR1 encompass one of two copies of functional site 2 and the DNA encoding CR1 15-17, already cloned into the *P. pastoris* expression vector pPIC9, was kindly provided by John Atkinson. Therefore transformation of this construct had been

entirely accomplished by the Atkinson group. A one-litre culture was grown up in a shaker flask (see Materials and Methods), cells removed by centrifugation, and protein of the correct size detected in the supernatant (Fig. 3.5A). An initial bench top cation-exchange chromatography step was carried out on diluted cell culture supernatant using 25 mM MES buffer (pH 5.0, the pI of CR1 15-17 is 8.46) as the binding buffer and eluting with a gradient to 1 M NaCl yielding the SDS-PAGE analysis shown in Figure 3.5B. Pooled protein-containing fractions from this step were buffer exchanged into PBS and subject to a size-exclusion (HiLoad Superdex 75 column) chromatography step (see Fig. 3.5C). The main peak (covering fractions 30 to 36) was analysed by SDS–PAGE and found to contain one very dominant band of the expected mass, under both reducing and non-reducing conditions (Figure 3.5D).

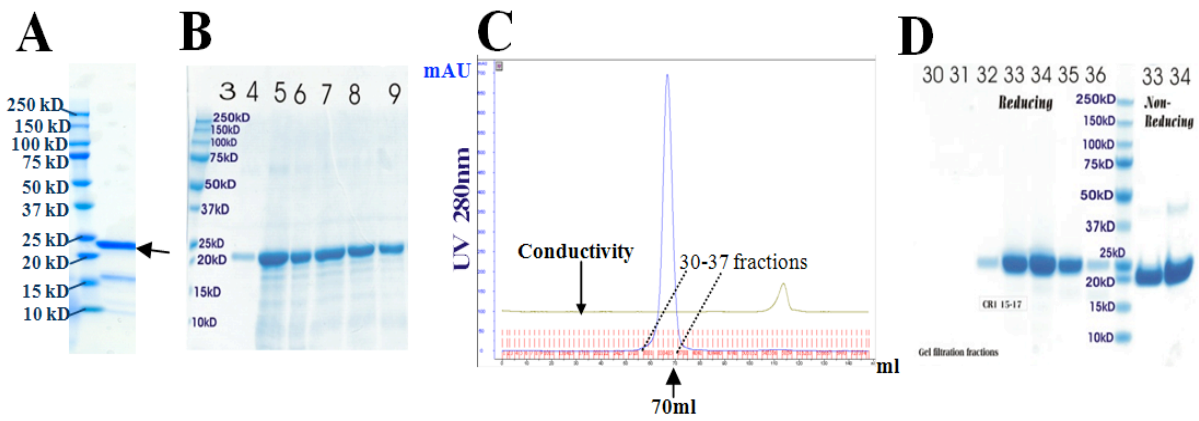



***Figure 3.5 Purification of CR1 15-17 by ion exchange and gel filtration***

(A) A Commassie-stained SDS-polacylamide gel of crude supernatant from a one-litre culture (12 µL of culture loaded in each well); arrow indicates position of the expected bands. (B) A SDS-polacylamide gel of the first ion-exchange based purification (SP sepharose). (C) Chromatogram following the final size-exclusion chromatographic (Superdex 75) purification of CR1 15-17. (D) Purity of CR1 15-17 as assessed by SDS-PAGE under reducing (left) and non-reducing conditions.

As expected for a well-folded protein with intact disulphide bonds, the reduced proteins runs slightly higher than the non-reducing ones. However, there does appear to be a contaminating band, of approximately dimer size, in the non-reducing lanes.

## 3.6 CR1 17 cloning, production and purification

Among the many antibodies that recognize CR1, J3B11 has CCP 17 as its epitope. Interestingly, J3B11 effectively disrupts rosettes. While all three CCPs (15-17) of site 2 are required for C3b interaction, soluble recombinant CR1 17 alone might have some effect on rosetting. The DNA encoding CCP 17 (for boundaries, see Table 3.2.) was amplified both from a cloned CR1 17-25-encoding template (see below) and from CR1 cDNA. All three lanes in Figure 3.6A (W1 and W2 from cDNA and W3 from CR1 17-25) contain DNA bands of an approximately appropriate size. Double digest (using *Pst*I and *Xba*I) of the DNA extracted from three of the TOPO®-cloned colonies are shown in Figure 3.6B wherein lanes 1 and 3 contained the digested DNA insert coding for CR1 17.

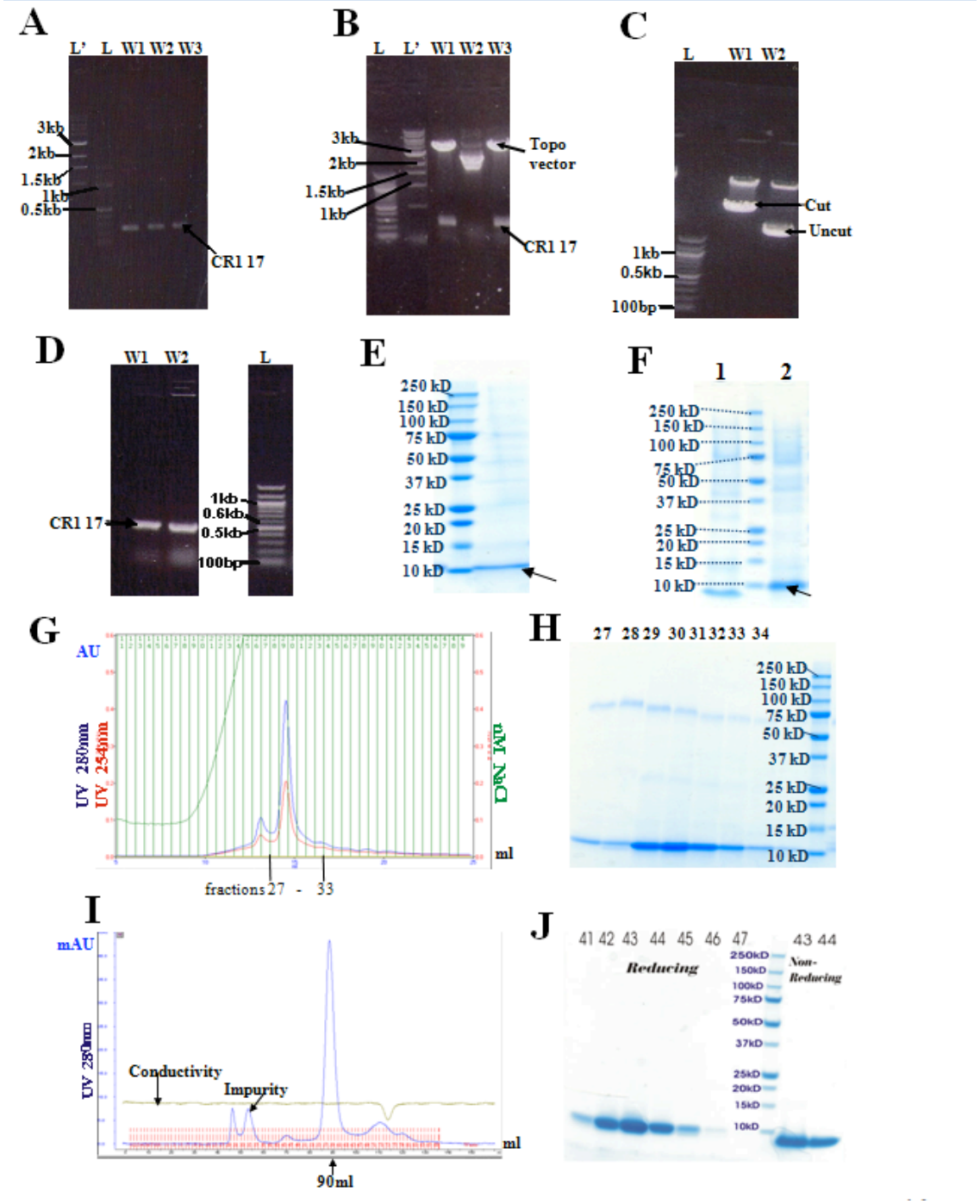



***Figure 3.6 Cloning, production and purification of CR1 17***

(A) PCR amplification of DNA coding for CR1 17 from cDNA for CR1 (well (W)1 & W2) and DNA encoding CR1 17-25 (W3). (B) Double digest of TOPO®-cloned CR1 17. (C) *SacI* digest of plasmid containing insert for CR1 17. (D) PCR-based screening of vectors containing DNA for CR1 17 from *P. pastoris* colonies. Note: L represents 100-bp ladder, L' represents 1-kbp ladder. (E) SDS-PAGE of unpurified supernatant from one-litre culture (shaker flask) for production of CR1 17 (12µL of culture supernatant loaded in each well). (F) SDS-PAGE analysis

of the outcome of first cation-exchange chromatographic purification performed as described earlier for CR1 21 and CR1 21-22. (G) Chromatogram to show outcome of the second cation-exchange step (monoS) ( pI =6.75 and hence pH = 4.0). (H) SDS-PAGE of fractions from the second cation-exchange purification of CR1 17. (I) Chromatogram obtained from size-exclusion chromatography of CR1 17. (J) SDS-PAGE performed on protein obtained from one of this final purification step.

As before, after ligation into pPicZ αB, screening and sequencing, plasmids containing DNA (for CR1 17) were extracted from the selected *E. coli* colony and *Sac*I digested (Fig. 3.6C), then purified and transformed into *P. pastoris*. Colonies obtained from the transformation were screened (Fig. 3.6D) and trial expression instigated.

The one-litre shaker flask supernatant was analysed by SDS PAGE (Fig. 3.6E), then subjected to cation-exchange chromatography (on a self-packed SP sepharose column as before). Since CR1 17 has a theoretical PI of 6.75, 20 mM sodium acetate buffer, pH 4, was used (Fig. 3.6F). As before, a second cation-exchange chromatographic step (on MonoS) was undertaken (see Figs. 3.6G and H), followed by a size-exclusion chromatographic step (Figs. 3.6I and J).

## 3.7 CR1 10-11 cloning, production and purification

Because of the near-identical sequences of LHRs B and C, it was possible that an attempt to amplify modules 17-18 from cDNA could result in amplification of modules 10-11. In anticipation of this difficulty, more than one tube were set up for the amplification of the requisite DNA region encoding modules 17-18 of CR1, one (well (W)3) used the CR1 17-25 coding sequence as its template whilst two others (W1 and W2) contained the cDNA of full-length CR1 as template (Figure 3.7A). The amplification of the cDNA for

CR1 worked well; after TOPO®-cloning and sequencing, however, it became apparent that the portion amplified from the cDNA was actually CR1 10-11 found in LHR-B and not the expected CR1 17-18 from LHR-C. Given the near-identical sequences, it was decided to continue anyway and make CR1 10-11 instead of CR1 17-18. A double digest was performed on the TOPO® plasmid vector (pCR®4Blunt-TOPO®) containing CR1 10-11-encoding DNA using *Pst*I and *Xba*I (Fig. 3.7B) as described in the previous sections. One of the successfully digested inserts (W1 and W3 of Fig. 3.7B) was then ligated into the pPICZ αB vector as before and this was used to transform *E. coli* cell. Once again, recombinant plasmids extracted (QIAprep Maxiprep Kit) from the tramsformed *E. coli* were linearised using *Sac*I (Fig. 3.7C) and transformed into *P. pastoris* strain KM71H. Successful screening on Agar-YPD plates containing 200 µg/ml Zeocin™ was performed, followed by PCR-based screening of *P. pastoris* colonies (Fig. 3.7D).

"Mini-scale" protein production trials indicated a good yield of highly glycosylated proteins. Enxouraged by this and using the best clone, a one-litre shaker flask culture was prepared, yielding the SDS-PAGE results shown in Figure 3.7E. As before, a bench-top self-poured SP Sepharose column was used to capture the glycsosylated protein from diluted supernatant (pI = 6.82, 20 mM sodium acetate buffer, pH 4.0). Following deglycosylation with Endo $H_f$ a second cation-exchange chromatographic step was performed as before (Figs. 3.7F and G, respectively). Figures 3.7H, I and J show the outcome of two sequential size-exclusion chromatographic steps. The second size-exclusion step was thought necessary in order to remove so far as possible any degraded product (see reducing side of Fig. 3.7H), and possibly the "ladder

like" impurities detected on the non-reducing side (Fig. 3.7H). From the one-litre culture, a solution of 0.61 mM purified protein in 1ml was obtained.

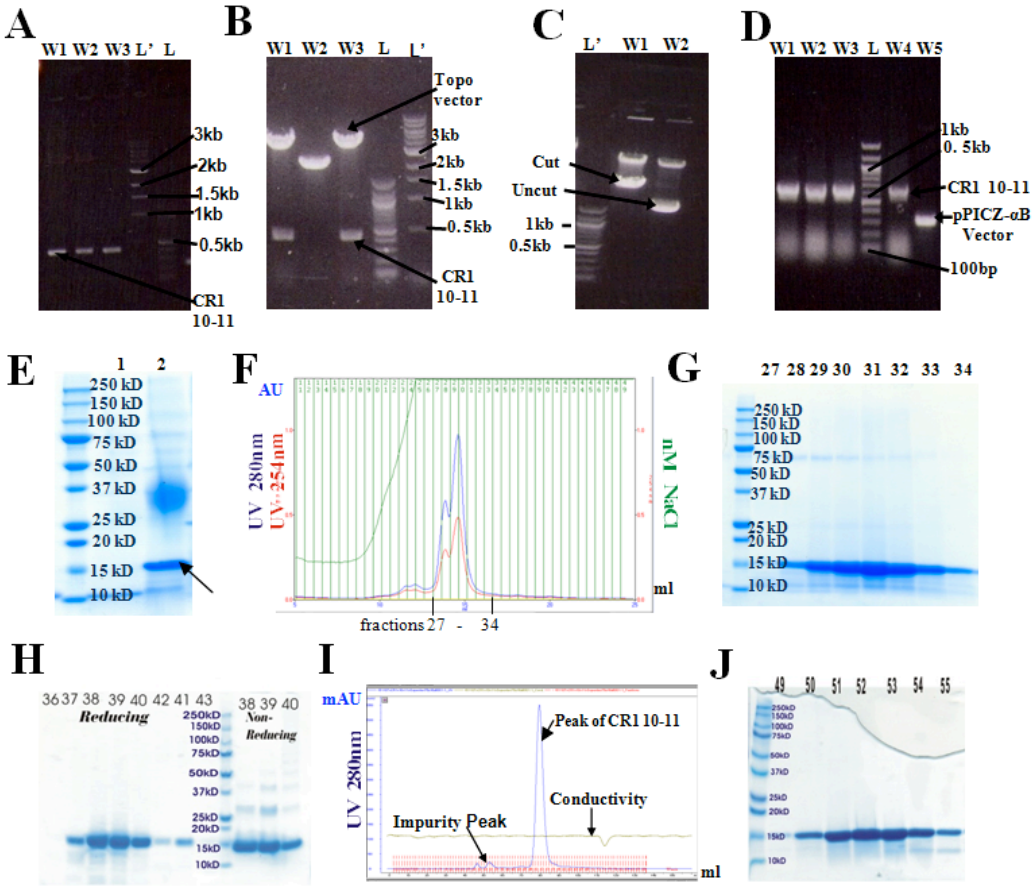



***Figure 3.7 Cloning, production and purification of CR1 10-11 (17-18)***

(A) PCR amplification of DNA encoding CR1 10-11 from cDNA (W1 and W2) and CR1 17-25 (W3). DNA of 417 bp was expected. (B) Double digest of CR1 10-11 DNA following TOPO®-cloning, using *Pst*I and *Xba*I. (C) *Sac*I llinearization of plasmid containing DNA for CR1 10-11. (D) PCR screening of CR1 10-11 insert following transformation of *P. pastoris*. Arrows indicate the expected bands. (E) Coomassie-stained SDS-PAGE performed on unpurified supernatant from a one-litre culture of CR1 10-11 (12 µL loaded); lane 2 contains glycsosylated protein and the main band has a smeared appearance at around 37 kD. (F) Chromatogram following cation-exchange purification (monoS). (G) Outcome of Coomassie-stained SDS-PAGE performed on cation-exchange chromatography fraction from the monoS column. (H) Coomassie-stained SDS PAGE following first size-exclusion purification of CR1 10-11. (I) Chromatogram following second size-exclusion purification of CR1 10-11. (J) Coomassie-stained SDS PAGE of fractions collected from second gel filtration chromatography step.

## 3.8 CR1 17-25 cloning, production and purification

Using the appropriately designed oligonucleotide primers (see Table 2.1), the DNA sequence coding for the Caucasian variant of this nine-module construct (CR1 17-25; was amplified (in triplicate – to increase chances of success since this is a much longer construct than had been made previously) using PCR from CR1 cDNA, and the product run out on an agarose gels (Figure 3.8). Two bands were observed in well 1 (W1 of Fig. 3.8A) because the primers were able to anneal with more than one part of the DNA template. For similar reasons, non-targeted amplification was observed in wells 2 and 3. The band of the expected size (about 1773 bp, table 3.1) was cut out and purified for further work.

TOPO®-cloning and double digestion using *Pst*1 and *Xba*1 were performed as described in previous sections (Fig. 3.8B). Colonies obtained from ligation into pPicZ αB and the transformation into Top 10 chemically competent *E. coli* cells for amplification were PCR screened (Figure 3.8C). Colonies 3, 4 and 7 screened positive and therefore were selected for sequencing. After selecting one of the sequenced colonies, the plasmid was amplified via a "maxi-prep", *Sac*I linearised (Figure 3.8D) and used to transform *P. pastoris*. The transformed *P. pastoris* colonies that had grew on agar-YPD plates with 300µg/ml Zeocin™ were re-screened using PCR (Figure 3.9E). Protein yields in small-scale production trials were low (Figure 3.8F). Therefore western blot was employed using the 3D7 CR1 antibody (Figure 3.8G). Lanes 2 to 4 reveal CR1 17-25 production by three colonies, and may be compared to lane 1 that shows production of CR1 20-23, serving as positive control.

# CHAPTER 3 PROTEIN EXPRESSION AND PURIFICATION

Once it has been established that this "Caucasian" variant of the construct (*i.e.* one that predominates in Caucasian populations and corresponds to CR1 17-25 containing K1590 and R1601, or CR1 17-25KR) had been successfully cloned, expressed and purified, attempts were made to introduce the point mutations needed to express the other Knops blood group polymorphic variants. Site-directed mutagenesis was carried out to replace the appropriate base pairs; thus an A4795G (*i.e.* A4795 to G) substitution leading to Glu1590 CR1 17-25ER, A4828G leading to Gly1601 CR1 17-25KG**,** and the double substitution leading to (Glu1590E, Gly1601G) CR1 17-25EG, were generated (see Materials and Methods). Plasmids were sequenced to ensure that the correct changes had been effected them amplified in Top 10 *E. coli* cells and linearised prior to *P. pastoris* transformations. Small-scale growths were performed for all three mutated versions (variants) (CR1 17-25ER, KG and EG). The results of SDS-PAGE performed on the products of the "mini-scale" production trials of CR1 17-25ER (left half of gel) and CR1 17-25KG (right half of gel) are shown in Fig. 3.8H, while Panel (I) is the product of ``mini-scale" production trials on CR1 17-25EG. To confirm expression of these proteins - that were produced in relatively low yields, more sensitivity was needed and so western blots (details in Materials and Methods) were performed following SDS-PAGE of harvested supernatants from two colonies of each variant (Figure 3.8J).

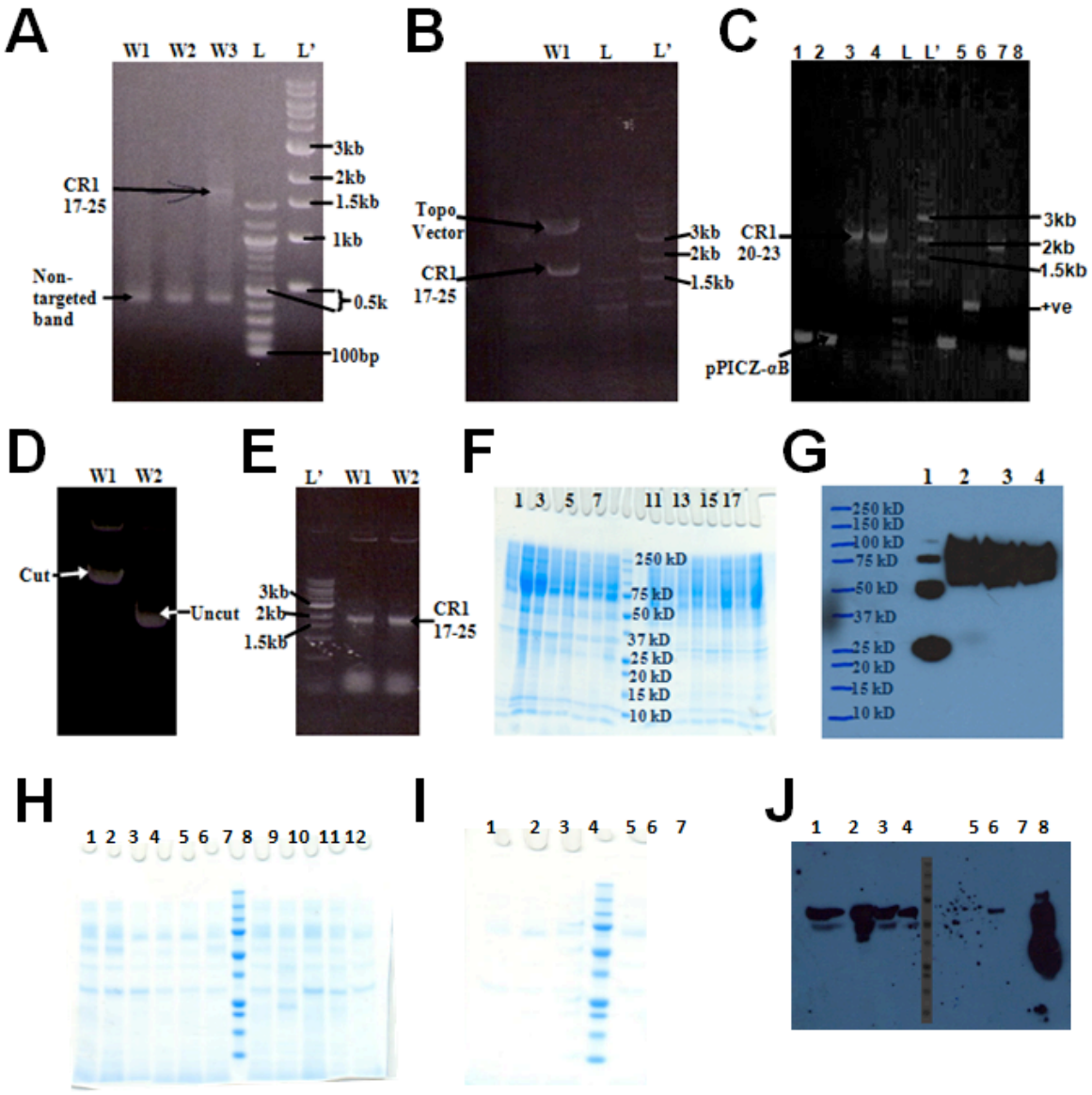



***Figure 3.8 Cloning, production and purification of CR1 17-25 and polymorphic forms***

(A) PCR amplification of DNA coding for CR1 17-25 from CR1 cDNA. Expected size of DNA = 1773 bp. (B) Double digest of TOPO®-cloned CR1 17-25 DNA using *Pst*I and *Xba*I. (C) Screening by PCR of CR1 17-25 from vector in *E. coli* Top10 cells. (D) *Sac*I linearisation of plasmid containing the DNA for CR1 17-25. (E) PCR Screening of plasmid containing CR1 17-25-encoding insert from *P. pastoris* cells. (F) SDS-PAGE analysis of small-scale production of CR1 17-25 (18 µl loaded of 20X concentrate from 1ml culture). (G) Western blot of proteins produced in small-scale trials after separation on an SDS-polyacrylamide gel. (H) SDS-PAGE analysis of "mini-scale" production trial of CR1 17-25ER (lanes 1 to 6) and of CR1 17-25KG (lanes 8 to 12) - the amount loaded is smilar to in panel F but only 10X concentrated. (I) CR1 17-25EG "mini-scale" product analysed by SDS-PAGE (same quantity as panel F). (J) Western blot of all variants derived from CR1 17-25KR. Lanes 1 and 2 for two colonies producing CR1 17-25ER, lanes 3 and 4 contain detected proteins produced from colonies of CR1 17KG. Lanes 5 and

6 were from a CR1 17-25EG-producing culture, but only lane 6 show a clear band. Lane 8 was loaded with CR1 20-23, which served as a positive control.

While the mutagensis work was still underway, a one-litre culture of a colony already confirmed to be producing CR1 17-25KR (Figs. 3.9F and G) was grown up. After a five-fold dilution, the supernatant was loaded onto the self-poured SP sepharose column (pI of CR1 17-25KR is 6.1, so buffer used was 20 mM sodium acetate, pH 4.0). The highly glycsosylated protein product, eluted with 1 M NaCl, was detected by Coomassie-stained SDS-PAGE (Fig. 3.9A). Enzymatic deglycosylation was accomplished using Endo $H_f$, Endo H or PNGase. The MWts of these last two are ~30 kDa from that of the target CR1 fragment; Endo $H_f$ on the other hand has an MWt similar to CR1 17-25, *i.e.* 68-70 kD. Under non-reducing conditions, the putative CR1 17-25 band runs just below the Endo$H_f$ band (lane 2 in Fig. 3.8B), but under reducing conditions, the two bands virtually overlap. After de-salting, a second cation-exchange procedure was carried out (Mono S, same buffers and gradient as described for the other constructs) (see Figs. 3.8C and D). Protein-containing fractions from the ion-exchange column (Fig. 3.9D), were buffer-exchanged into PBS for size-exclusion chromatography using Superdex 75 (see Fig. 3.9E). Under reducing conditions some degradation is visible that is not present under non–reducing conditions. This is consistent with proteolysis occurring within modules such that under non-reducing conditions the cleaved portions of the protein are held together by the disulfide bonds, of which they are two per module.

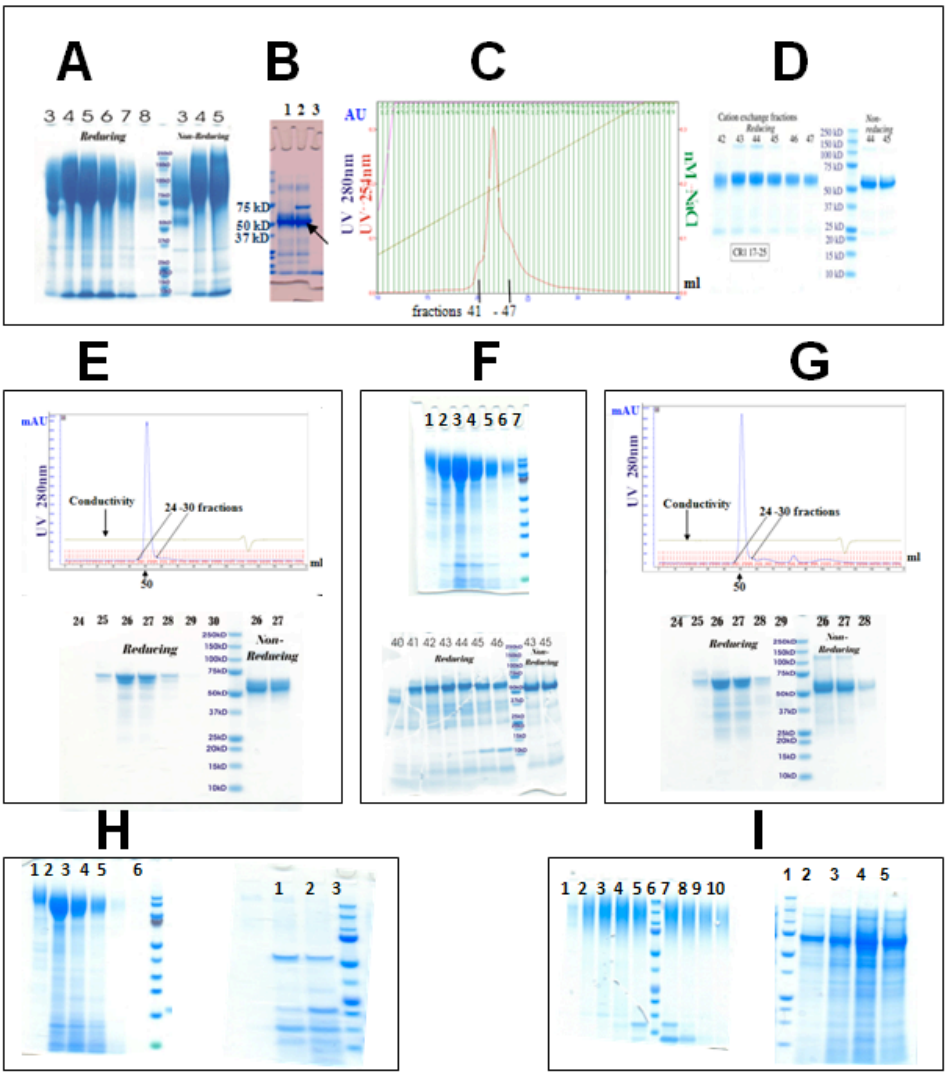



***Figure 3.9 Purification of CR1 17-25 polymorphic forms***

(A) Commassie-stained SDS-polyacrylamide gel following the first cation-exchange chromatographic step. (B) As in (A) but after using EndoH$_f$ / PNGase (lane 1) and EndoH$_f$ (lane 2). (C) Chromatogram following cation-exchange purification (monoS). (D) SDS-PAGE of fractions collected during the chromatographic step shown in panel C for CR1 17-25KR (*i.e.* the form predominant amongst Caucasian populations) (15 µL loaded). (E) Chromatogram for size-exclusion purification step (Superdex 75) of CR1 17-25KR and Coomassie-stained SDS-polyacrylamide gel of the fractions obtained from this chromatographic step (for CR1 17-25KR). (F) SDS-PAGE analysis of cation-exchange chromatographic purification products of CR1 17-25KG. Glycosylated product is shown in the upper gel of the panel whiles the lower gel is deglycocylated using PNGase. (G) Upper, chromatograph and lower, SDS-PAGE obtained from similar purification step to that shown in panel E, except that these are products from the CR1 17-25RG culture. (H) SDS-PAGE gel of CR1 17-25ER, showing (left) the glycosylated product following first cation-exchange step, reducing eventhough this does not mean much because of the glycosylation; also shown is the deglycosylated gel run under non-reducing conditions (right). (I) Similar what is shown in panel H but in this case results are for CR1 17-25EG, i.e. left –

glycsosylated product; right - deglycocylated form of the product purified by cation-exchange chromatography and run under reducing conditions.

One-litre cultures of each of the three Knops blood group variants were also grown up and protein purified as before. The glycosylated fractions obtained from the second cation-exchange purification step for CR1 17-25KG are shown in the upper part of Figure 3.9F; displayed below are the deglycosylated samples. Panel G of Figure 3.9 shows the results of the subsequent size-exclusion chromatographic step for the CR1 17-25KG variant. The results of SDS-PAGE performed on glycosylated and deglycosylated samples, following cation-exchange, are presented in Figure 3.9H for CR1 17-25ER and Fig. 3.9I for CR1 17-25EG, respectively. Purity as seen on these gels were poor but these could possibly be improved when bigger volumes areprepared and purified.

## 3.9 Cloning, production and purification of CR1 15-25

The most ambitious target of the current work was the eleven-module segment of CR1 encompassing both a C3b(/C4b)-binding site (CR1 15-17) and the site of the McC and Sl Knops blood group polymorphisms (in CCP 25). To enhance the chances of attaining this challenging target, amplification of the CR1 15-25KR coding sequence from the "Caucasian variant of CR1"-encoding cDNA (see Table 3.2 for domain boundaries) was carried out in triplicate. Since there is an *Xba*I restriction site within the DNA segment coding for CCP modules 15-16 of this construct, the primers were designed with *Not*I restriction sites instead of *Xba*I ones. However the *Pst*1 restriction enzyme site was maintained (See Table 2.1 for the primer sequences used). Anticipating the potential

difficulties with amplifying such long DNA fragments, the extension time of the PCR cycle was lengthened and the number of cycles increased.

Lanes 2 and 3 of Figure 3.10A demonstrates amplification of products that correspond to the expected size of CR1 15-25KR. After ethanol precipitation of the DNA product, TOPO®-cloning, ligation into pPICZ αB and transformation of *E. coli* Top10 cells were carried out, as before. DNA was extracted from selected colonies and restriction enzyme digestion was used to identify, on a product-size basis, the colony most likely to have the insert. In Figure 3.10B, Lane 1 contains a correctly sized, albeit rather weak, band. Extensive screening and sequencing was then carried out in the hope of confirming that ligation into the pPICZ αB vector had been achieved. Screening with performed using the α-factor forward primers and the AOX-derived reverse primers, always adding, as controls, both an empty pPICZ αB vector (as seen in lane 8 of Fig. 3.10C), and another CR1 construct such as CR1 20-23 (as seen in lanes 2, 3 and 5 of Fig. 3.10C). Lane 1 of the gel in Figure 3.10C showed evidence of amplification, therefore DNA was extracted from this colony for sequencing. To ensure that the entire length of the CR1 15-25 had the appropriate sequence, several primers (*i.e.* in addition to the α-factor forward and AOX-promoter reverse primers) were employed. These included oligonucleotides directed towards the 5' end and the 3' end of the DNA coding for module 21 along with the 5' and 3' primers used previously for amplifying CR1 20-23 (for details of primers see Table 2.1).

Sequence-positive colonies were prepared for glycerol stock while DNA was extracted from one of them for linearization and *P. pastoris* transformation. An aliquot of this extracted DNA (*i.e.* encoding the CR1 15-25KR variant) was subsequently used for

site-directed mutagenesis (as was done in the case of the CR1 17-25 constructs and using the same oligonuceotides) to introduce the base-pair changes needed for the production of the other Knops blood group polymorphic variants (see Materials and Methods) and appendix D2 fro base pair changes. Lanes 1 and 2 of the gel in Figure 3.10D show bands for the K-to-E and R-to-G mutagenesis products.

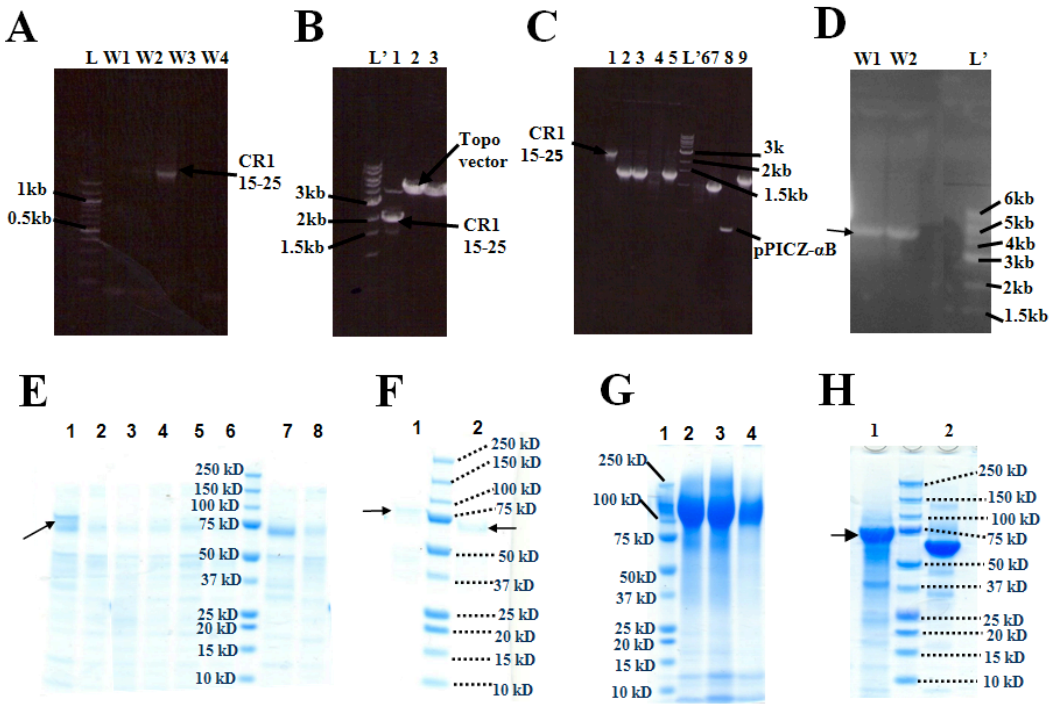



***Figure 3.10 Cloning, production and purification of CR1 15-25 and polymorphic forms***

(A) PCR amplification of DNA encoding CR1 15-25 from cDNA. Expected DNA size was 2.133 kilo-bp. (B) Double digest of TOPO®-cloned CR1 15-25-encoding DNA using *Pst*I and *Not*I. (C) PCR-based screening of plasmids containing CR1 15-25 from *E. coli* Top10 cells. (D) "QuickChange" site-directed mutagenesis of CR1 15-25KR DNA to produce DNA coding for CR1 15-25ER (W1) and CR1 15-25KG (W2). (E) SDS PAGE analysis of small-scale production trial of CR1 15-25KR (18µL loaded of 20X concentration from 1-ml cell culture supernatants). (F) SDS-PAGE analysis of crude deglycosylated material in supernatant from a larger-scale culture (18 µL loaded from a one-litre fermenter). (G) Coomassie-stained SDS-polyacrylamide gel of fractions collected from the first cation-exchange chromatographic step; highly glycosylated proteins were formed. (H) SDS PAGE of de-glycosylated form of the fraction in panel G.

Specifically in the case of this eleven-module construct, an attempt was made to boost copy numbers by re-streaking colonies that grew on 100 mg/ml Zeocin™-YPD onto 300 mg/ml-Zeocin™ plates before selection for further processing. Small-scale protein production trials of CR1 15-25KR yielded only faint bands (Fig. 3.10E). Nonetheless, colony 1 was picked for scale-up. Analysis by SDS-PAGE of the crude deglycosylated supernatant of a one-litre culture produced faint bands that were candidates for CR1 15-25KR; lanes 1 and 2 (Fig. 3.10F) contain appropriately sized protein bands, under reducing and non-reducing conditions, respectively.

In trial purifications (on a one-litre scale) of CR1 15-25KR, the standard initial cation-exchange chromatography step was carried out and glycoprotein-containing fractions were identified by SDS-PAGE (Fig. 3.10G). This glycosylated protein produced a smeary band running higher than might be expected, but deglycosylation yielded sharper bands of the expected masses - about 75 kD under reducing conditions, and reassuringly a single band also under non-reducing conditions (running just below the 75-kD marker) (Fig. 3.10H). Panels E, F, G and H of Figure 3.10, which show results for CR1 15-25KR, follow the sequence of steps through which all four variants were taken before the final purification step (usually, size-exclusion chromatography). Similar results for the other three variants (from one-litre fermentations) were obtained but not shown. In order to increase the yields of protein produced (for all variants of CR1 25-25) with the aim of carrying out multiple experiments, larger-scale production attempts were made as described in the subsequent sections.

### 3.9.1 Purification of CR1 15-25KR (Caucasian)

Based on its calculated pI of 6.86, CR1 15-25KR was amenable to cation-exchange chromatography in sodium acetate buffer, pH 4, and an XK 16/20 column packed with 25 ml of SP-Sepharose resin (see Materials and Methods for details). The chromatogram (Fig. 3.11A) for this first purification step reveals poorly bound contaminating proteins eluting off the SP-Sepharose before the protein of interest. Because the target protein was glycsosylated all lanes in the SDS polyacrylamide gel contain "smeary" or poorly defined bands and species migrate more slowly (appear higher on the gel) than would be expected from the molecular weights calculated for the protein portion of the molecule (Figure 3.11B). Deglycosylation using Endo $H_f$ was performed immediately after the appropriate glycoprotein-containing fractions from cation-exchnage chromatography had been pooled. It will be appreciated that the molecular weights of this construct (*i.e.* CR1 15-25) and that of Endo $H_f$ are similar (78 Kd and 70 Kd for CR115-25 and Endo $H_f$, respectively) and difficult to resolve on size-exclusion resin. Therefore, to the standard purification protocol, was added a step that involved passing the deglycosylated product over a mannose-binding resin, thus getting rid of the Endo $H_f$. Subsequently, size exclusion purification was performed.

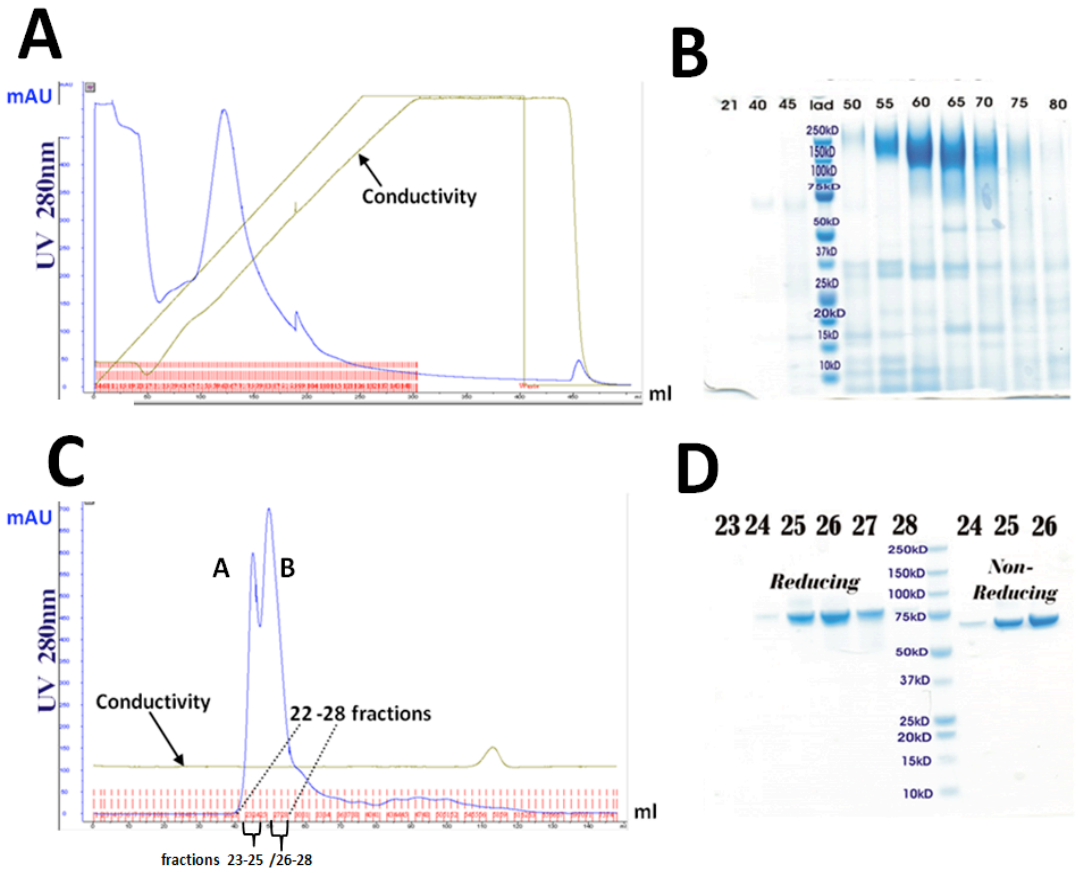



***Figure 3.11 Purification of CR1 15-25KR***

(A) Chromatogram for first cation-exchange purification step of CR1 15-25KR (details in Materials and Methods). (B) Commassie-stained SDS-polyacrylamide gel of glycoprotein-containing fractions from the ion-exchange purification step. (C) Gel-filtration chromatogram of CR1 15-25KR (run over HiLoad Superdex75); peaks A and B in the chromatogram correspond to two species that become evident after deglycosylation and removal of EndoH$_f$. Peak A corresponds to fractions 23 to 25, while peak B corresponds to fraction 26 to 28 as indicated in panel D; (D) The SDS-polyacrylamide gel loaded with frations (as shown) from the size-exclusion chromatography and electrophoresed.

The final "polishing" step (on a size-exclusion resin) was similar to that used previously (Figure 3.11C). There appear to be two overlapping peaks corresponding to species (A and B) of different sizes, perhaps reflecting incomplete deglycosylation. Indeeds there could be O-linked glycosylation present (that is not susceptible to deglycosylation by Endo H$_f$). Therefore fractions corresponding to peak A were not pooled with those of

peak B and the two species were individually tested for biological activities or binding abilities. The corresponding gel is shown in Figure 3.11D, with the proteins running at the expected molecular weight of 78.1 kD, irrespective of whether protein fractions corresponded to peaks A or B. This implies that the two species might be different in conformation or that they might correspond to dimers and monomers. Note that in the polyacrylamide gels, single bands are obtained under both reducing and non-reducing conditions with the non-reduced protein running a little faster, consistent with what would be expected of proteins containing disulfides.

### 3.9.2 Purification of CR1 15-25ER

Expression and production of CR1 15-25ER was performed using similar steps to those chosen to purify the CR1 15-25KR variant described above. CR1 15-25ER obviously has a Glu residue instead of Lys at amino acid position 1590 but still has an Arg at position 1601 (theoretical pI is 6.62). Figure 3.12A shows the cation-exchange chromatogram of CR1 15-25ER performed on the same column as was used for CR1 15-25ER (see Materials and Methods for details). Though glycsosylated (and hence running higher than expected molecular weight of the protein component), an exceptionally clean protein is evident in Figure 3.12B. This may have been thanks to modification of the protocol to include a 2.5-times column-volume low-salt buffer wash prior to the start of the salt gradient that helped to eliminate weakly binding contaminants.

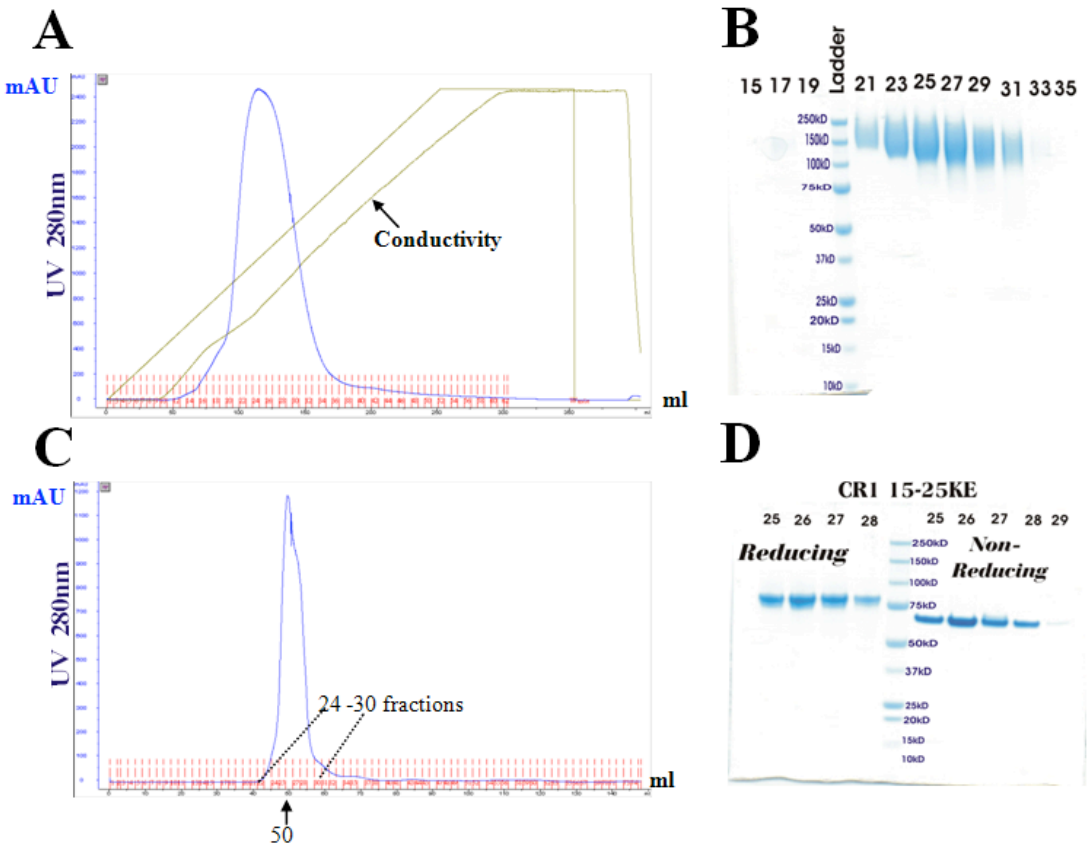



***Figure 3.12 Purification of CR1 15-25ER***

(A) Chromatogram for cation-exchange purification of CR1 15-25ER (Details in Materials and Methods). (B) Commassie-stained SDS-polyacrylamide gel of eletrophoresed glycoprotein-containing fractions from this ion-exchange purification step. (C) Size-exclusion chromatography of CR1 15-25ER (on HiLoad Superdex-75) after deglycosylation and removal of Endo $H_f$. (D) Outcome of SDS-PAGE peformed on protein-containing fragments.

After deglycosylation with Endo $H_f$ as described previously, and passing the product through mannose-affinity beads to remove the fusion protein, size-exclusion chromatography yielded the results shown in Figures 3.12C and D. The protein ran at the expected molecular weight of ~78.1 Kd and was a single band under both reducing and non-reducing conditions (Figure 3.12D).

### 3.9.3 CR1 15-25KG Purification

This protein has a Lys residue at amino acid 1590 but a Gly replaces the Arg at position number 1601. The expression, production and purification procedures were similar to those used for the other variants as described above. Hence cation-exchange chromatography of CR1 15-25KG yielded the results shown in Figure 3.13A. As expected for a construct with N-linked glycosylated sites, the fractions as analysed by SDS-PAGE (Fig. 3.13B) contain glycosylated proteins.

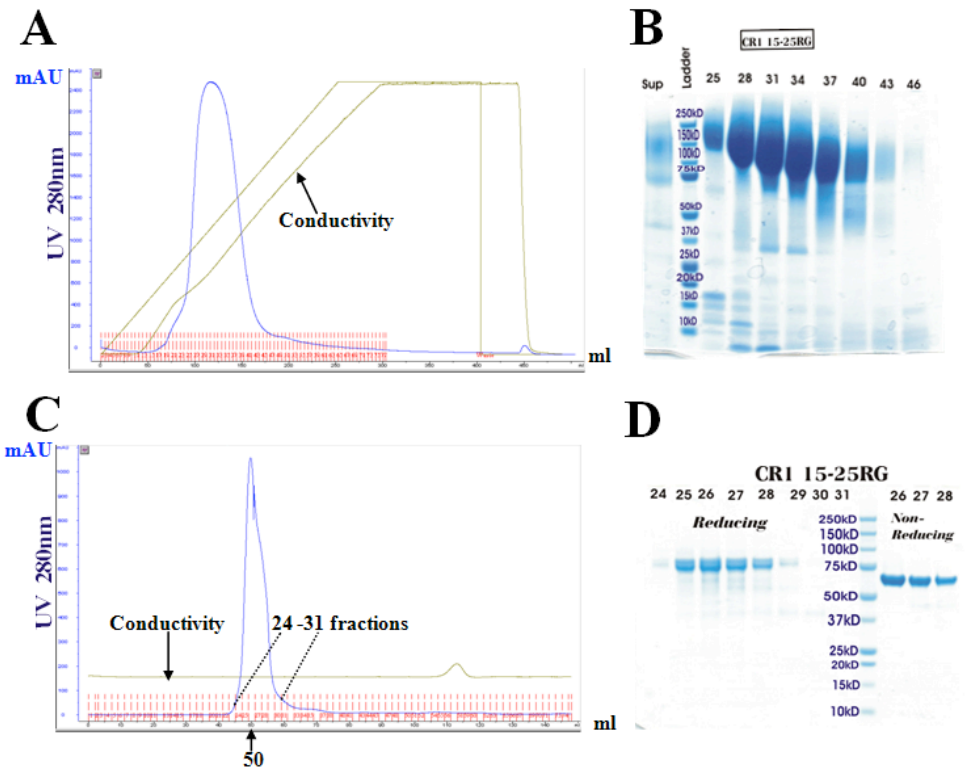



***Figure 3.13 Purification of CR1 15-25KG***

(A) Chromatogram of cation-exchange purification of CR1 15-25KG. (B) Commassie-stained SDS-polyacrylamide gel of glycoprotein-containing fractions from ion-exchange purification step. (C) Size-exclusion chromatogram of CR1 15-25KG following deglycosylation and removal of Endo $H_f$. (D) The corresponding SDS-PAGE. Details of all methods are provided in Chapter 2.

The chromatogram of the final size-exclusion step performed on the CR1 15-25KG sample and the corresponding SDS-PAGE analysis are shown in Figures 3.13 C and D.

The purified protein runs on the gel at the expected size (78.1 kD), and there was an appropriate difference between the migration rates (upon SDS-PAGE) of bands under reducing and non-reducing conditions. Under reducing condition, however, there appeared to be two bands on the polyacrylamide gel close to each other. These appear to collapse into a single band in non-reducing conditions implying that a proportion of the protein molecules had been proteolytically "nicked", presumably within one of the modules, but the separate polypeptides remains held together by disulphide bonds. Alternatively, and less likely, it is possible that the difference in the length of the potential cloning artefact (EA or EAEAA, as explained in the Introduction) could result in two similar species that are resolvable by SDS-PAGE under certain conditions and not others.

### 3.9.4 CR1 15-25KG Purification

The CR1 15-25EKG variant was cloned, produced and purified in essentially the same way as the other variants. Cation-exchange chromatography was used for inital harvesting and again for a first purification step as summarised in Figures 3.14A and B. As before, extensively washing the cation-exchange resin prior to starting an elution gradient (0 – 1 M NaCl) improved the efficacy of this procedure. Subsequently, deglycosylation and a size-exclusion chromatographic "polishing" step (results shown in Fig. 3.14C and D) afforded samples that upon SDS-PAGE yielded bands migrating with the expected masses of ~78 kD. Careful examination of the polyacrylamide gel reveals the presence of two bands running close together under reducing conditions, which become one band under non-reducing conditions. As discussed above, this presumably

arose from limited proteolysis within a module (near to one or other end of the protein) such that the two cleavage products are tethered via disulfide bond(s) and would therefore run as single species unless reduced.

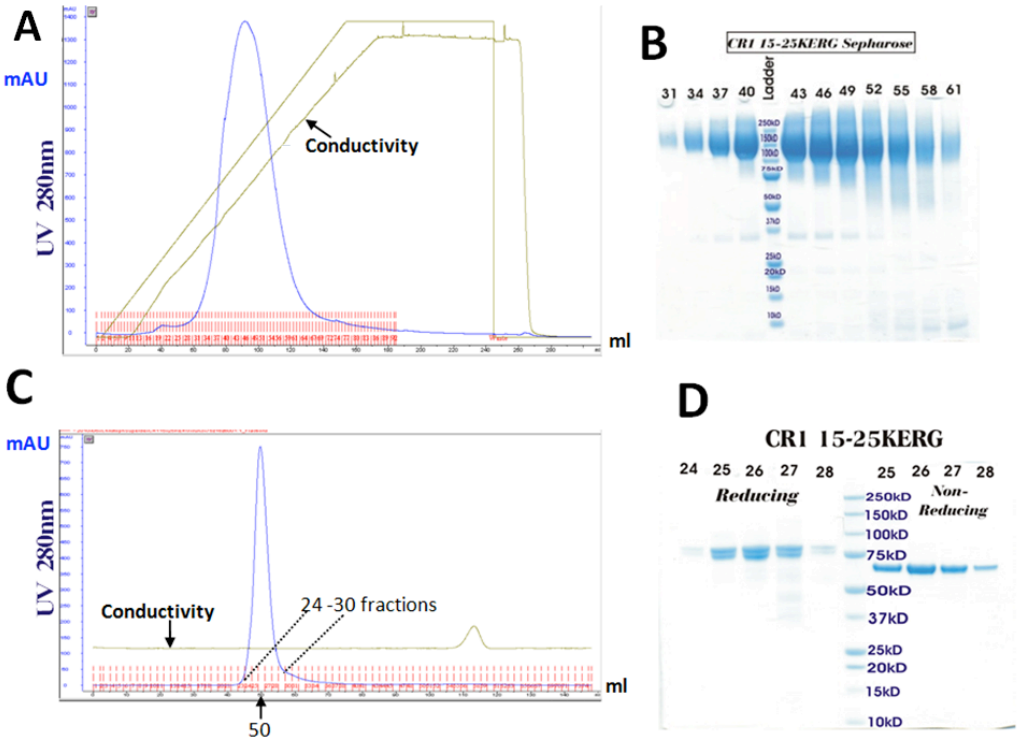



***Figure 3.14 Purification of CR1 15-25EG***

(A) Chromatogram resulting from purification, on cation-exchange resin, of CR1 15-25EG. (B) Commassie-stained SDS-polyacrylamide gel following electrophoresis of glycoprotein-containing fractions from the aforementioned ion-exchange purification step. (C) Chromatogram for CR1 15-25EG eluted from a HiLoad Superdex-75 size-exclusion column. (D) Results of SDS-PAGE performed on fractions from size-exclusion chromatography.

## 3.10 CR1 24-25 cloning, production and purification

The recombainantly produced CR1 24-25 fragment (for domain boundaries see Table 3.2) contains both the McC and Sl Knops blood group variants. Primers (Table 2.1) were

designed to amplify the relevant DNA segments from the DNA sequences coding for the four CR1 15-25 variants described above. Thus (in Fig. 3.16A) lanes 1 to 4 contain the correctly amplified DNA segments coding for the four polymorphic forms of CR1 24-25. After going through the processes of PCR, subcloning into the TOPO® plasmid vector (pCR®4Blunt-TOPO®), double digestion and ligation into pPICZ αB (as before, see Materials and Methods), successful transformed *E. coli* amplification hosts that screened positive were sequenced. Plasmids with the correct sequences were amplified by a "maxi-prep", *Sac*I linearised and used to transform *P. pastoris*, all as before.

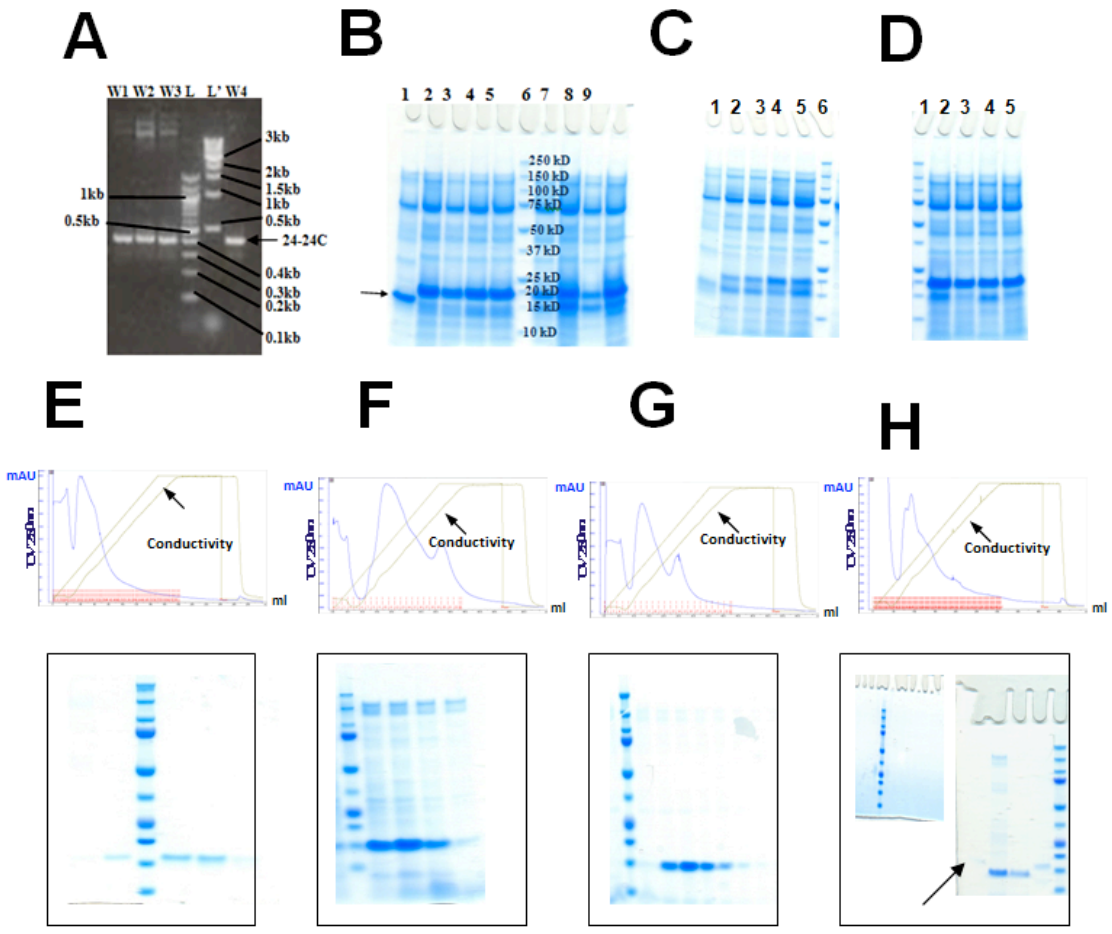



***Figure 3.15 Cloning, production and purification of CR1 24-25 and polymorphic forms***

(A) Products of amplification by PCR of DNA segments encoding four variants of CR1 24-25 from the DNA coding for CR1 15-25 (see text). The expected size of the DNA insert is 411 bp. (B) The result of SDS-PAGE performed on supernatants from small-scale production trials of CR1 24-25ER (lanes 1-5) and CR1 24-25KG (lanes 6-9); lanes 1 and 2 contain protein samples from the same culture but were run under non-reducing and reducing conditions, respectively. (C) SDS-PAGE performed following "mini-scale" trials for production of CR1 24-25EG (18 µL of 20x concentrate of supernatant from a 1-ml cell culture). Lanes 1 to 5 represent expressions from the five selected colonies screened. (D) Similar outcome to that of panel C, except that lanes 2 to 5 show expressions from selected colonies producing CR1 17-25ER Amount of material loaded is same as in panel C. (E) Cation-exchange chromatogram (upper sub-panel)) and selected fractions of CR1 24-25KR, ran on SDS-PAGE (lower sub-panel) are shown Similar cation-exchange (SP-sepharose) chromatograms and their corresponding polacylamide gels were obtained for CR1 24-25KG (panel F), CR1 24-25KG (panel G), and CR1 24-25ER(panel H) (left is non-concentrated gel while arrow in right points to the pooled fraction that has been concentrated). Volumes loaded on gels were the same in each case. (Note that an XK column packed with 25 ml of resin was used for panels (E) and (H), while the one used for (F) and (G) was packed with only 7 ml of SP-Sepharose.

Based on the reasonably good yields of target proteins revealed by SDS-PAGE following "mini-scale" production trials (Fig. 3.16), a one-litre culture for each variant was set up. As per normal, the resultant supernatants after harvesting were each diluted 1-in-5 and loaded onto self-packed (with SP-sepharose) XK column. Panels of Figure 3.16 (E for CR1 24-25KR, F for CR1 24-25KG, G for CR1 24-25 EG and H for CR1 24-25 ER) show the outcome of these protein production runs. Out of the four variants, production of CR1 24-25ER (Fig. 3.16H) was the poorest although evidence of a low level of producton is discernible as a faint band indicated by the arrow.

### 3.11 Attempted Production of domains of PfEMP1

The protocols employed for cloning, protein production and purification for the two DBLα domains depicted in Figure 3.17A were similar to those used for the CR1 constructs in Figure 3.1. However, since the DBLα domains are "AT rich" and could pose a challenge for the *P. pastoris* expression system, a codon-optimized gene was purchased from GeneArt. The gene was subcloned into the TOPO® plasmid and transformed into chemically competent *E.coli* TOP10 cells. After "miniprep", double digest, ligation into pPic ZαB and amplification (in *E.coli* cells), linearised plasmid was used to transform P. pastoris cells. All of these steps were carried out in the same way as was described earlier for the other recombinant protein targets.

Cell culture, harvesting of protein and purification were also carried out as before. Therefore after harvesting (and addition of EDTA and PMSF to inhibit proteases) ion-exchange chromatography and size-exclusion chromatography were employed. In this

case, two differently sized peaks were obtained ("small" and "big" peaks, chromatogram not shown).The fractions corresponding to small and big peaks were independently pooled, and each pool was subjected to a further round of cation-exchange chromatography using the monoS. While two promising chromatograms were recorded, (see Fig. 3.17 B – up and C-up), subsequent SDS-PAGE (Fig. 3.17 A-down and B-down) suggested both batches of the protein had degraded, with potentially proteolytic fragments appearing as bands under both reducing and non-reducing conditions. A possible explanation for this observation derives from inspection of the sequence (see Appendix F) which suggests the odd number of cysteine residues present might result in disulfide shuffling and the potential non-formation of native disulfides. Fractions of the big peak were, nonetheless, used in the early SPR experiments involving CR1, C3b and DBLα, where it was described as DBLα(P) (Section 5.6.2)

After the disappointing results obtained from attempts at DBLα expression in the *P. pastoris* system (and early failures to observe physiologically significant binding to either DBLα(P) or DBLα(M)) an attempt was made to produce the three-domain construct (DBLα-CIDR-DBLγ) using domain boundaries that result in a recombinant product that has an even number of cysteine residues and hence a lower chance of free thiols. Another codon-optimized gene, which was then ligated into pPic ZαB. was purchased from GeneArt and transformed into *E.coli* cells. After maxipreping to extract the amplified DNA, linearization was performed and the product prepared for transformation into *P. pastoris* as previously described (Section 2.2.14).

Figure 3.17D shows that a DNA band of the expected size (6.783 kb = 3.183kb of the DBLα-CIDR-DBLγ gene inserted + 3.600 kb of the pPicZ-αB vector) was obtained

from the linearised plasmid. Subsequent yields of recombinant protein were, however, very low and bands obtained by SDS-PAGE of appropriate molecular weight could only be identified using Western blot. Results obtained using DBL antibody and antihistidine antibody are shown in panel D of Figure 3.17D (lower sub-panel).

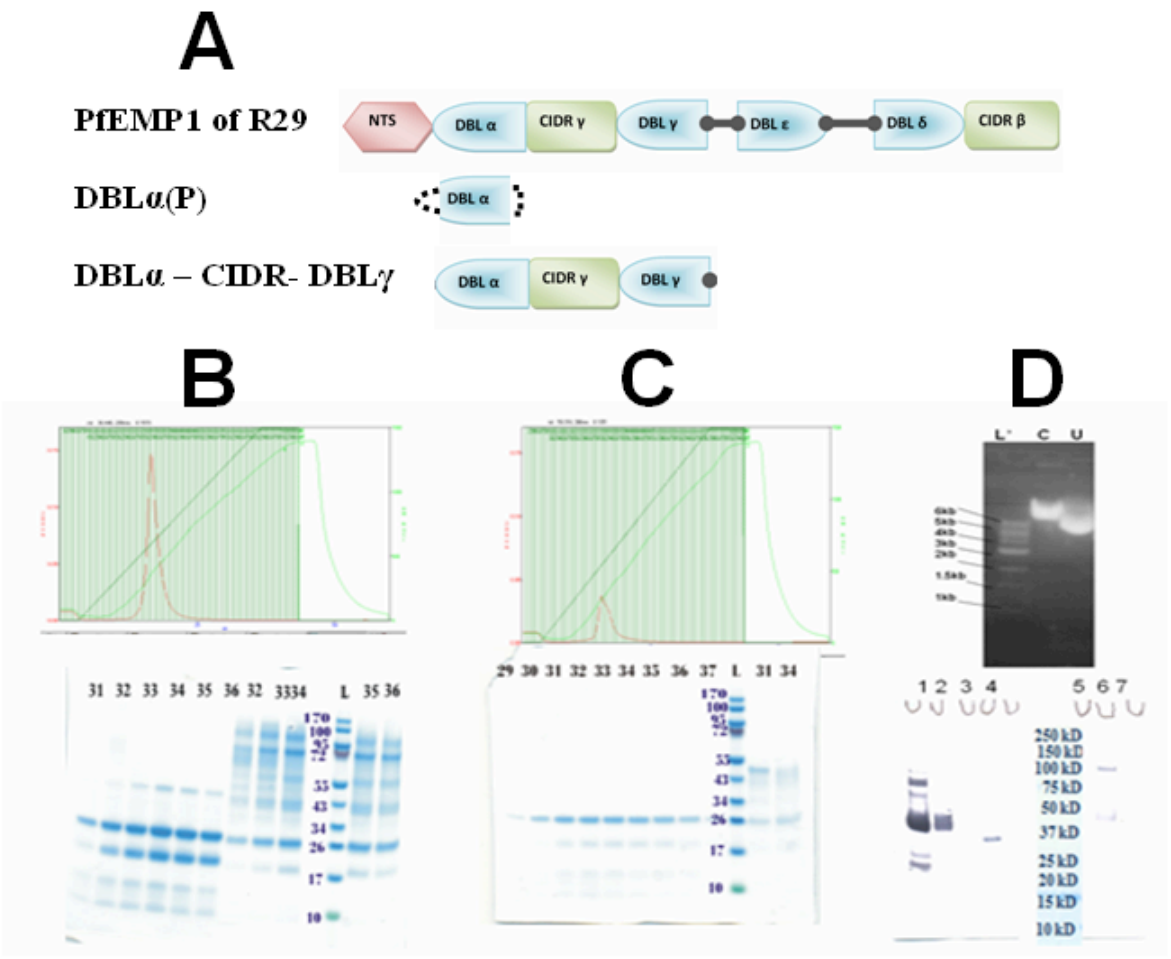



***Figure 3.17 Attempted production of PfEMP1 domains***.

(A) Domains of PfEMP1 used in this study. NTS is N-terminal segment, DBL means Duffy-Like binding domain and CIDR is Cysteine-rich interdomain region. (M) and (P) are letters designated to the different DBLs with respect to those who produced them. DBLα(M) was produced by Dr. Matt Higgins and hence the letter (M) whiles DBLα(P) was produced by Patience Tetteh-Quarcoo, hence the letter (P). Cation-exchange chromatograph and SDS-PAGE analysis of (B) the "big" peak (named DBLa Pat (P)). First six lanes from the left are reducing whiles the rest

are non reducing and (C) the "small" peak obtained from a previous attempt at purification involving cation exchange and size-exclusion chromatographic steps (D) Agarose gel (above) of linearised DNA encoding DBLα-CIDR-DBLγ in pPicZα B (C-Cut and U-Uncut) and (below) Western blot of the protein expressed (to identif the C-terminus, an anti-His tag antibody was used in lanes 1-4 while a DBLα-recognising antibody was used (lanes 5-7), to identify the N-terminus).

# CHAPTER FOUR
# BIOPHYSICAL STUDIES

# CHAPTER 4 BIOPHYSICAL STUDIES

## 4.1 Overview

This chapter covers the results obtained from extensive biophysical characterisation of the recombinant proteins produced in this study. Efforts - utilising a range of structural techniques including nuclear magnetic spectroscopy (NMR), dynamic light scattering (DLS), analytical ultracentrifugation (AUC) and small-angle X-ray scattering (SAXS) - were aimed at addressing two main questions:

(i) Does the long linking sequence unique to the boundary between LHR-C and LHR-D induce a bend in the CR1 molecule that causes these two regions to fold back against one another (such that sequence variations in LHR-D could modulate the functional activities resident in LHR-C)?

(ii) Do sequence variations in the 24$^{th}$ and 25$^{th}$ CCPs of CR1, arising from the Knops blood group McC and Sl polymorphisms, correlate with differences, either in gross molecular architecture or in self-association tendencies.

Note that where there was a requirement for isotopic labelling during protein production, the results of this procedure are also included in the relevant sections of this chapter.

## 4.2 NMR-based investigations of CR1-21, CR1 21-22 and CR1 20-23

The approach employed here was to use chemical shift-perturbation mapping to interrogate the extent of interactions between neighbouring CCP modules in the LHR-C/LHR-D boundary region. Each CCP module resembles a prolate ellipsoid in overall

shape with one long axis and two shorter ones. An end-to-end interaction (with long axes more or less aligned) between modules generally results in an extended conformation involving only a small intermodular interface. On the other hand, a side-by-side interaction between modules results in a more compact arrangement with more extensive intermodular contacts. In the case of a side-by-side interaction then, one would expect major differences in the chemical shifts of a module (*e.g.* CR1 21) studied on its own *versus* the chemical shifts of the same module but studied in its double module context (*e.g.* CR1 21-22). As described in the Introduction, while in CR1 28 of the intermodular tethers (three or four residues) are too short to allow a side-by-side interaction (due to the steric bulk of the modules themselves), this is not the case for the linker between CCPs 21 and 22.

Comparison of $^{1}H,^{15}N$-HSQC spectra for the variously-sized fragments was used as a relatively rapid and straightforward means of assessing the context-dependency of chemical shifts of backbone amide protons and nitrogen nuclei in a given CCP module. Some backbone assignment of nuclei was carried out to enhance the interperetation of the data. Isotopic labelling of proteins was needed for these studies, and therefore most of the results sections below show protein gels obtained for the purified isotopically enriched protein fragments. Labelling included $^{15}N$-enrichment only (for CR1 21 and CR1 21-22), dual $^{2}H,^{15}N$ labelling ( for CR1 20-23) and production of a double-labelled ($^{15}N,^{13}C$) version of CR1 21-22.

### 4.2.1 Production and purification of isotopic labelled CR1 21

The production of isotopically ($^{15}N$) labelled production CR1 21 in a fermentor was performed similarly to non-labelling fermentor runs, with a few adjustments to the protocol as described in Chapter 2 (Section 2.3.6). After the supernatant from a one-litre fermentation of a $^{15}N$-labelled growth had been harvested, purification commenced in a similar fashion to that described for the non-labelled CR1 21 protein (see Sections 3.2 and 2.3.7).

The chromatogram shown in Figure 4.1A was obtained during the purification, of CR1 21 on the MonoS cation-exchange column. SDS-PAGE of protein-containing fractions and Coomassie staining yielded the gel shown in Figure 4.1B. As was observed for the the non-labelled preparation (Section 3.2), under reducing condition the major band, presumably corresponding to CR1 21, runs close to the 10-kD size marker, while fractions also contain an impurity of appriximately 70 kD. This was removed by size-exclusion chromatography resulting in the chromatogram shown in Figure 4.1C.

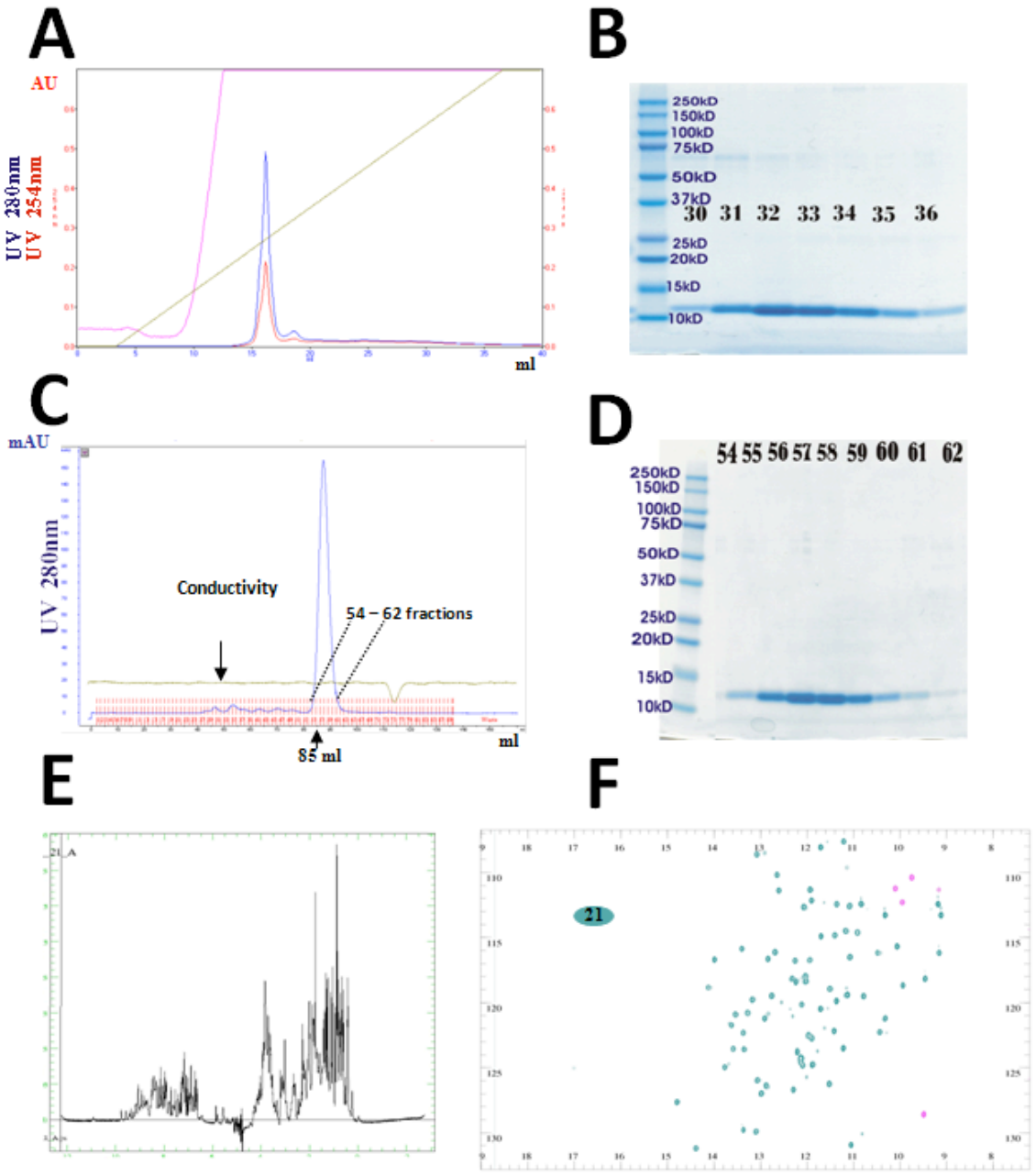



***Figure 4.1 Preparation and NMR spectra for $^{15}$N CR1 21***

(A) Chromatogram obtained following the cation-exchange purification of $^{15}$N-CR1 21 (eluting with a gradient of 20 mM sodium acetate buffer, pH 4 and the same buffer with 1 M NaCl). (B) Commassie-stained gel following SDS-PAGE of the protein-containing fractions obtained from the cation-exchange chromatography. (C) Chromatogram following size-exclusion purification of $^{15}$N-CR1 21. (D) Commassie-stained gel following SDS-PAGE of the protein-containing fractions obtained from size-exclusion chromatography. (E) One-dimensional $^{1}$H-spectrum of 233 μM $^{15}$N-CR1 21. (F) $^{15}$N,$^{1}$H HSQC spectrum of the same $^{15}$N-labelled protein construct as in panel E.

# CHAPTER 4 BIOPHYSICAL STUDIES

Unless specified otherwise, NMR experiments were run on samples containing 20 mM potassium phosphate buffer, pH 6.6 (with no NaCl). Therefore, since the $^{15}$N-CR1 21 was produced specifically with NMR experiments in mind, the buffer used for size-exclusion chromatography was 20 mM potassium phosphate, pH 6.6, and 500 mM NaCl. In this way desalting, but no buffer-exchange, was necessary prior to transfer to the NMR tube. The Coomassie-stained gel, obtained following SDS-PAGE of the fractions from size-exlusion chromatography of $^{15}$N-CR1 21, is shown in Figure 4.1D.

The overall yield of labelled protein from this one-litre fermentation was a useful ~7 mg. To a purified sample of 500 μL of $^{15}$N-CR1 21was added 50 μL of $D_2O$ to act as a "frequency lock", and then the resultant 233 μM protein sample was transferred into a 5-mm NMR tube for data collection in a high-field NMR spectrometer. A globular and properly folded, non-degraded and stable protein was indicated by the quality of the $^{1}$H spectrum (Fig. 4.1E) and $^{15}$N,$^{1}$H-HSQC spectrum (Fig. 4.1F**)** of $^{15}$N-CR1 21. For example, the sharpness and dispersion of peaks in Figure 4.1E and the high ratio of up-field shifted methyl peaks (at around 0-0.5 ppm) to those at a random coil position (*i.e.* giving signals at 1-2 ppm) are consistent with formation of a compact hydrophobic core containing aromatic side chains. Additionally, the HSQC spectrum in Figure 4.1F has roughly the number of cross peaks (~60, i.e. one for each amide (NH) group) expected for a pure sample of CR1 21, and the peaks are all of approximately uniform intensity and are well dispersed.

### 4.2.2 Production and purification of isotopically labelled CR1 21-22

An identical protocol to that used for purification of CR1 21 was also employed for the cation-exchange purification of $^{15}$N-CR1 22-22 from a one-litre fermentation. The corresponding Coomassie-stained gel, following SDS-PAGE under reducing conditions, is shown in Figure 4.2A. The chromatogram and protein gel for the subsequent size-exclusion chromatography step are displayed in Figures 4.2B and C. The total yield of labelled protein from a one-litre fermentation was 18 mg.

Following addition to 10 % (v/v) of $D_2O$, a total volume of 550 μL of $^{15}$N-CR1 21-22 solution (at the relatively high concentration for an NMR sample of 2.6 mM) was transferred to a 5-mm NMR tube**.** After obtaining a 1D $^{1}$H spectrum for this sample (Fig. 4.2D), a $^{15}$N,$^{1}$H-HSQC spectrum was also recorded (see Fig. 4.2E). It was judged that these spectra were consistent with a well-folded, stable and non-degraded protein as had been the case for CR1 21. The central portion of the HSQC spectrum is quite crowded, which is to be expected given that this sample contains some 125 amides, but overall dispersion is good and the number of peaks matches the expected number of amides.

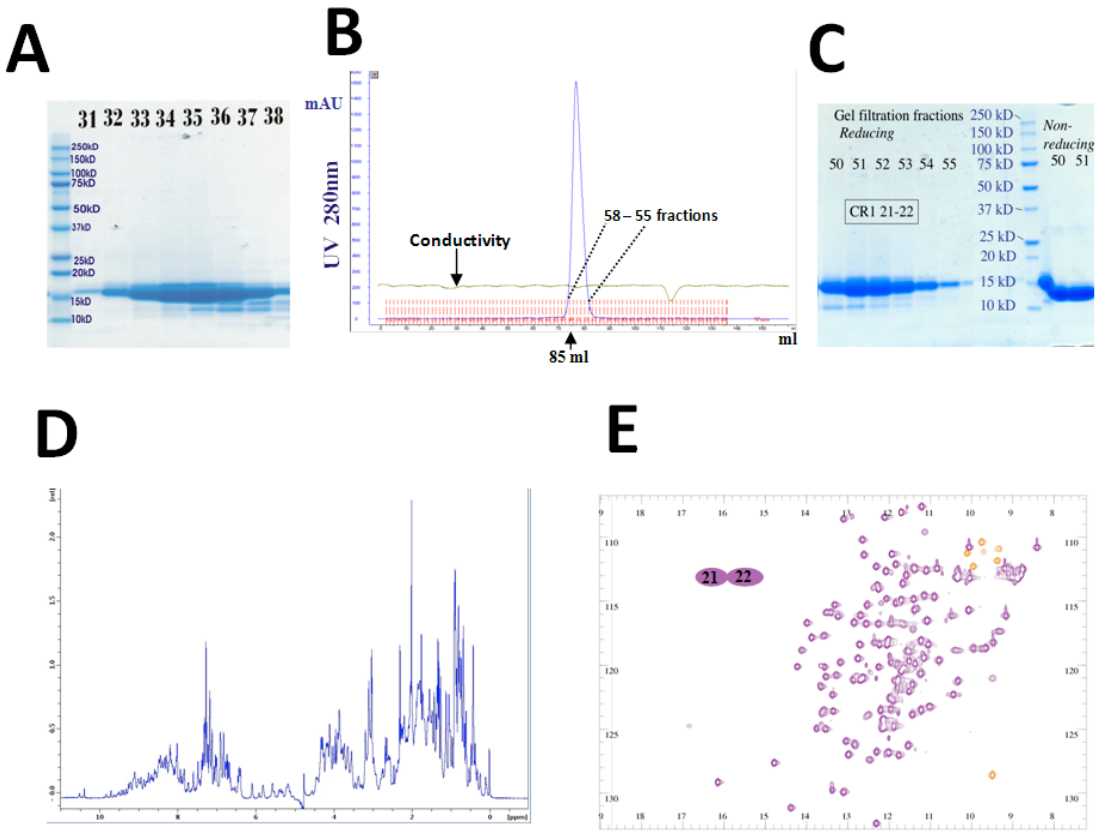



***Figure 4.2 Preparation of, and NMR spectra for, $^{15}$N-CR1 21-22***

(A) Commassie-stained gel following SDS-PAGE of protein-containing fraction obtained from cation-exchange purification (not shown) of $^{15}$N-CR1 21-22. (B) Chromatogram obtained from the subsequent size-exclusion purification. (C) SDS-polyacrylamide gel following electrophoresis of protein-containing fractions indicated in panel B. (D) 1D $^{1}$H spectrum of $^{15}$N-CR1 21-22. (E) $^{15}$N,$^{1}$H-HSQC spectrum of $^{15}$N-CR1 21-22.

### 4.2.3 Comparison of $^{15}$N,$^{1}$H- HSQC spectra of $^{15}$N-CR1 21 and $^{15}$N-CR1 21-22

An overlay of the $^{15}$N,$^{1}$H-HSQC spectra of CR1 21 and CR1 21-22 is displayed in Figure 4.3. With a few exceptions, the cross peaks observed in the spectrum of CR1 21 corresponded to cross peaks at near-identical positions in in the spectrum of CR1 21-22. The latter must correspond to CR1 21-derived cross peaks in the spectrum of CR1 21-22 that have the same chemical shifts in spectra of both the single-module and double-module CR1 fragments. This conservation of chemical shifts in the two contexts suggests

that the attachment of CCP 22 does not affect the magnetic environments of any substantial surface regions of CCP 21. Such a result is not consistent with a conformation of the double module-construct in which the two modules participate in extensive side-by-side interactions; rather they suggest an end-to-end arrangement with a relatively small intermodular interface. Note, however, that these results do not preclude a tilted arrangement of the modules in which the linker forms a small hydrophobic pocket that "glues" the two in a particular orientation (See Discussion in Chapter 6).

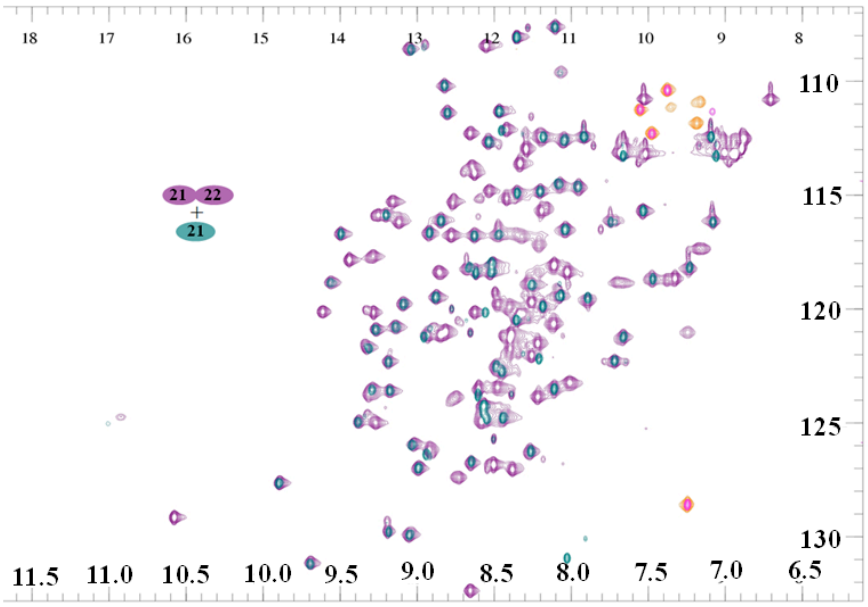



***Figure 4.3 Overlay of HSQC spectra of $^{15}$N-CR1 21 and $^{15}$N-CR1 21-22***

Positive cross peaks in the spectra are colour-coded as indicated by the cartoons *i.e.* the CR1 21 spectrum is cyan while the CR1 21-22 spectrum is purple. Negative peaks for CR1 21 are red while those of CR1 21-22 are yellow.

A question mark remained over the biological relevance of this conclusion, however, since additional CCP modules before module 21 and after module 22 could, in theory, interact with one another to stabilise a fully bent back 21-22 structure not seen in the isolated double module. To check whether the architecture of 21-22 changes in the context of a quadruple-module construct, a $^{15}$N-labelled sample of CR1 20-23 was prepared.

### 4.2.4 Production and purification of $^{15}N,^{2}H$-CR1 20 -23

The quadruple-module construct has a molecular weight of 28.76 kD, and if it were extended would be expected to tumble slowly, producing broad lines and consequently low signal intensities. To circumvent this anticipated outcome, a dual deuterated ($^{2}H$) and $^{15}N$-labelled sample was produced, as described in Section 2.3.6.5. This was subsequently purified *via* a similar method to the one used for the non-labelled sample (see Sections 3.4 and 2.3.7). The resultant partially deuterated sample - in which the backbone amides inevitable become protonated by chemical exchange with aqueous ($^{1}H_2O$) solvent - was expected to yield improved spectra since the remaining protons should relax more slowly due to having fewer potential relaxation partners.

The chromatogram obtained from cation-exchange chromatography and the Coomassie-stained gel following SDS-PAGE of protein-containing fragments, are shown in Figures 4.4A and B. After size-exclusion chromatography (Figs. 4.4C and D), the resultant purified protein preparations were pooled and buffer-exchanged for NMR data collection. An estimated 8.6 mg of protein was obtained from the one–litre cell culture. A 300-μM NMR sample of 550 μl containing 10 % (v/v) $D_2O$, was transferred to the NMR tube and the $^{1}H$ and $^{15}N,^{1}H$ HSQC spectra of $^{15}N,^{2}H$ CR1 20-23 were recorded (Fig. 4.4E and Fig. 4.4F, respectively). These promising results represented the first NMR spectra collected (so far as we know) for a quadruple CCP module. Both spectra were of unexpectedly high quality for a relatively high-molecular weight and potentially extended protein of ~28 kD, despite the crowded nature of the middle region of the HSQC spectrum (Figure 4.4F).

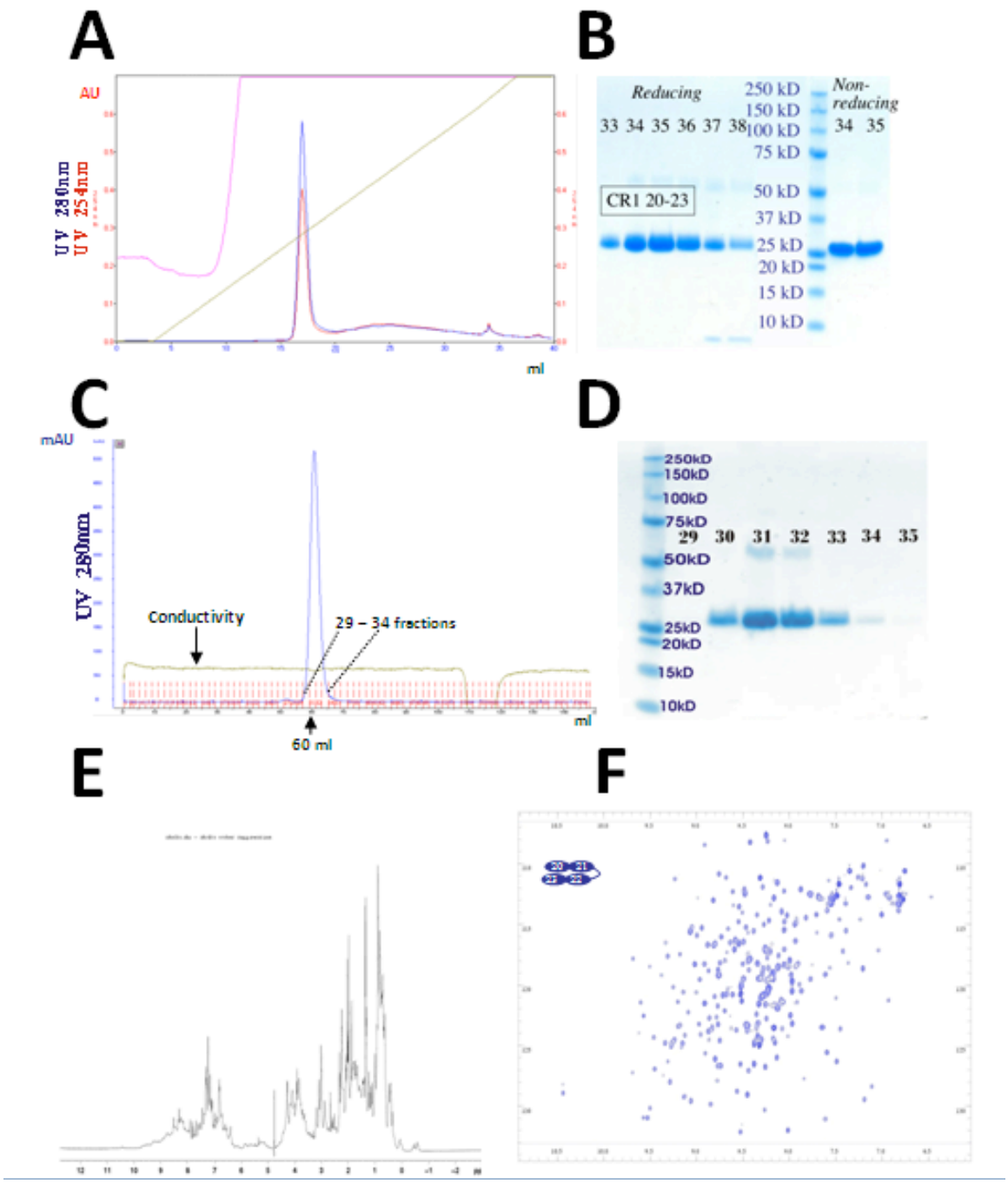



***Figure 4.4 Preparation of, and NMR spectra for, $^{15}N,^{2}H$-CR1 20-23***

(A) Chromatogram obtained for the cation-exchange purification of $^{15}N,^{2}H$-CR1 20-23 (elution with a gradient from 20 mM sodium acetate buffer, pH 4 to the same buffer but with 1 M NaCl). (B) Commassie-stained gel following SDS-PAGE of protein-containing fragments from the cation-exchange chromatography. (C) Chromatogram for size-exclusion purification step (using Superdex-75; 20 mM potassium phosphate, pH 6.6, 0.5 M NaCl.). (D) SDS-polyacrylamide gel corresponding to the protein-containing fractions obtained from the size-

exclusion step. (E) A $^{1}H$ spectrum of $^{15}N,^{2}H$-CR1 20-23. (F) $^{15}N,^{1}H$-HSQC spectrum of $^{15}N,^{2}H$-CR1 20-23.

The high quality of these spectra is, of course, indicative of non-degraded and stable protein in which all of the modules are compactly folded. The HSQC spectrum contains approximately the expected number of cross peaks for the size of protein, are they are of generally quite uniform intensity and well dispersed.

### 4.2.5 Overlay of $^{15}N,^{1}H$-HSQC spectra of $^{15}N$-CR1 21-22 and $^{15}N,^{2}H$ CR1 20-23

The two HSQC spectra (of CR1 21-22 and CR1 20-23) were compared, by overlaying them (see Fig. 4.5), in an attempt to assess the effects of the additional modules (CCPs 20 and 23) on the structure of CR1 21-22, Unlike the comparison of the HSQC spectra of CR1 21 and CR1 21-22, in which it was readily apparent that there was only minimal contact between modules, the comparison of CR1 21-22 and CR1 20-23 is not as easily interpreted. The majority of the cross peaks from CR1 21-22 are close to cross peaks in the CR1 20-23 spectrum, so most of these likely correspond to residues that are not influenced by attachment of modules 20 and 23. There are, however, a proportion (roughly 20%) of CR1 21-22 cross peaks that do not match to cross peaks in the CR1 20-23 spectrum.

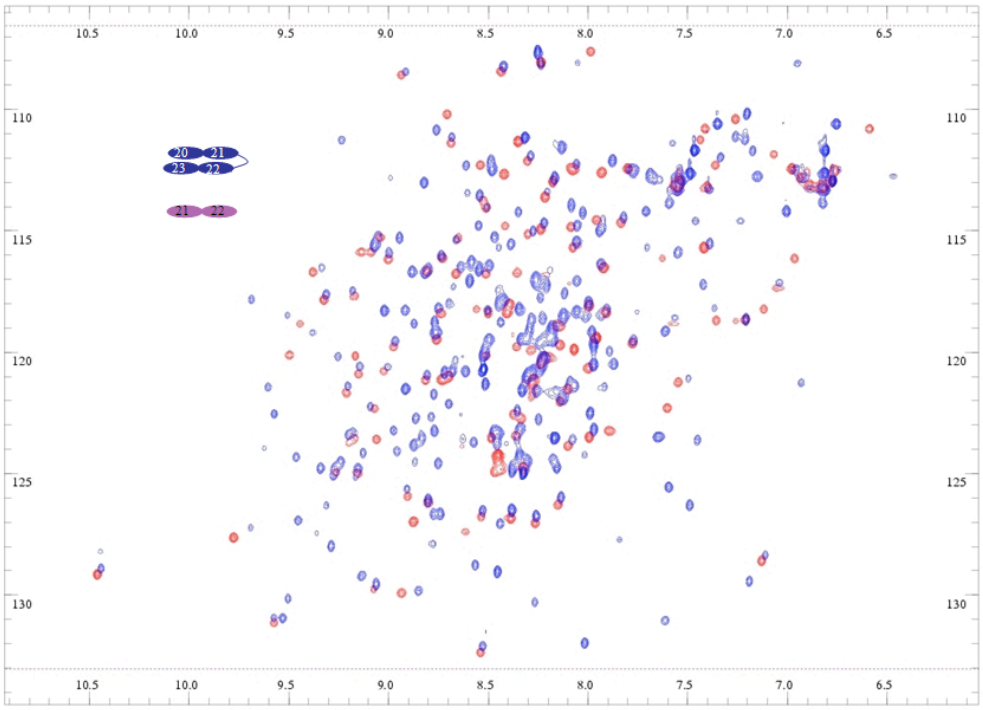



***Figure 4.5 Overlay of HSQC spectra of $^{15}$N-CR1 21-22 and $^{2}$H,$^{15}$N-CR1 20-23***

Cross peaks (positive and negative) in the CR1 21-22 spectrum are red while those from the CR1 20-23 spectrum are purple.

Minor differences in peak positions might have arisen from slight changes in temperature or buffer composition despite efforts to ensure these were consistent. On the other hand, a couple of dozen cross peaks observed in the double-module context seem to have moved substantially in the spectrum of the quadruple-module fragment These might correspond to residues at the N-terminus of CCP 21 and/or the C-terminus of CCP 22 that are involved in end-to-end interfaces with CCPs 20 and 23, respectively, while the relative orientations of CCPs 21 and 22 remains unaffected (note that the short linkers between CCPs 20 and 21, and between CCPs 22 and 23, would seem to preclude a side-by-side interaction, so this possibility was deemed unlikely and not really considered.) Alternatively, some of these perturbations might arise in residues at the CCP 21-22

linker/interface were it to undergo rearrangement in the longer construct. To discriminate between these possibilities, some backbone assignments were required to allow identification of the perturbed residues. In turn this required production of a double-labelled ($^{15}$N,$^{13}$C) sample of CR1 21-22.

### 4.2.6 Production and NMR-based studies of $^{15}$N,$^{13}$C-CR1 21-22

A one-litre fermentation was carried out for production of double-labelled CR1 21-22, following a similar procedure to that employed for preparation of the single labelled ($^{15}$N) protein, but with the appropriate modification as described in Section 2.3.6.4. The processes and conditions for harvesting, and cation-exchange and size-exclusion chromatographies, were all highly similar to those described in the case of $^{15}$N-CR1 21-22. The size-exclusion chromatogram and its corresponding, Coomassie-stained, protein gel (following SDS-PAGE under reducing conditions) are shown in Figures 4.6A and 4.6B (the results of the cation-exchange step are not shown). When the size-exclusion chromatographic fractions were subjected to SDS-PAGE and Coomassie stained, the resultant gel was consistent with a relatively pure and non-degraded protein (Fig. 4.6B).

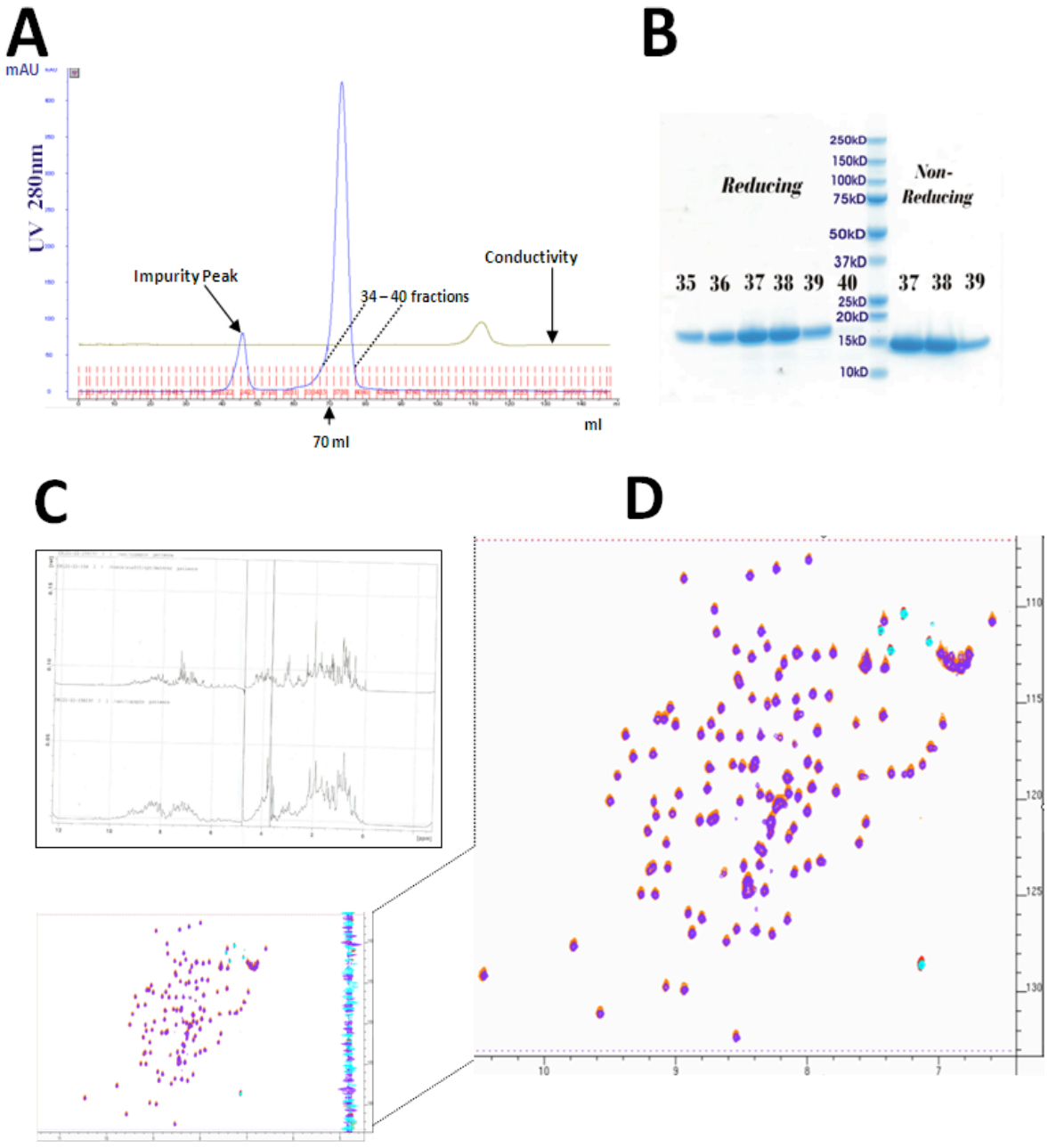



***Figure 4.6 Preparation and NMR studies of $^{15}N,^{13}C$-CR1 21-22***

(A) Size-exclusion chromatography of $^{15}N,^{13}C$-CR1 21-22. Using Superdex-75, the 15 kD protein predictably elutes at 75 ml, and is separated from larger-Mwt contaminating material that elutes as an earlier, smaller, peak. (B) The corresponding Coomassie-stained, protein gel following SDS-PAGE of the protein-containing fragments indicated in panel A. (C) Overlay of the $^{1}H$ spectra collected on single-labelled and double-labelled CR1 21-22 (top) and overlay of the corresponding HSQC spectra (bottom). (D) "Zoom in" on a well-populated region of the overlaid HSQC spectra shown in panel C. Shown in orange are the $^{15}N,^{13}C$-CR1 21-22

cross peaks (negative peaks in red), while in purple are shown cross peaks arising from the $^{15}$N CR1 21-22 sample (negative peaks in cyan).

A large total of 30 mg of labelled protein was obtained from the initial one-liter fermentation; therefore this was divided into two NMR samples of about 1 mM each. One of these was used for the recording all of the NMR experiments needed for the purposes of assignment while the other was kept for future use by shock freezing and storage in a -80° C freezer. Thus, $^{1}$H and $^{15}$N,$^{1}$H HSQC spectra were collected initially, and these showed (Figs. 4.6 C and D) that the double-labeled protein was almost certainly chemically identical to the previously prepared singly labelled CR1 21-22 sample.

### 4.2.7 Partial assignment of 15N, 13C CR1 21-22

Using the $^{15}$N,$^{1}$H HSQC spectrum, along with the CACB(CO)NH and CACBNH NMR experiments (performed on the aforementioned $^{15}$N,$^{13}$C-CR1 21-22 sample), an attempt was made to assign the all backbone nuclei. As detailed in Section 2.5.1, this task was undertaken in a stepwise procedure. To begin with, all cross peaks in the HSQC spectrum were picked and assigned arbitrary numbers (Fig. 4.7). These picked cross peaks (in the HSQC spectrum) were subsequently linked with their corresponding Cαs and Cβs by inspecting the processed spectra from the two complementary NMR experiments (CBCA(CO)NH and CBCANH).

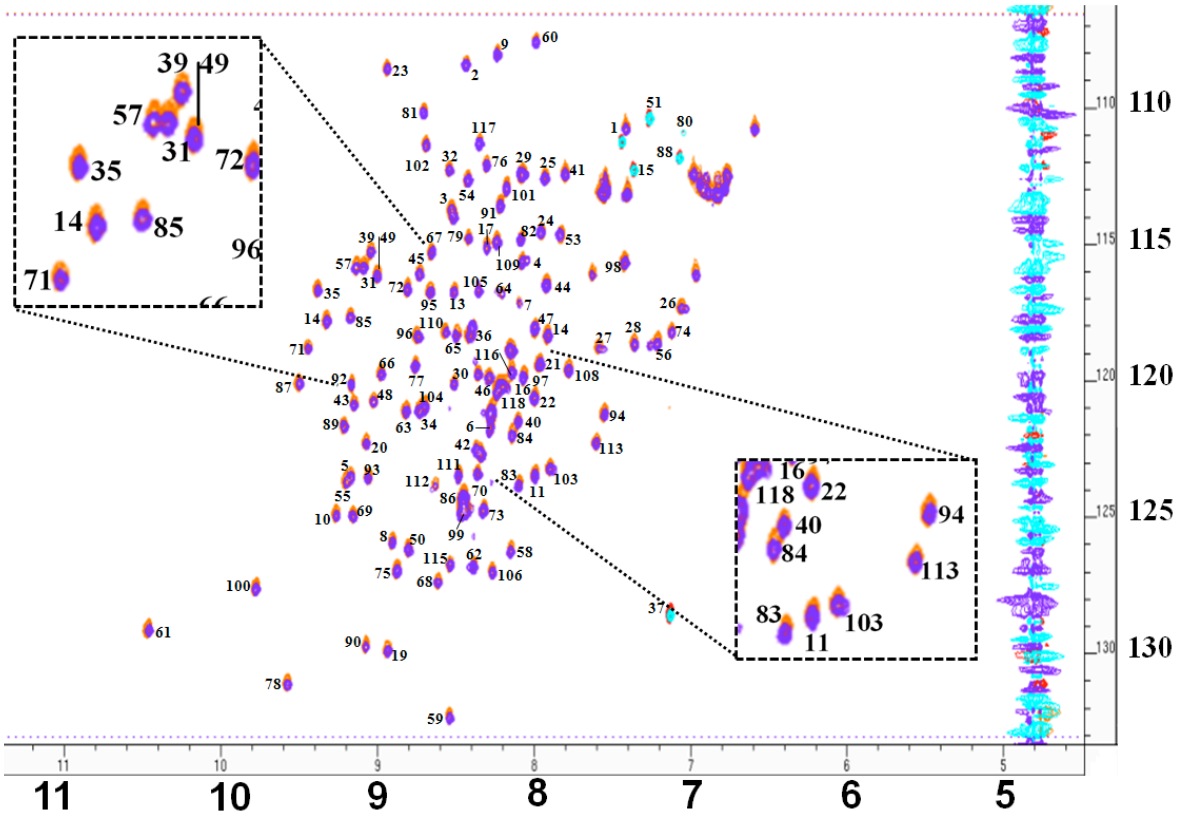



***Figure 4.7 Early step in partial assignment of NMR spectra of CR1 21-22***

Peaks in the HSQC spectrum were assigned arbitrary numbers before subsequent linkage to Cα and Cβ resonances.

**Figure 4.8** demonstrates the stepwise procedure generally used for backbone assignments in which the CBCACONH and CBCANH experiments were employed to help identify the Cα and Cβ of residue *i -1* as well as the equivalent nuclei for residue *i*. By way of exemplification, Figure 4.8 shows the assignment of the cysteine residue that occurs at the C terminus of CCP module 21 and initiates the relatively long linking sequence (eight residues) between modules 21 and 22, culminating in the cysteine that represents the first residue of module 22.

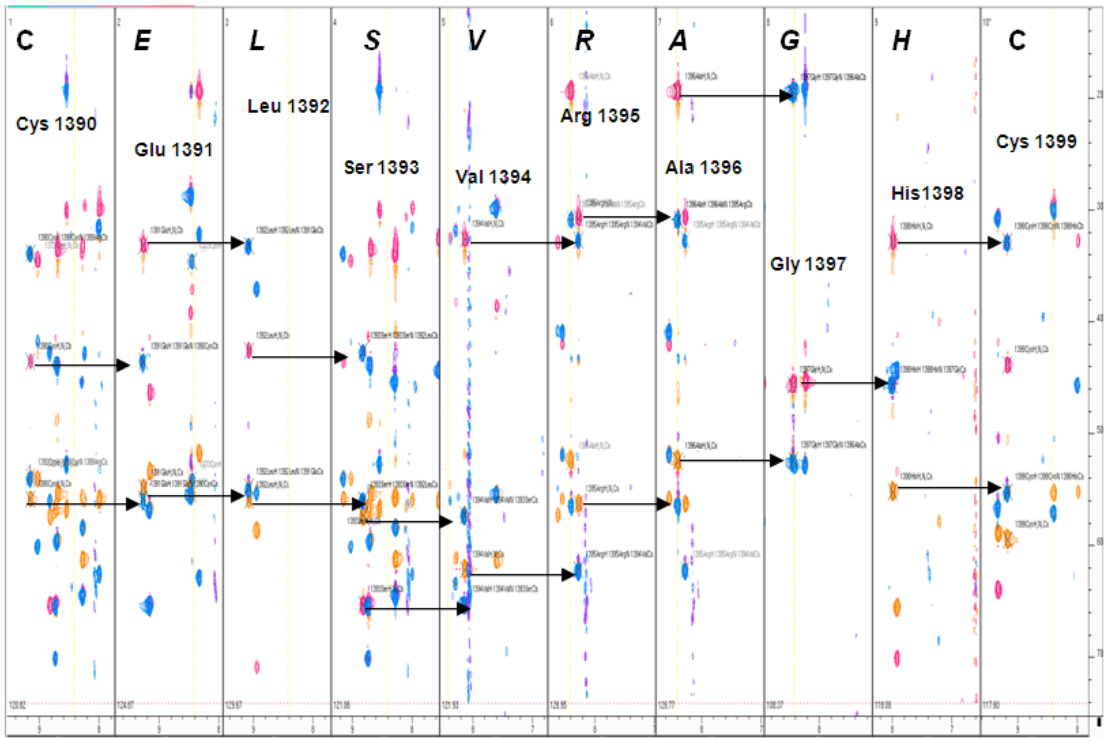



***Figure 4.8 Example of "sequential walk" for backbone assignment of CR1 21-22***

Assignment of the long linking sequence between modules 21 and 22 of CR1 Arrows show connectivities (used for assignment) from the Cα and Cβ of the *i-1* residue to that of residue *i*. The Cα and Cβ in the CBCA(CO)NH expreriment are in blue, highlighting mainly the *i-1* cross peaks. The Cα and Cβ of both residue *i* and residue *i-1*, obtained from the CBCANH experiment, are shown in pink and yellow: yellow for putative Cα and pink for putative Cβ (based on chemical shifts).

In all, about 90% of backbone nuclei in CR1 21,22 were confidently assigned to specific residues while a further 5% were only assigned tentatively (see Fig. 4.9 and the table in appendix D ).

Based on such a partial assignment it was possible to ascertain which residues in CR1 21-22 were affected, in terms of their backbone chemical shifts, by attachment of modules 20 and 23. Figure 4.9 shows the partially assigned HSQC spectrum of CR1 21-22 with residues of the 21-22 linker labelled in a blue font.

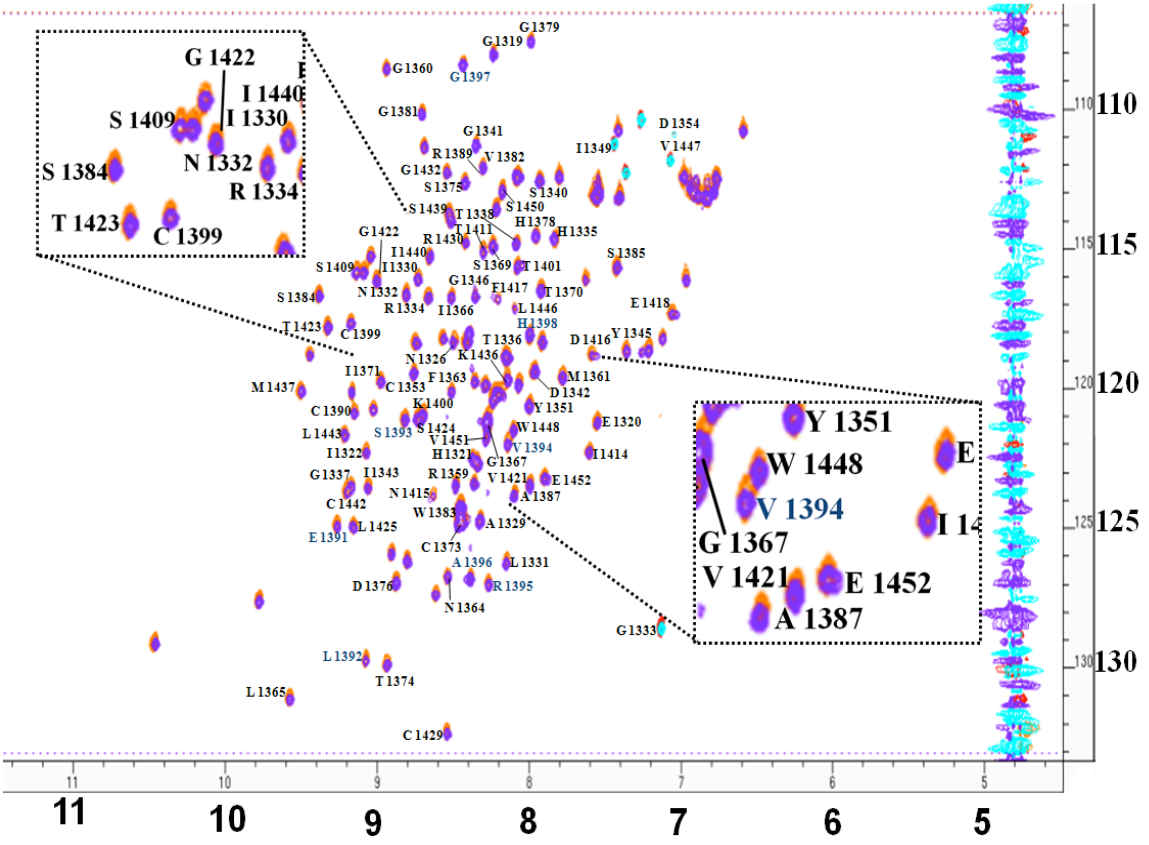



***Figure 4.9 Partial assignment of backbone amides of CR1 21-22***

The $^{15}N$,$^{1}H$-HSQC spectra for single-labelled (purple are for positive and cyan for negative peaks, respectively) and double-labelled CR1 21-22 (orange and red for positive and negative peaks, respectively) are overlaid in this figure. The peaks that are labelled in a blue font are the eight amino acid residues in the linker region.

It would be expected that if the preferred mutual arrangement of CCPs 21 and 22 were to change upon addition of CCP 20 and 23 then the chemical shifts of the linker between CCPs 20 and 21 would change when these extra modules are added to the double-module fragment.

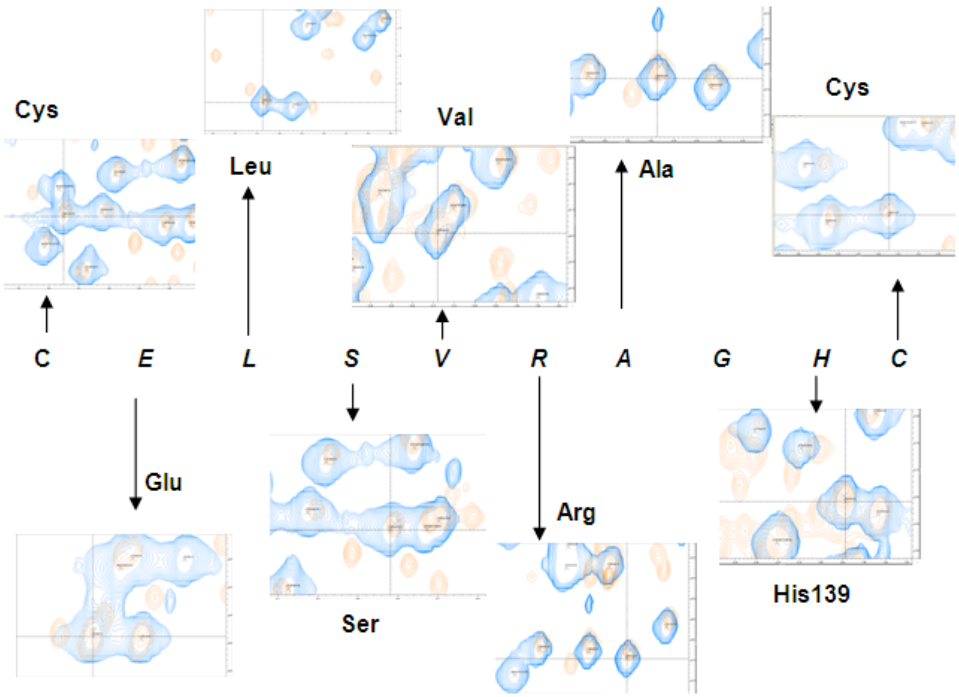



***Figure 4.10 Peaks representing the amino acid of the linker between CCPs 21 and 22***

The linker is written out in single-letter code from left to right along the page. Arrow identify (in the "zoomed-in" pictures) the cross peak (marked with a cross) corresponding to each amino acid residue in the linker.

Figure 4.10 shows expanded regions, containing linker-assigned cross peaks, extracted from the spectrum of Figure 4.11**.** Importantly, from inspection of this figure it is clear that none of these linker residues in fact experience context-dependent chemical perturbations. Since the presence of the additional module (20 and 23) have no effect on the magnetic environment of the eight amino acids in the linker, it may be concluded that these two module retain a non-intimate, probably end-to-end arrangement both as an isolated pair and when within the CR1 molecule. This very strongly reinforces the notion that a side-by-side interaction is highly unlikely between these modules. These observations do not support the hypothesis that LHR-D folds back onto, and thereby functionally modulates, LHR-C.

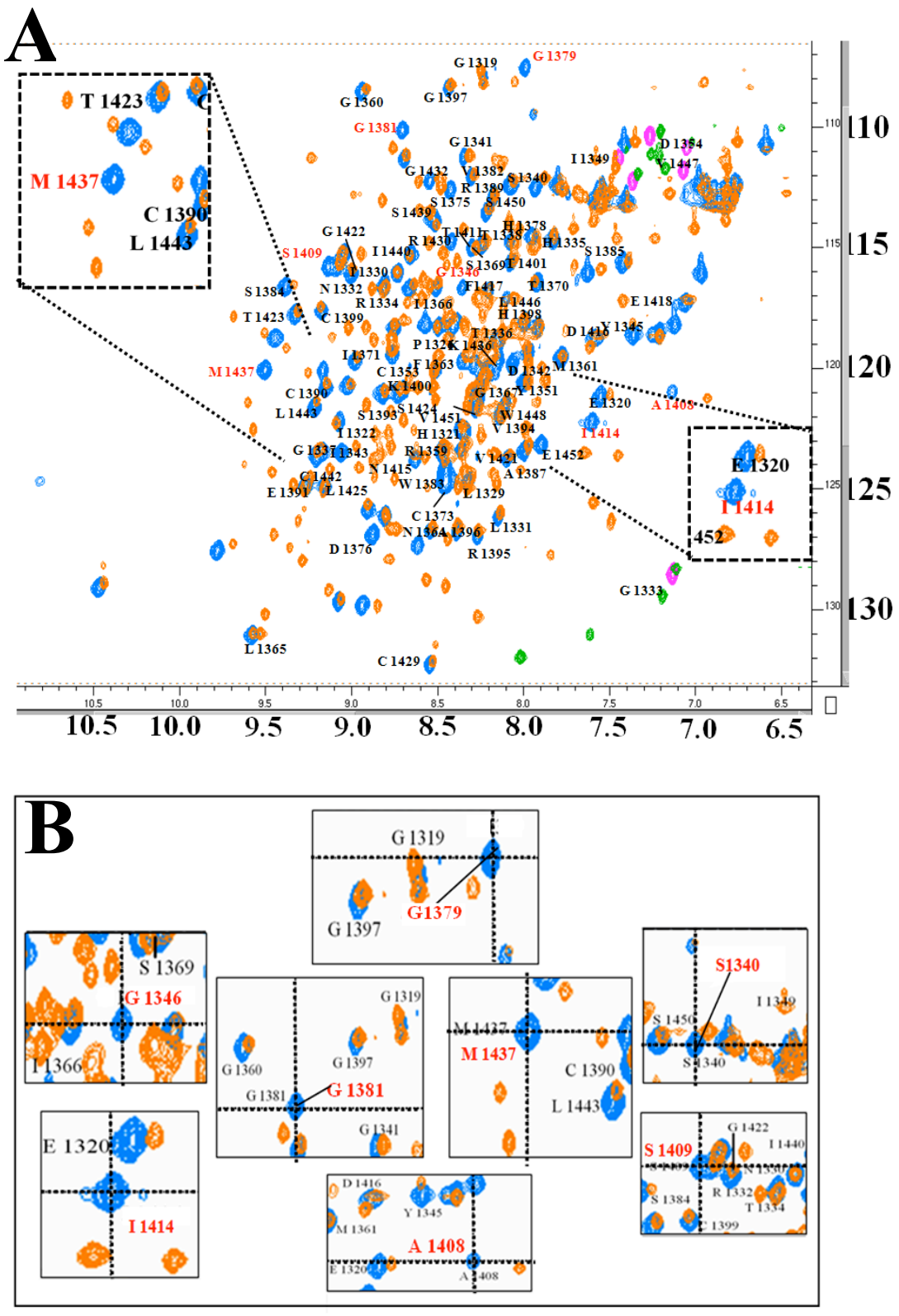



*Figure 4.11 Assessing chemical shift perturbations of residues in the linker between CCPs 21 and 22 arising from attachment at either end of CCPs 20 and 23*

(A) Overlay of spectra for the $^{15}N,^{13}C$-CR1 21-22 sample and the $^{15}N,^{2}H$-CR1 20-23 sample. Positive cross peaks from the spectrum of CR1 21-22 are shown in blue (negative peaks in pink); orange is used to indicate positive cross peaks from the spectrum of the CR1 20-23 fragment (green for negative peaks). (B) Enlarged versions of selected zones to highlight key residues. Red labels were used for the peaks in the CR1 21-22 spectrum that did not overlap with peaks in the CR1 20-23 HSQC spectrum.

As may be judged from Figure 4.11, while backbone chemical shifts within the intermodular linker of CR1 21-22 seemed not to have been affected by the additional modules, the amide signals of non-linker residues including Ser1340, Gly1346, Ala1408, Ser1409, Ile1414 and Met1437, have been substantially "perturbed". Most of these belong to CCP module 22.

Thus the results of these NMR-based studies are fairly conclusive. They rule out the presene of a side-by-side arrangement of CCPs 21 and 22 even though these two modules have a tether between them of sufficient length to allow such a bent-back conformation. Some marked shifts in CCP 22 upon addition of modules 20 and 23 to CR1 21-22 is most easily explained by an intimate end-to-end association between modules 22 and 23 as seen in some other module pairs with four-residue linkers.

## 4.3 CR1 15-25 polymorphic forms behave similarly upon subjection to size-exclusion chromatography

Size-exclusion (or gel-filtration) chromatography has been extensively employed in this study, prinicipally as a final-step purification technique. A well-calibrated column such as the one used in the study ( HiLoad™ 16/60 Superdex™ 75 prep grade column from Amersham Biosciences®) , however, can achieve more than just purification. The volume

at which a prote elutes corresponds to its Stokes radius. A comparison of elution positions of a series of variants should thus reveal the existence of any major differences in shape or self-association, although it might be difficult to disentangle these two properties based on analysis of size-exclusion chromatography alone. It is thus interesting to observe (see Figure. 4.12) that all four variants migrate at a similar rate through the resin (ignoring peak A of CR1 15-25KR), and emerge from the column at very similar elution volumes. This resuls indicaties that all four have near-identical Stokes radii and suggests that they are unlikely to have different intermodular conformational preferences.

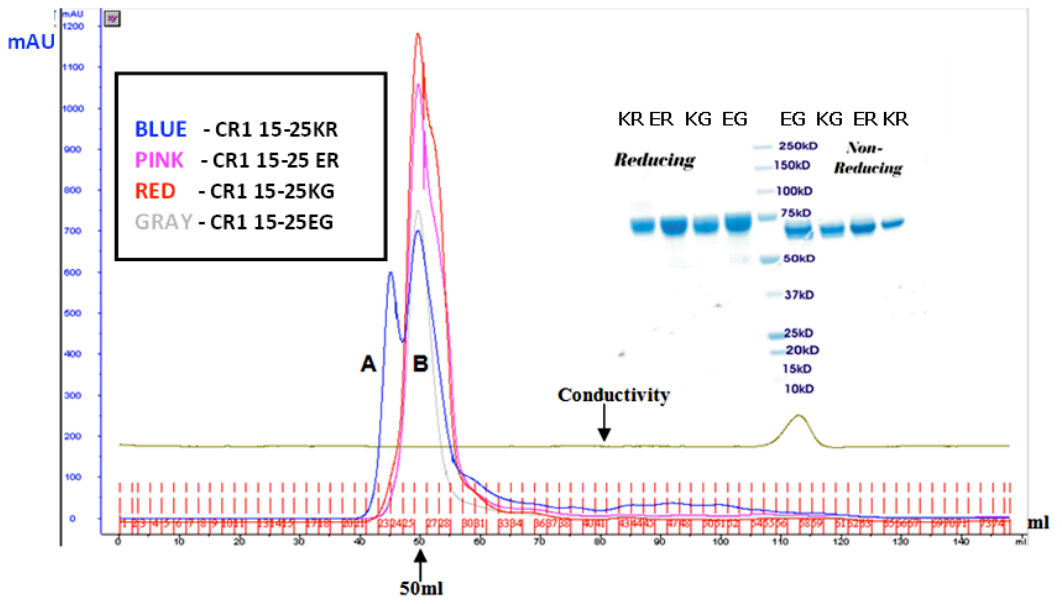



***Figure 4.12 Overlay of size-exclusion chromatographic elution profiles for CR1 15-25 variants***

Colour coded chromatograms are overlaid illustrating co-elution of the four Knops blood group variants. The insert is a Coomassie-stained SDS-polyacrylamide gel following electrophoresis of all four purified variants.

## 4.4 Dynamic light scattering (DLS)

Dynamic light scattering (DLS) is a readily available and easy-to-apply technique that may be used for rapidly assessing the distribution of particle sizes (based on their Stoke's radii) in a solution of protein molecules. Like size-exclusion chromatography, it is particular useful for comparing sets of mutants, or series of fragments, but less useful for determining directly molecular dimensions or molecular weights (since a smaller, but more extended protein will scatter to a similar extent as a larger, globular protein). The set of recombinant fragments of CR1 produced in this study were subjected to DLS to: (a) check them for homogeneity; (b) compare the DLS-derived data for fragments of unknown structure with equivalent data collected for fragments of CR1 and other RCA proteins of known structures; and (c) compare the scattering properties of the four polymorphic variants of CR1 15-25 as a means of elucidating any differences in their overall architctures or self-associative properties. Because the technique is extremely sensitive to particle size, all the protein solutions tested were filtered (0.2 μm pore size.).

### 4.4.1 Study by DLS of CR1 15-17 and CR1 21-22 and comparison with other fragments.

In an experiment designed to provide "baseline" results, the CR1 15-17 sample yielded DLS results consistent with a largely homogeneous sample (judged by inspection of its particle-size distribution as a function of the total volume of particles sampled, rather than as a function of the total scattering intensity of the sample– see legend to Fig. 4.13). The mean particle diameter (estimated from the position of the top of the centre of the left-hand peak in the "by intensity" profile) was in the region of 5.9 nm or ~1.9-2.0 nm

for each of the three modules present in the fragment. This value may be compared with 1.9 nm per module for CCPs 6-8 of factor H (FH) and 1.6 nm per module for FH CCPs 1-4 (all the DLS data for FH fragments quoted in this section were supplied by Dr Christoph Schmidt (Edinburgh) and are unpublished). Experimentally derived 3D structures exist for all three of these proteins; they are all monomeric and largely extended although FH CCPs 6-8 forms a curved ("banana-like") structure and in FH CCPs 1-4 there is a kink or bend between CCPs 3 and 4.

The DLS data collected for the CR1 21-22 sample was also consistent with a largely homogenous preparation and monomeric protein, although there were more "contaminating" larger particles – probably aggregates – than in the CR1 15-17 preparation. In the case of CR1 21-22, the mean particle diameter was in the region of 4.5–4.6 nm or ~2.2–2.3 nm per module. This is somewhat higher than the value of 2.1 nm per module for the almost fully extended and rod-like (experimentally determined) structure of FH CCPs 19-20 (with its three-residue linker). It very strongly suggests that these two modules do not form a compact structure (like, for example, FH CCPs 10-15 for which a value of 1.05 nm per module was observed).

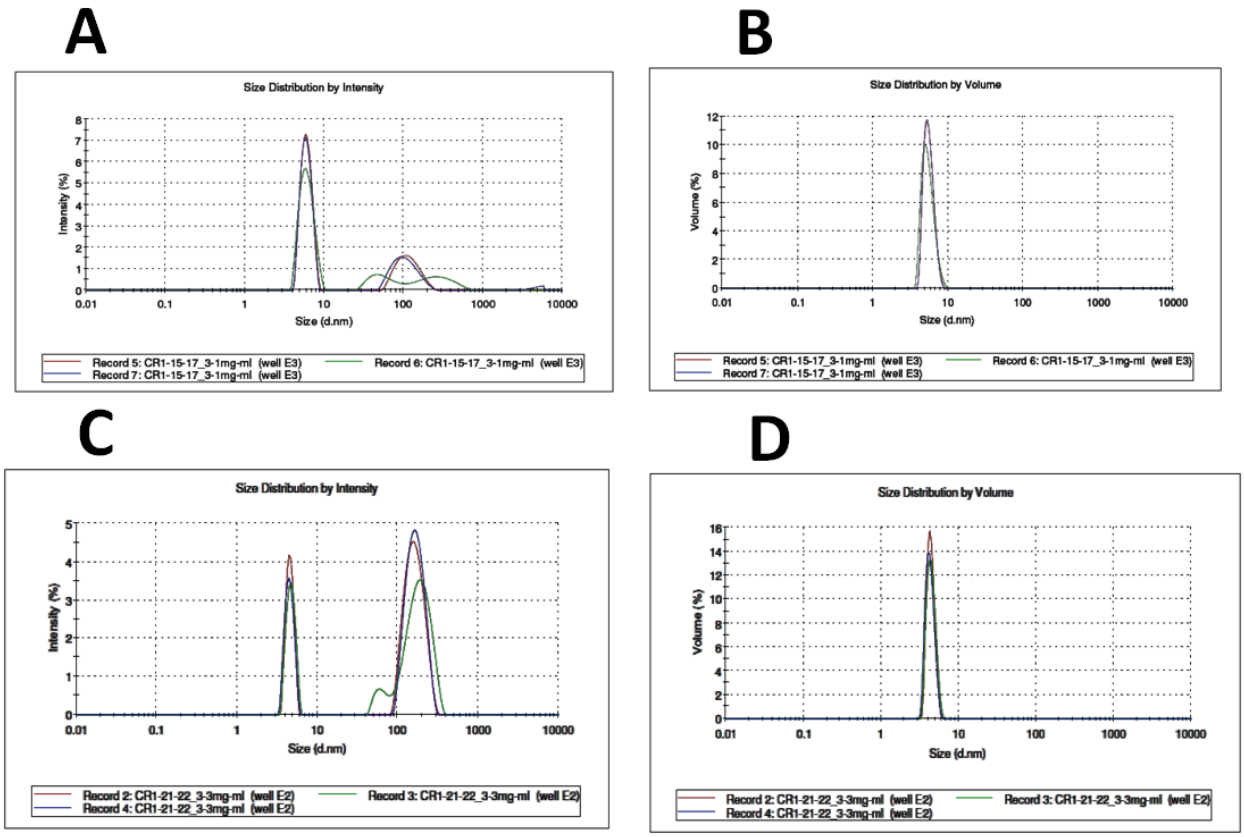



***Figure 4.13 Results of DLS compared for CR1-15-17 and CR1 21- 22***

Shown are DLS-derived particle size profiles by, respectively, intensity and volume for: (A) & (B) CR1-15-17 (C) & (D): CR1 21-22. Note that distributions "by intensity" emphasise the presence of any larger particles due to their disproportionate strength of scattering– hence the distribution "by volume" is generally used to assess the homogeneity of the sample. On the other hand the plot of size distribution by intensity is useful for estimating the average size of the smaller particles in the preparation since they give rise to a distinct peak that is easily analysed.

### 4.4.2 DLS results for CR1-15-25 polymorphic forms

In this batch of experiment, all four Knops blood group variants of CR1 15-25 were assessed by considering particle size distribution both in terms of the total scattering intensity and also in terms of the total volume of particles present. In general, the four variants exhibited particle size profiles (for triplicate experiments) consistent with homogeneous and largely non-aggregated proteins (see Fig. 4.13) but some differences in

profiles are apparent. As explained below, these are probably sample preparation-related and not very significant.

As is apparent from Figures 4.12A-D both the CR1 15-25EG and CR1 15-25ER preparation yielded particularly high-quality data with no indications of any self-association or inhomogeneity. The CR1 15-25KG sample also looked to be largely free of aggregates, even though peaks representing bigger particles are present in the distribution by intensity (Figs. 4.13 E and F). As stated earlier (see legend to Fig. 4.12), even a miniscule population of the bigger particle would be sufficient to generate the signals on the right of the plot in Figure 4.12F. Hence this protein (CR1 15-25KG) behaves in a similar manner to CR1 15-25EG and CR1 15-25ER. In the case of CR1 15-25KR, however, (Figs. 4.13 G and H) the results of DLS are less clear-cut. In this case some significant levels of self-association or aggregation are present; moreover the scattering data for the triplicate measurements were not consistent with “record 10” appearing to be an outlier.

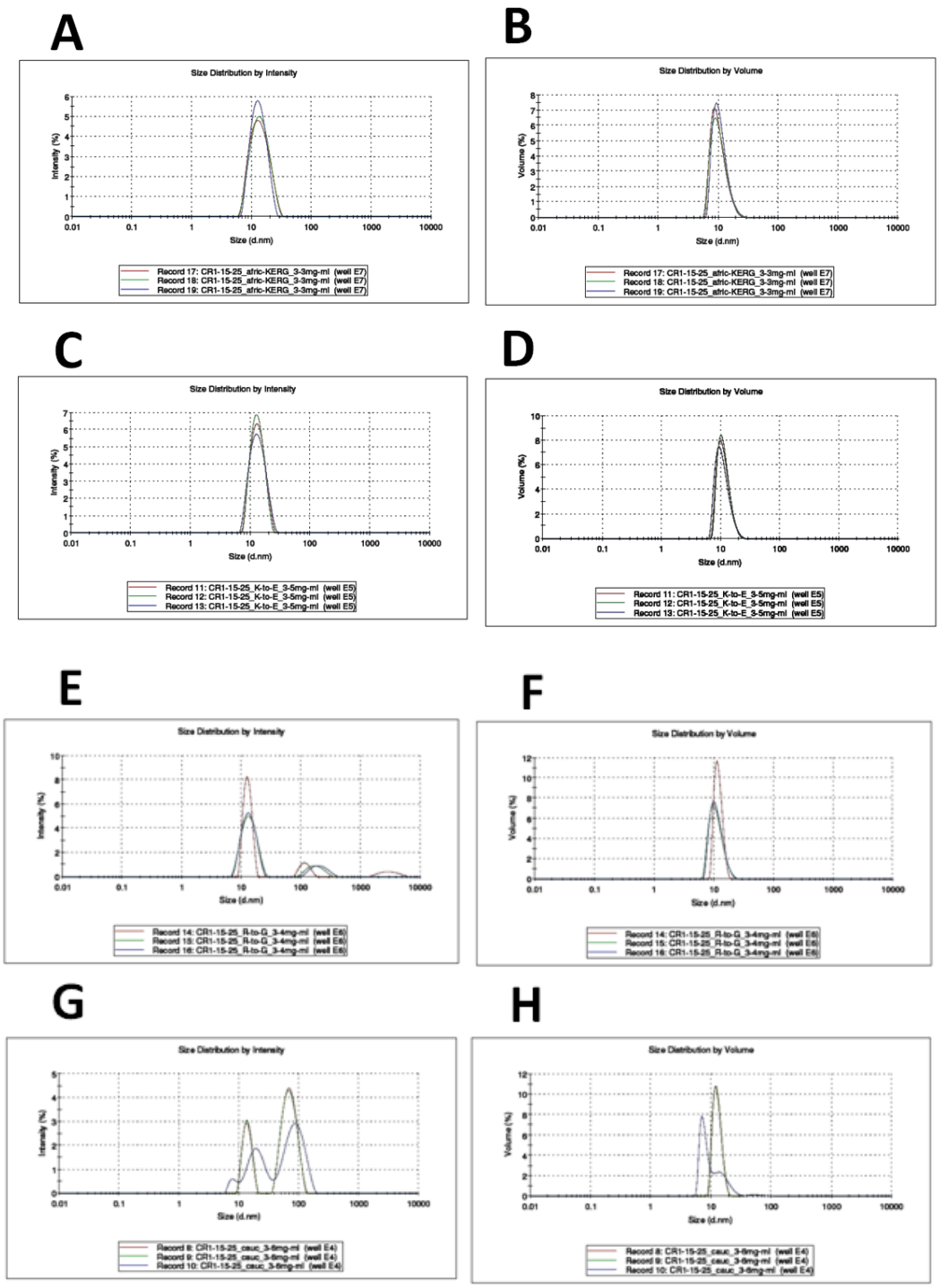



***Figure 4.14 Results of DLS conducted on CR1-15-25 variants***

Size distribution profiles are shown by, respectively, intensity and volume for: (A) & (B): CR1 15-25EG; (C) & (D): CR1 15-25ER; (E) and (F): CR1 15-25KG; (G) & (H): CR1 15-25KR.

In a direct comparison of the four polymorphic forms of CR1 15-25, fresh recordings (one for each, collected consecutively) were made and the profiles overlaid (Figs. 4.14A and B). Here it may be seen that the left-hand peaks in the "by intensity" profiles overlay very well for all four variants with a mean particle size of about 13.5 nm, or 1.2 nm per module. Interestingly, this value is slightly greater than 1.05 nm per module for FH CCPs 10-15 and significantly greater than 0.8 nm per module for FH CCPs 8-15. It suggests that while the eleven modules of CR1 15-25 do not form an elongated structure nor do they form a bent-back structure as envisaged for FH CCPs 8-15.

Note that the "by volume" profiles indicate somewhat smaller and less consistent average particle sizes. This apparent discrepancy was not further investigated due to time constraints.

Note also that parallel studies (below) using analytical ultracentrifugation suggested all four proteins were very similar both in terms of their mass, shape and degree of self-association. Therefore the evidence in Figures 4.13G and H for the presence of oligomers or aggregates in the CR1 15-25KR sample may be attributed to one of the steps in preparation or storage. The DLS data collections were not repeated on freshly prepared samples and so the results for CR1 15-25KR do not constitute good evidence for any physiologically meaningful differences between this and the other three variants in their propensity to self-associate.

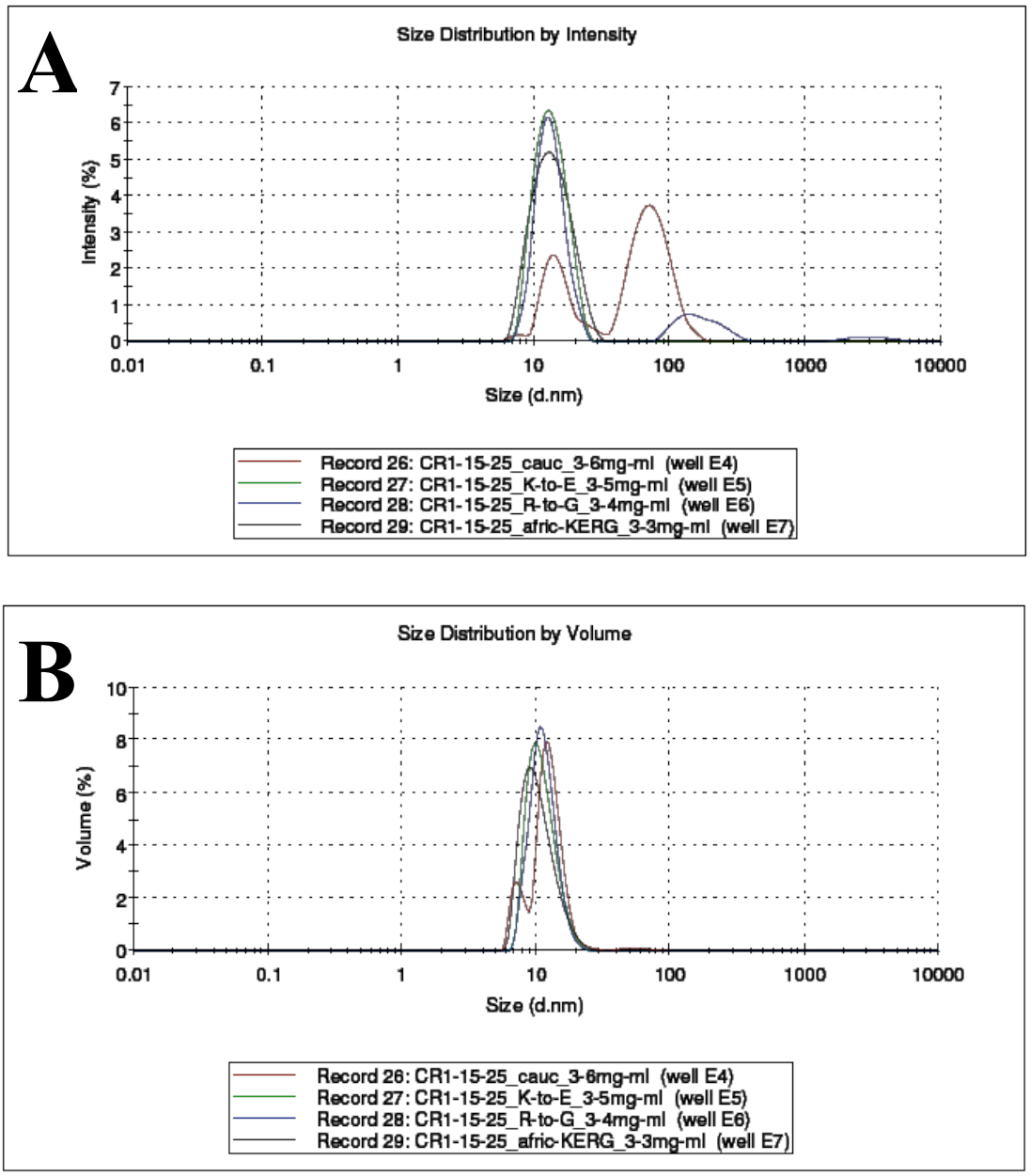



***Figure 4.15 Comparison of DLS-derived particle size profiles for CR1 15-25 variants***

An overlay of CR1-15-25 particle size distributions (by intensity, (A) and volume, (B)) of all polymorphic forms Cauc = KR is CR1 15-25 K1590 and R1601 (predominantly Caucasian), K to E = ER is CR1 15-25 E1590 and R1601, R to G = KG is CR1 15-25 K1590 and G1601 and afric KERG = EG is the double change, E1590 and G1601 (African type).

## 4.5 AUC and small-angle X-ray scattering

Analytical ultracentrifugation (AUC) and small-angle X-ray Scattering (SAXS) each provide information about both the size and shape of proteins. Although they only furnish low-resolution information they proved valuable in the present work to: (a) Furnish evidence (orthogonal to the NMR work) regarding the overall shape of CR1 20-23 – is it extended or folded back; and (b) allow a comparison of the overall architecture of the CR1 15-25 variants that is more sophisticated than what could be achieved with DLS.

### 4.5.1 Ultracentrifugation and scattering data for CR1 20-23

The AUC results (data collected and analysed by Professor Arthur Rowe, Sutton Bonnington) suggested that CR1 20-23 is a highly extended protein especially when they were considered alongside the equivalent data for other multiple-CCP module fragments (Table 4.1). With its four modules, CR1 20-23 had an estimated axial ratio of 6.7. This value may be compared with axial ratios of 4.8 for four-module FH CCPs 11-14 and only 3.4 for six-module FH CCPs 10-15, both of which have been reported to adopt compact structures.

| Sample | Axial ratio (AUC) | |
|---|---|---|
| fH 11-14 | 4.75 | |
| fH 10-15 | 3.39 | |
| CR1 20-23 | 6.72* | |
| fH 12-13 | 1.00 | |
| fH 7-8 | 3.24 | |
| fH 7 | 1.00 | * Red font highlights CCPs of interest |

*Table 4.1 Axial ratios (derived from AUC data) for various multiple CCP module protein fragments*

Further evidence for the extended structure of CR1 20-23 was obtained from analysis of SAXS data. Despite radiation damage-mediated polymerisation upon repeated exposures posing a problem, satisfactory data of sufficient quality were collected (on 1.9 mg/ml and 3.3 mg/ml samples) to allow an ab initio model of CR1 20-23 (see Fig. 4.16) to be fitted to the scattering curve (using the program DAMMIF – this work was carried out by our collaborator Dr Hardyn Mertens at DESY in Hamburg).

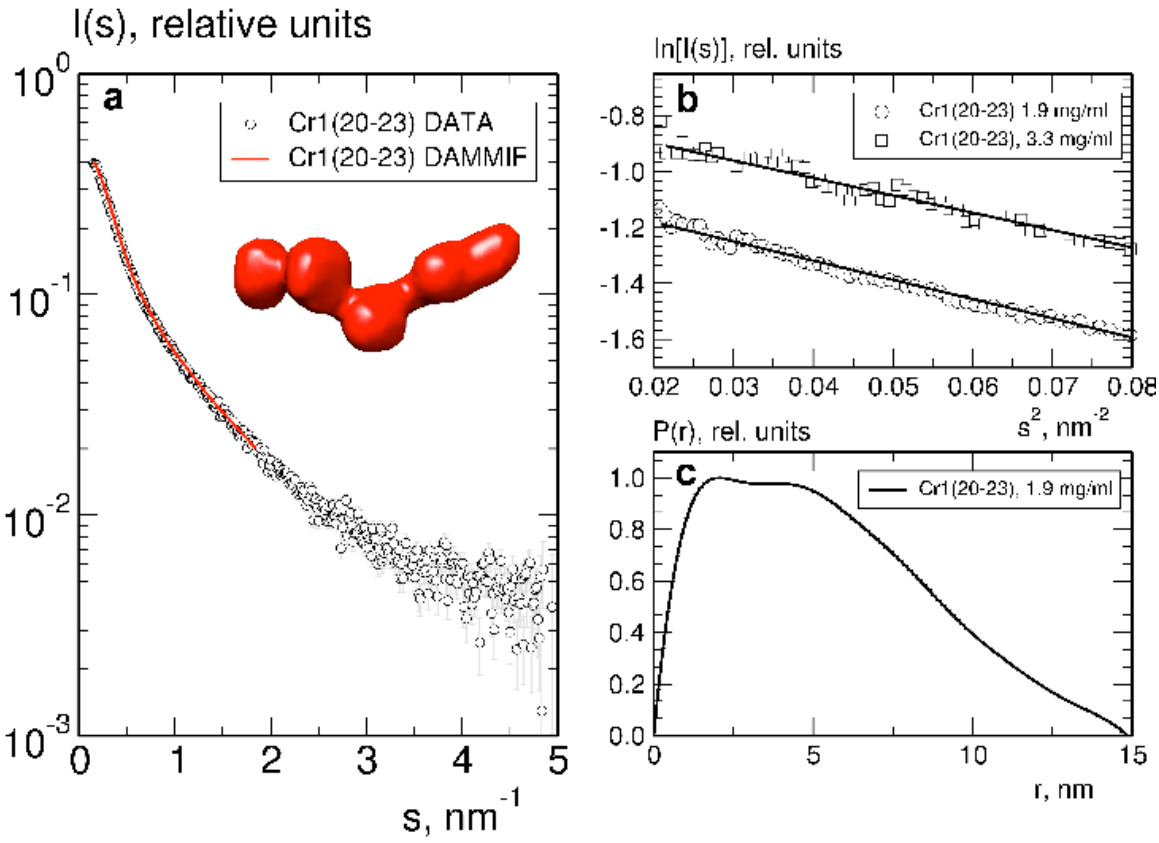



***Figure 4.16 SAXS data and model for CR1 20-23***

(A) SAXS profile and fit of the DAMMIF *ab initio* model to the CR1 20-23 data (1.9 mg/ml). The CR1 20-23 DAMMIF model is shown as a red surface. (B) Guinier plot of the CR1 20-23 SAXS data at 1.9 mg/ml (open circles) and 3.3 mg/ml (open squares), showing no significant

concentration dependence of the SAXS parameters. Plots are displaced on the vertical axis for clarity. (C) Distance distribution function, p(r) for CR1 20-23.

### 4.5.2 Comparison of CR1 15-25 polymorphic variants by AUC

All four variants, as 0.5 mg/ml solutions, were subjected to AUC (see Chapter 2 for more details; work performed and analysed by Professor Arthur Rowe, Nottingham). The outcome is summarised in Table 4.2. Their estimated sedimentation coefficients (related to molecular weight, and hence to oligomerisation state) were all very similar and lay in a narrow range from 3.13 S to 3.22 S. Their frictional ratios were also similar, ranging from 1.81 to 2.00. Frictional ratios are related to axial ratios (more detailed calculations are underway in Professor Rowe's lab but were not available at the time of writing and were beyond the scope of this thesis). Thus the AUC data are consistent with all four variants having the same monomeric oligomerisation state and molecular dimensions. For comparison, the frictional ratio obtained for the compactly organised FH CCPs 10-15 fragment was 1.14. Hence these data suggest that CR1 15-25 (in all four of the variations investigated) adopts a more extended structure than the central modules of FH.

| Sample | mg/ml | s (S) | s(20,w) (S) | f/f0 SEDFIT |
|---|---|---|---|---|
| KR | 0.5 | 3.13 | 3.25 | 2.10 |
| EG | 0.5 | 3.16 | 3.28 | 1.94 |
| ER | 0.5 | 3.22 | 3.33 | 1.81 |
| KG | 0.5 | 3.13 | 3.25 | 2.00 |

*Table 4.2 Comparisons of AUC data for four polymorphic forms of CR1 15-25*
***(Perform by Prof. Arthur Rowe's Laboratory)***

## 4.6 Summary of Structural Studies

By using an array of biophysical techniques it has been possible to answer most of the questions that motivated this part of the study. LHR-C and LHR-D are unlikely to be in close spatial proximity in the non-liganded form of CR1. Despite the potential offered by the uniquely long linker between these regions of the exodomain, none of NMR chemical shift perturbations, AUC nor SAXS support the hypothesis that the modules on either side of the long linker fold back against one another to create a U-turn in the molecule. This picture holds true for all four of the Knops blood group variants investigated since – in the CR1 15-25 context - all have very similar dynamic light scattering properties and migrate similarly in the ultracentrifuge. Moreover, all four CR1 15-25 variants are monomeric and there is no evidence that the polymorphisms module self-associative properties. The implications of these findings when considered in conjunction with the functional data are further discussed in Chapter 6.

# CHAPTER FIVE
# BIOLOGICAL AND FUNCTIONAL STUDIES

# CHAPTER 5 BIOLOGICAL AND FUNCTIONAL STUDIES

## 5.1 Overview of biological and functional studies

This chapter covers results obtained from the biological and functional aspects of the current project. These studies aimed to find out whether the Knops blood-group polymorphisms located in CCP modules 24-25 influence how CR1 interacts with its principal known ligands. A range of assays was employed including experiments that utilised cultures of the parasite (*e.g.* rosette-disruption assay and invasion-inhibition assays), measurements of complement regulatory activity on purified proteins, and *in vitro* experiments to measure directly protein–protein interaction (*e.g.* surface plasmon resonance and ELISA). As indicated below, some of this work was done in collaboration with labs elsewhere in Edinburgh and in Australia and the USA.

## 5.2 Co-factor assays

As was described in the Introduction, CR1 is an important regulator of the complement system that has both decay-accelerating and co-factor activities directed at both alternative and classical pathways of complement activation. It is the only complement regulator that (under physiological ionic strength conditions) acts as co-factor both for cleavage of C3b to iC3b, and of iC3b to C3dg/C3d. The co-factor activity resides primarily in functional sites 2 of CR1 (modules 8-10 and 15-17), hence several of the constructs prepared for this study were expected to be active in fluid-phase co-factor assays. These were carried out by observing (*via* SDS-PAGE) the cleavage products of purified C3b upon incubation with factor I and the various CR1 fragments.

In the current study, the co-factor assay was used initially (Figs. 5.1A-D) to confirm whether the purified over-expressed proteins were active and then to compare

activities of the four polymorphic variants of CR1 15-25 with one another. Note, however, that this assay was not well suited to discerning, in a quantitative manner, small differences in activity since it is an "end-point" assay (*i.e.* it does not monitor product in a continuous manner over time); moreover it was not straightforward to establish conditions under which the amount of product, produced at a particular time-point, was proportional to the amount of cofactor activity present.

Initial fluid-phase co-factor assays were performed on samples (total volume 20 μL) containing a mixture of C3b (2.5 μg), Factor I (0.1 μg) and either 0.25 μg or 1 μg of the appropriate CR1 construct, followed by incubation for either 15 minutes or 1 hour. At the end of the incubation, the reaction was stopped by adding 6 μL NuPAGE reducing buffer, pH 8.4, containing lithium dodecyl sulfate (Invitrogen). Samples were then heated and run on SDS-PAGE.

Coomassie staining was used to reveal the presence of the 68-kD and 43-kD proteolytic fragments derived from the initial factor I-catalysed cleavage of the α'-chain of C3b and, potentially, the products of the (second) proteolytic cleavage of the 68-kD fragment (expected only when CR1, or a biologically active recombinant fragment of CR1, is the cofactor), namely C3dg (~39 kD) and a 29-30-kD fragment (Figs. 5.1A-D).

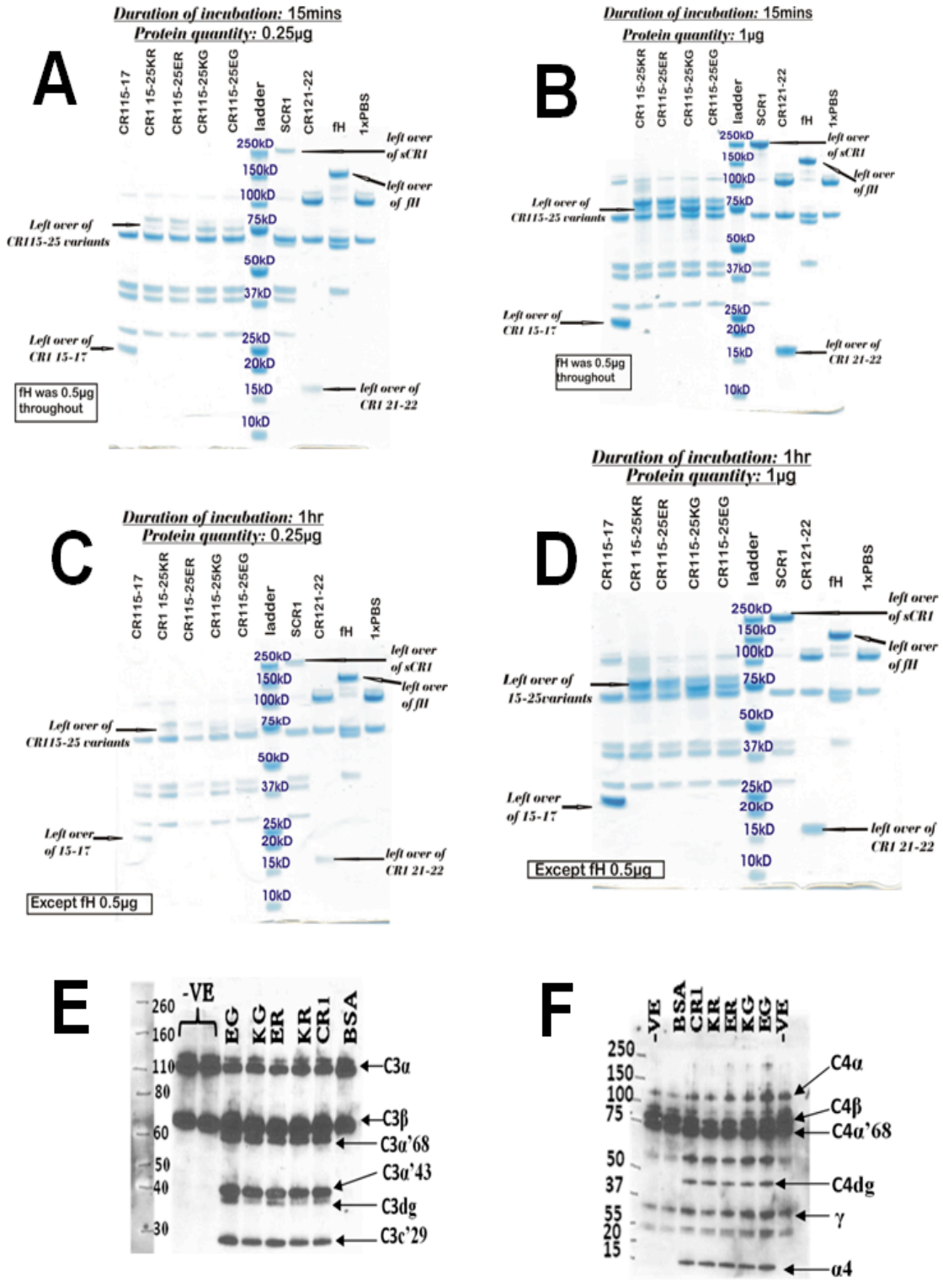



***Figure 5.1 Cofactor Assay testing the polymorphic forms of CR1 15-25***

In this assay activity is evidenced by detection of proteolytic cleavage products of the C3b α' chain at 68, 43, 39 and 29 kD. ***(A)*** Incubation for 15 mins and addition of 0.25 µg of co-factor; ***(B***) 15 min duration as in (A) but addition of 1 µg of co-factor; ***(C)*** Incubation for 1 hour with

0.25 μg of added co-factor; ***(D)*** Incubation for a duration of 1 hour, following addition of 1 μg of co-factor. ***(E)*** Result of cofactor assay (for cleavage of C3b) carried out in the Atkinson lab using a more sensitive protein detection method and therefore smaller quantities of reagents (see Materials and Methods). (***F)*** Same as panel (E), except that in this panel the protein being degraded is C4b. In panels (E) and (F) a slightly different nomenclature was used for labelling and KR is CR1 15-25KR, ER is CR1 15-25ER, KG is CR1 15-25KG and EG is CR1 15-25EG.

Three cleavage products of the C3b α'chain (43, 37 and 29 kD) were observed (following electrophoresis and Coomassie staining, see Figs. 5.1A-D) after 15 minutes incubation following addition of 0.25 μg of all four polymorphic variants of CR1 15-25. Thus unlike the negative controls (*i.e.* CR1 21-22, or buffer only), each variant acted as co-factor for cleavage by factor I, just like the positive control (sCR1) and the previously characterised CR1 15-17 fragment. This confirms that correctly folded modules of site 2 (CCPs 15-17) are present within each of the longer fragments and that all four variants show comparable co-factor activity, with no evidence for modulation of functionality *via* residues in CCP 25.

Note from the results in Figures 5.1A-D that, as expected, sCR1 and CR1 constructs containing functional site 2, are co-factors for cleavage of the C3b α'-chain (111 kDa) into the α'-chain fragments of iC3b (43 and 68 kDa) and then its further cleavage to yield C3dg and the 29-kD α'-chain remnant that belongs to C3c; on the other hand factor H cofactor activity is limited to cleavage to iC3b. Similar results were obtained for both shorter and longer (15 mins and 60 mins) incubation periods and for both low and high additions of CR1 construct (0.25 μg or 1 μg), indicating that we had not arrived at assay conditions under which cofactor amounts were limiting and hence would not have detected minor differences between the variants.

Further co-factor assays were carried out in the Atkinson laboratory. For these

experiments a more sensitive protein-detection methods was used: In the C3b (Fig. 5.1E) cofactor assay, 25 pg of biotinylated C3b, 8.2 pg of FI, and 10 pg of CR1 15-25, (or 27 pg of sCR1 2.7) were mixed in a final reaction volume of 30 μl of phosphate buffer, pH 7.3, containing 50 mM NaCl. For the C4b assay (Fig. 5.1F) 37.5 pg of C4b, 3.8 pg of FI and 50 pg of CR1 15-25 (or 135 pg of sCR1) were mixed in 30 ml of the same buffer as above. Incubations were carried out for one hour then stopped as above and run on 4-12% NuPAGE gels (Invitrogen). Protein bands were detected *via* avidin-horse radish peroxidase and electrochemical luminescence as described in Chapter 2.

## 5.3 Rosette-disruption assays

All experiments whose results will be shown in this section were performed in Professor Alex Rowe's laboratory and under her kind supervision. The details of the rosette-disruption assay and the principles behind it have been presented in the Materials and Methods chapter. As indicated, a successful outcome of this assay depends largely upon managing to maintain a viable "rosetting culture" *i.e.* a preparation of erythrocytes in which some cells are infected with the *P. falciparum* parasites and wherein rosettes have been (or are being) formed between parasitized and non-infected cells. Two of the key processes in maintaining such a culture are selection and synchronisation, both of which were detailed in Chapter 2. Synchronisation was achieved when a mixed culture of ring and trophozoic forms (see Fig. 5.2A) were treated with sorbitol (for details explanations see section 2.4.4.2). The resultant culture, with many ring forms, is Giemsa stained (to reveal parasitized cells) as displayed in Figure 5.2B.

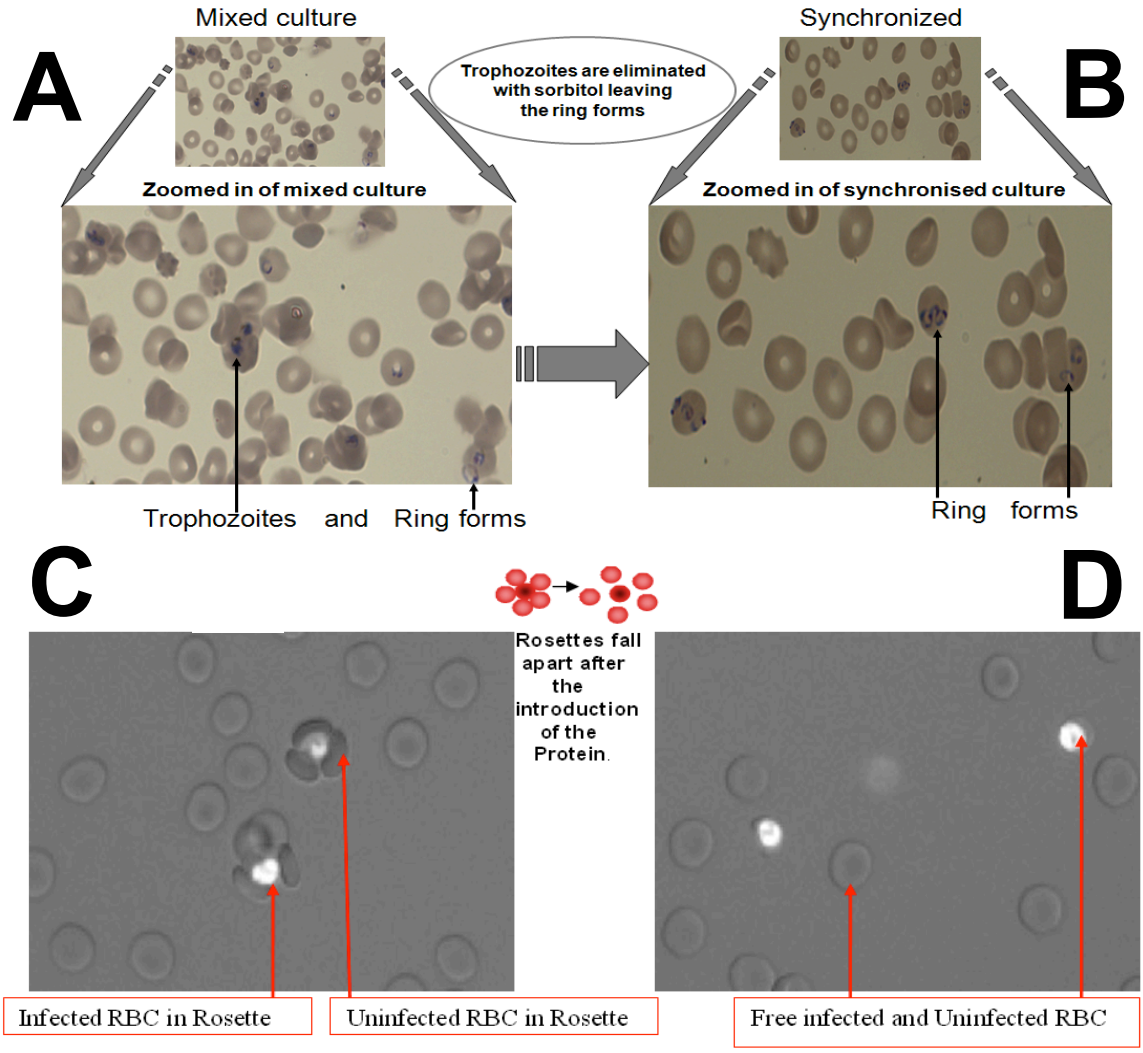



***Figure 5.2 Plasmodium falciparum in vitro culture techniques and rosetting assay***

**(A)** A Giemsa-stained slide of a mixed culture of ring and trophozoite forms. **(B)** A Giemsa-stained slide of a synchronised culture showing, primarily, ring forms. Lower slides are "zoom ins" of the slides shown above. **(C)** Ethidium bromide preparation of rosetting culture before addition of any potential rosette-disrupting proteins. **(D)** Ethidium bromide preparation of cells (as in (C)) after addition of rosette-disrupting protein.

Within the next cycle, a culture that contains numerous rosettes ("rosette-rich") is achieved from the trophozoite-containing red blood cells. Figure 5.2C shows an ethidium bromide-stained rosette-rich culture, examined under the fluorescence microscope.

Infected cells stain positively (fluoresence) due to presence of parasite DNA in the anucleate red blood cell. Normal light microscope can also be used to visualize erythrocytes irrespective of their infection status. Thus a combination of light and fluorescence microscopy helps visualize rosettes and makes counting possible. Using this technique, between three and four uninfected red blood cells can be seen attached to each of the two infected ones in Figure 5.2C. Figure 5.2D demonstrates the disruptive effect on rosettes of sCR1. Destruction of rosettes is assumed to be the result of competition between sCR1 and CR1 on the erythrocyte surface for binding to PfEMP1 on the surface of infected erythrocytes.

The experimental results featured in Figure 5.3A, B and C were obtained in assays performed with shorter constructs to define the minimum-size fragment capable of disruption of rosettes but also as a proof of principle that the longer constructs could be tested in this assay with reliable results. From sampling various protein concentrations it was found that 20 μM CR1 15-17 (Site 2) had significant disruptive effect on the rosettes, unlike the negative control of 20 mM potassium phosphate buffer alone. The J3B11 antibody, which recognises at epitope in module 17 of CR1, served as a positive control throughout the three experiments summarised in Figure 5.2 A, B & C. It consistently had a dramatic negative effect on the number of rosettes present. This experiment further demonstrated that CR1 10-11 (*i.e.* equivalent to CR1 17-18) and (perhaps unexpectedly) single-module CR1 17 reproducibly disrupted rosettes as effectively as CR1 15-17 and J3B11.

To prove that not all recombinant versions of CR1 fragments had negative effects on rosettes, CR1 21, CR1 21-22 and CR1 20-23 were included as negative controls

(Figure 5.3B). While the positive control (J3B11 antibody) worked well, this experiment showed as expected that none of these fragments had rosette-disrupting properties.

In addition, protein constructs from factor H, which is also a member of the regulators of complement activation family, were also tested in this assay (Figure 5.3C). Again the positive control, J3B11, worked well, but factor H fragments did not significantly disrupt rosettes. Addition of full-length factor H resulted in a small reduction in rosettes that was significant when compared to the binding-medium negative control, but not significant when compared to the PBS negative control. The result obtained with full-length factor H are therefore equivocal and the experiment would need to be repeated.

Overall these results served to establish a workable, practical protocol, and suitable negative and positive controls, for subsequent functional testing of the polymorphic variants of CR1 15-25. All four polymorphic forms of CR1 15-25 had a considerable disruptive effect on rosettes that was comparable to that of CR1 15-17 and highly significant against negative controls with PBS and CR1 21-22 (Fig. 5.3D). (Note that antibody J3B11 was unfortunately not available to serve as the positive control in this last experiment). Interestingly, there was no measurable difference between the variants in their ability to disrupt rosettes and hence it may be inferred that all four have similar affinities for PfEMP1(For analytical tables and raw data see appendix E ). The conclusion of this study is therefore that the propensity to form rosettes – a property that is associated with susceptibility to severe forms of malaria – has not exerted selective pressure on the Sl1/2 and $McC^{a/b}$ alleles in the *CR1* gene and cannot explain their non-uniform geographical distribution.

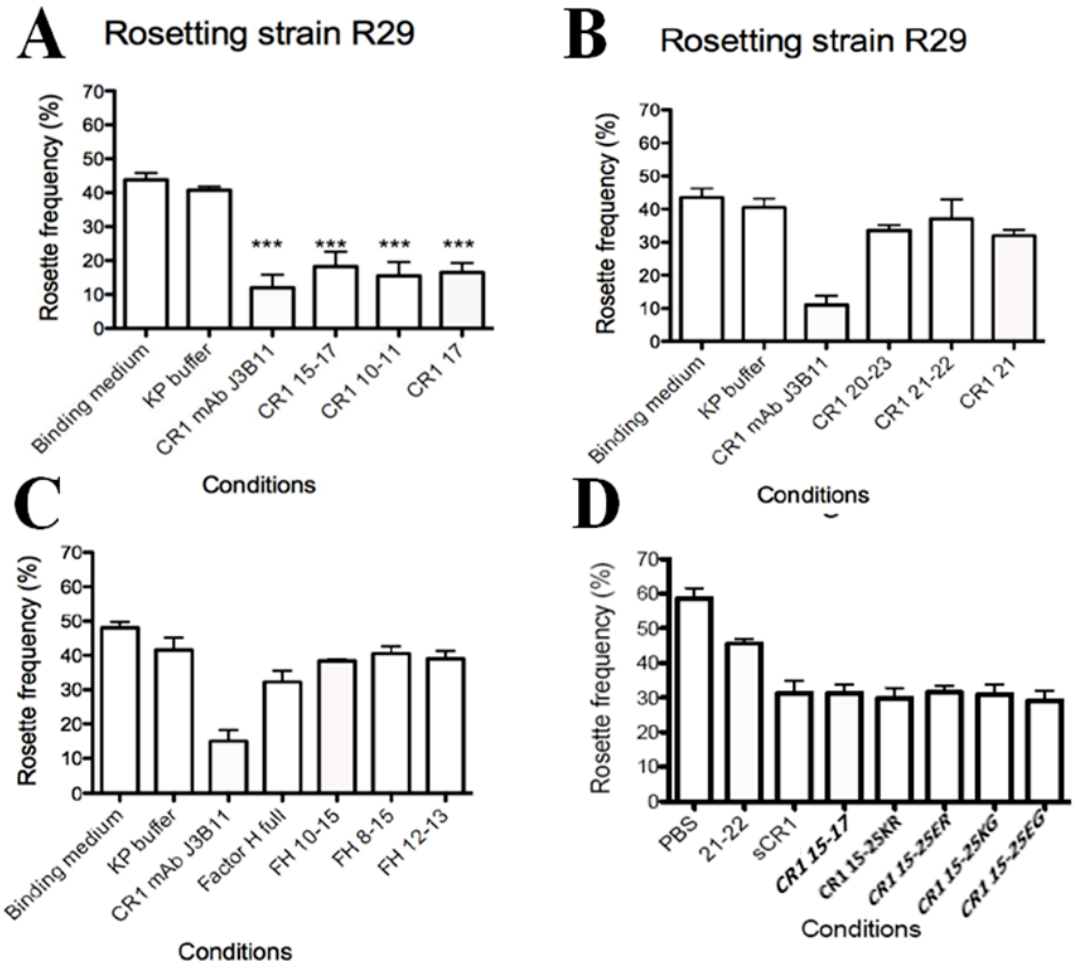



***Figure 5.3 Rosette disruption assays***

***(A)*** Experimental results of adding CR1 15-17, CR1 17 and CR1 10-11 alongside the J3B11 antibody (positive) control and buffer (PBS)-only negative control.*** P value <0.005 (***B)*** Other CR1 constructs tested alongside J3B11 antibody. (***C)*** Factor H constructs tested in this assay, along with antibody (to CCP 17) J3B11 (***D)*** Results obtained for the four polymorphic forms of CR1 15-25 alongside sCR1, CR1 15-17 (positive controls) and CR1 21-22 (negative control).

## 5.4 Erythrocyte-invasion assays

These experiments were carried out with Edinburgh-produced and purified proteins that had been dispatched to Professor Alan Cowman's laboratory in Melbourne. The assays were performed by Dr Wai-hong Tham according to the procedure briefly outlined in Materials and Methods. Invasion experiments were conducted using cultures of 3D7 and W2mefΔRh4 strains of *P. falciparum*. The latter are deficient in the protein, Rh4, suspected to interact with CR1. Since it is the sialic acid-independent invasion pathway that was suspected of being mediated by CR1, neuraminidase treatment was employed to remove access to other potential invasion pathways.

Thus when PBS (*i.e.* the buffer-only negative control) is added, invasion efficiency into 3D7 cells is independent of neuraminidase treatment and is set to 100% (Fig. 5.4A). On the other hand neuraminidase treatment results in loss of invasion capability of W2mefΔRh4 cells. Addition of sCR1 (*i.e.* the full-length ectodomain, CCPs 1-30) has a very dramatic effect on invasion of neuraminidase-treated red blood cells by 3D7 parasites. This was presumed to arise from competition between sCR1 and erythrocyte-borne CR1 for binding to Rh4 (since sCR1 had no effect on invasion by W2mefΔRh4 cells into untreated red blood cells). This experiment also demonstrated that CR1 1-3 (a construct prepared in Edinburgh by Dr Mara Guariento) was almost as effective as sCR1 in the invasion-inhibition assay. This construct corresponds to functional site 1. Interestingly, the CR1 15-17, prepared in the current study, (i.e.

functional site 2) had no significant invasion-inhibition properties despite the aforementioned very high levels of sequence similarity between functional sites 1 and 2.

None of the four polymorphic variants of CR1 15-25 had a major effect on invasion by 3D7 of neuraminidase-treated red blood cells. Although addition of CR1 15-25KR appeared to have some effect, this is barely significant with respect to controls and not significant compared to the other variants. Such small discrepancies between variants could arise from slightly different buffer conditions or as the result of differences observed during the purification; further experiments would be required to investigate this.

In summary, the Edinburgh-produced set of recombinant CR1 fragments allowed the delineation of the binding site of Rh4 on CR1, which is required for sialic acid-independent invasion of red blood cells, to functional site 1. The constructs made in the current study (CR1 10-11, 15-17, 17, 20-21 and 15-25) were crucial to demonstrating that the affinity of CCPs 1-3 for Rh4 is unique amongst the modules of CR1. Moreover, the availability of all four variants of CR1 15-25 established that this observation holds true for geographically diverse populations. The lack of any difference between the variants suggests that susceptibility to invasion has not exerted selective pressure on the Knops blood-group antigens in CCP 25.

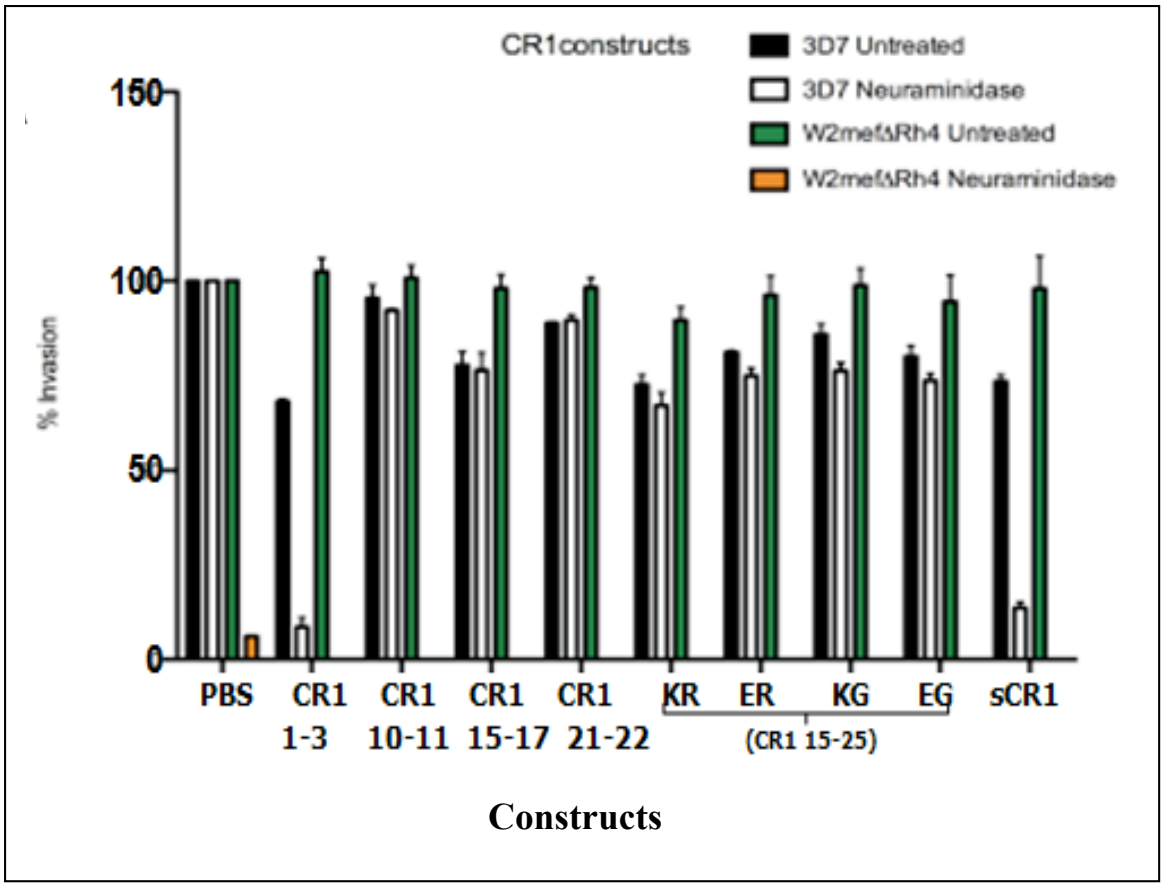



***Figure 5.4 Erythrocyte-invasion assays and ELISA***

(A) Histogram to show the effects of various CR1 fragments on invasion of erythrocytes by *P. falciparum* merozoites. The horizontal axis of the graph indicates the constructs used in the invasion assay. The colour-coding indicates the strain of *P. falciparum* investigated (see text) and whether or not cells were treated with neuraminidase to remove sialic acids and hence block the sialic acid-dependent invasion pathway (that does not utilse Rh4:CR1). PBS = phosphate buffered saline; sCR1 = soluble CR1.

# CHAPTER 5 BIOLOGICAL AND FUNCTIONAL STUDIES

## 5.5 Acquisition of *P. falciparum* proteins

In this study, protein-protein interaction was assessed not only in culture-medium settings as described earlier, but also by direct measurement of protein:protein interactions. For surface plasmon resonance (SPR) studies, purified proteins are essential. Since all the requisite CR1 constructs had been produced successfully as described in Chapter 3, the next task was to acquire the binding partners in purified form. The three complement proteins used (C3b, C4b and C1q) were purchased from Complement Technology, Texas. On the other hand, the DBLα (Figure 5.5D, and see Introduction for further details of PfEMP1) believed to be the domain of PfEMP1 that interacts with CR1 during rosetting, was kindly provided by Dr. Matt Haggin's laboratory at the University of Cambridge (now at University of Oxford). This gift was timely after attempts to make useful amounts of Dblα-containing constructs in Edinburgh failed to yield the hoped for results (Figure 3.17). For completion, these frustrated attempts are described in Chapter 3, section 3.11.

Having much more experience in making, handling and characterising the DBLα domains, Dr. Matt Higgins generously suggested the domain boundaries he had selected when producing his versions of DBLα. He also generously provided some of his recombinantly prepared DBLα(M) (Figure 5.5C) for SPR experimental trials. Subsequently he provided further recombnant proteins containing DBLα domains (through his collaboration with Alex Rowe's group) including NTS-DBL and NTS-DBLα-CIDR (Figure 5.5D). Finally, Rh4.9, the binding domain of the merozoite surface

protein used for erythrocyte invasion through the sialic acid-independent pathway, was provided by Professor Alan Cowman and Dr Waihong Tham (Figure 5. 5D).

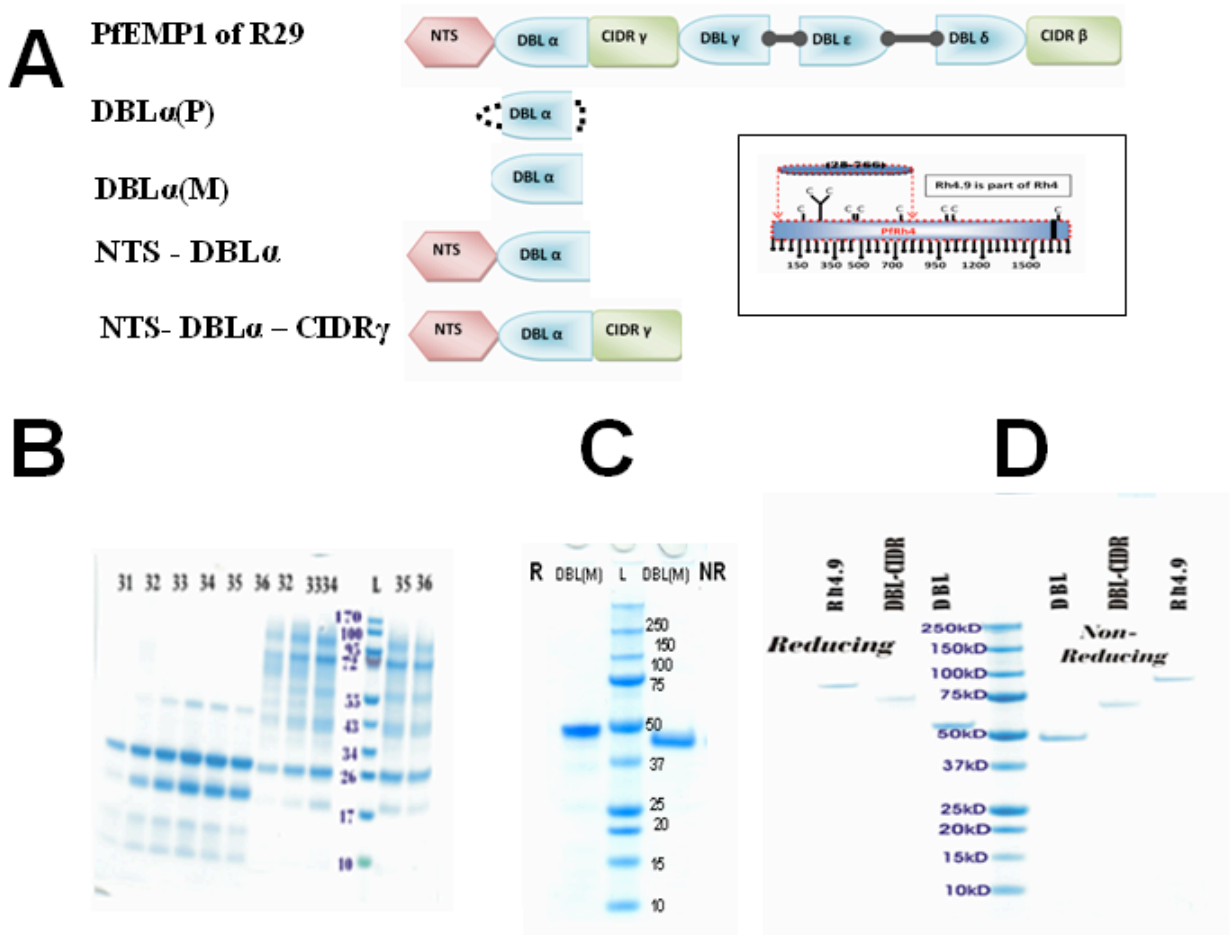



***Figure 5.5 Malaria related proteins.***

(A) Domains of PfEMP1 used in this study. NTS is N-terminal segment, DBL means Duffy-Like binding domain and CIDR is Cysteine-rich interdomain region. (M) and (P) are letters designated to the different DBLs with respect to those who produced them. DBLα(M) was produced by Dr. Matt Higgins and hence the letter (M) whiles DBLα(P) was produced by Patience Tetteh-Quarcoo, hence the letter (P). (A-insert) is representation of PfRH4. (Modified from Tham *et al*., 2009). (B) DBLα(P).The sequence of this domain has odd number of cysteines. (C) DBLα(M). Has even number of cysteine. R and NR represents reducing and non-reducing conditions respectively. (D) DBLs with NTS and Rh4.9. Left side of the marker is reducing condition while the right side of the marker is non reducing condition . (From Left) Lane 1- Rh4.9, Lane 2 – NTS-DBLα-CIDR and Lane 3- DBLα. (From right) Lane 1- Rh4.9, Lane 2 – NTS-DBLα-CIDR and Lane 3- DBLα.

# CHAPTER 5 BIOLOGICAL AND FUNCTIONAL STUDIES

## 5.6 Surface plasmon resonance-based affinity measurements I (early experiments)

Surface plasmon resonance (SPR) affords the opportunity to observe protein-protein interactions in real time and to measure rate and equilibrium constants in a quantitative manner. It thereby complements results obtained from cell-based studies such as the rosette-disruption and invasion-inhibition assays from which the specific protein-protein interactions and their strengths can only be inferred. SPR thus provides a direct and accurate means of assessing whether sequence variations in CR1 influence binding affinities for major ligands.

It was crucial to identify the best conditions under which to perform SPR measurements, particular since there existed little precedent for SPR-derived investigations of the proteins in the current study. While there was experience (in the Barlow group) of SPR-based investigations into the interaction of C3 fragments with complement regulators, interactions with putative binding domains from the parasite-encoded proteins, DBLα and Rh4.9, represented new territory. The type of buffer to be used, the pH of the buffer and the type of sensor chip had to be optimized while at the same time it was highly desirable to perform the studies under physiologically relevant conditions. As discussed under Materials and Methods in Chapter 2, several different sensor chips were available for assessing various types of interaction. The CM5 and the C1 sensor chips had been used successfully to measure the affinity of the interaction between C3b immobilised by amine-coupling, and fragments (or full-length versions) of FH and CR1 (Schmidt *et al.,* 2008*,* Mallin, 2003). A major objective of the current study

was to compare binding affinities, for both host proteins and parasite ligands, between the Knops blood-group polymorphic forms of CR1. It was decided that using similar sensor chips for both host and parasite proteins would provide the best basis for comparison. The strategy adopted was to further optimize conditions previously developed for measuring the affinities of C3b for various FH fragments.

| **Description** | **Chip Type** | **Flow Cell 1** | **Flow Cell 2** | **Flow Cell 3** | **Flow Cell 4** |
|---|---|---|---|---|---|
| Chip 1 | CM 5 | Blank (0 RU) | C3b 1510 RU | DBLα(P) 3023 RU | C3b 2992 RU |
| Chip 2 | CM 5 | Blank (0 RU) | C3b 1500 RU | DBLα(M) 3000 RU | C3b 3000 RU |
| Chip 3 | C1 | Blank (0 RU) | C3b 500 (RU) | C3b 750 (RU) | DBL α(M) 750 (RU) |
| **Chips for sections described as " SPR measurement I"** □ | | | | | |
| | | | | | |
| **Description** | **Chip Type** | **Flow Cell 1** | **Flow Cell 2** | **Flow Cell 3** | **Flow Cell 4** |
| Complement Chip | CM 5 | Blank (0 RU) | C3b (1620 RU) | C4b (1518 RU) | C1Q (2125 RU) |
| Malaria Chip | CM 5 | Blank (0 RU) | NTS-DBL-CIDR (1500 RU) | Rh4.9 (408 RU) | Rh4.9 (1476 RU) |
| **Chips for sections described "SPR measurement II "** | | | | | |

*Table 5.1 Summary of Chips used for SPR-based studies*

### 5.6.1 Binding of CR1 site 2 to DBLα (versus binding to C3b) on CM5 and C1 chips

A buffer with a pH similar to physiological conditions ("HBS-EP+" buffer - see Materials and Methods for details) was selected as a starting point. Proteins were dialysed into this buffer prior to SPR. Utility of the CM5 and C1 chips were assessed. Figure 5.6

shows the results of some of the trial processes that were carried out to establish an optimised protocol suitable for accurate testing of the polymorphic forms of CR1 15-25.

### 5.6.2 Using a CM5 chip

The first SPR trial experiment was performed using the constructs that tested positive in the rosette-disruption assay (*i.e.* CR1 17, CR1 15-17 and CR1 10-11). A CM5 chip was used with blank (0 response units (RU)) coating on flow cell (FC) 1 and 3023 RU of DBLα(P) immobilised, by amine coupling (see Materials and Methods for details), on FC 3. Flow cells 2 and 4 were coated (amine-coupled) with 1510 RU and 2992 RU of C3b respectively (Table 5.1). Note that DBLα(P) refers to the version produced by the author (see Chapter 3, section 3.11) and has an odd number of cysteine residues and that it appeared to be degraded by SDS-PAGE under reducing conditions but was intact (*i.e.* held together by disulfides) under non-reducing conditions (see Fig. 5.5B) Known binders of C3b such as FH CCPs 19-20 (FH 19-20) and FH CCPs 1-4 (FH 1-4) were used as positive controls; they were flowed over the chip and the responses produced were compared to those obtained with similar concentrations of the CR1 fragments. The HBS-EP+ buffer served to provide a baseline control.

Two concentrations of proteins (2 μM and 20 μM) were flowed over the four channels of the CM5 chip. The responses ("sensorgrams") obtained are displayed in Figures 4.6A and B. As expected, CR1 15-17 binds well to C3b (on FCs 2 and 4) as does FH 19-20 and FH 1-4. The single module construct CR1 17 does not bind in a detectable fashion to C3b, nor does the double-module construct CR1 10-11 that is identical to CR1 17-18. These results are entirely consistent with previous work

suggesting that the functional site 2 of CR1 consists of a minimum of three modules at the N-terminus of LHR-B or LHR-C, and that neither the third module alone nor a conjunction of third and fourth modules (as in CR1 10-11), is sufficient to achieve measurable affinity. When the CR1 15-17 – previously implicated in rosetting – was flowed over FC 3 (bearing 3023 RU of immobilised DBLα(P)), only a very interaction was observed that is almost certainly not physiologically meaningful (see Fig. 5.6B). Strangely, though, FH 19-20 produces the largest response of all the proteins tested (see Fig. 5.6B). While the interaction is not strong and has a very rapid off-rate (note the $K_D$ was not measured), the size (and profile) of response is not much smaller than the response observed when an equivalent concentration of FH 19-20 was flowed over its well-established ligand C3b (Fig. 5.6B) ($K_D$ = ~4 μM). There has not, to date, been any indication of involvement of factor H in rosette-disruption so this observation remains an enigma for the time being.

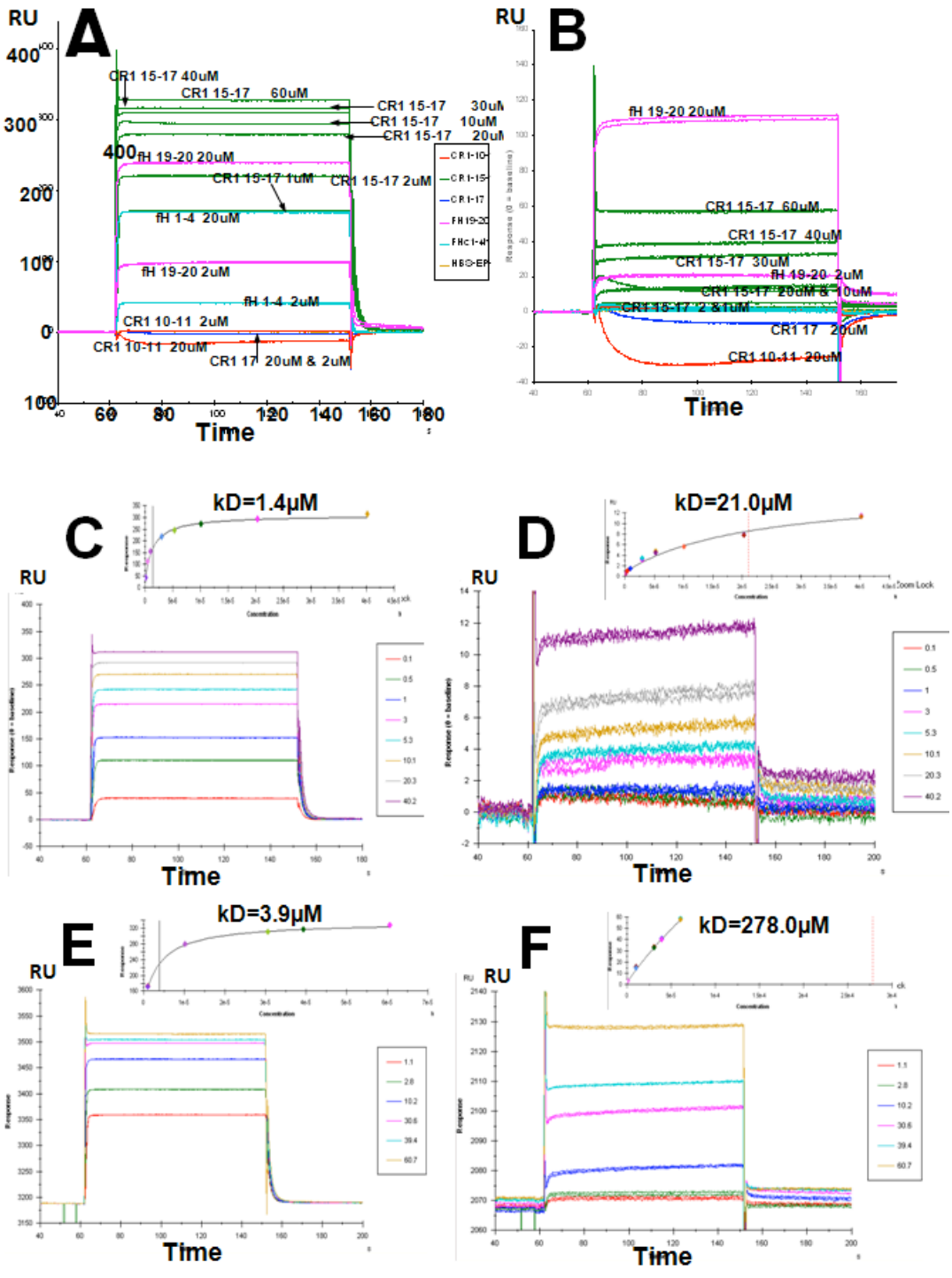



***Figure 5.6 Binding of Site 2 (and control proteins from FH) to ligands bound to a CM5 Chip***

***(A)*** Overlaid sensorgrams obtained for multiple potential ligands flowed over the C3b-coated CM5 chip surface (2992 RU on FC 4 of Chip 1, see Table 5.1); the identities of constructs

(colour-coded) and their concentrations are shown. All data (in this Figure) are baseline (HBS-EP+ buffer)-subtracted. (***B)*** As in panel (A) but in this case the chip surface (FC 3) was coated with 3023 RU of DBLα(P) . (***C)*** A concentration series (up to 40.2 μM ) of CR1 15-17 was flowed over 3000 RU C3b-coated FC 2 to generate the response versus concentration curve (inset) that was used to estimate a $K_D$ for the CR1 15-17:C3b interaction. (***D)*** As in panel (C) except the concentration series of CR1 15-17 was flowed over FC 3 that carried 3000 RU of DBL-α(P). The $K_D$ could not be calculated accurately (inset) due to failure to reach saturation. (***E)*** As in panel (C) except the CR1 15-17 samples were flowed over the FC (4) (of Chip 2 in Table 5.1) bearing 3000 RU of immobilised C3b. (***F)*** CR1 15-17 series of concentrations (up to 60.7 μM) flowed over the 3000 RU of the DBL-α(M)–coated surface (FC 3 of Chip 2 in Table 5.1) to establish that there was only very weak direct interaction between CR1 15-17 and the DBLα - $K_D$ could not be calculated (inset) due to failure to reach saturation (even at 60.7 μM).

To quantify the CR1 functional site 2:C3b interaction, a series of concentrations of CR1 15-17 were flowed over FC 2 and FC 4 yielding (see Fig. 5.6C and E) $K_D$ values of around 2-4 μM. A similar experiment performed on FC 3 (bearing 3000 RU of DBLα(P) yielded very much smaller responses and a dubious (given weak responses and failure to reach saturation) $K_D$ value of about 20 μM.

Thus these studied very clearly demonstrated that functional site 2 of CR1 interacts with C3b very much more strongly than with DBLα(P) and indeed the latter interaction was too weak to be defended as physiologically meaningful. Either CR1 interacts only very weakly with PfEMP1, or the choice of domains used in these experiments was inappropriate. (Note that no direct interaction between these proteins has yet been demonstrated in the literature). It was in fact not unlikely that DBLα(P) was improperly folded or had been cleaved at a critical site (based on previously mentioned SDS-PAGE, Fig. 5.5B and 3.17 ) and therefore it was timely that Dr. Matt Higgins (now at University of Oxford) kindly provided an alternative construct of DBLα termed DBLα(M) for the purposes of this study, which had been recombinantly produced in *E. coli* cells (using the Origami strain). This protein domain behaves appropriately on SDS-PAGE producing a dominant main band that ran slightly higher under reducing

conditions compared to non-reducing conditions (Figure 5.5C). Using similar conditions (see chip 2 of Table 5.1) to those employed for DBLα(P), the interaction between CR1 15-17 and DBLα(M) was assessed and the results presented in Figure 5.6F. As was observed for DBLα(P), however, the interaction between CR1 15-17 and DBLα(M) was very weak (compared to that with C3b, see Fig. 5.6E); no saturation could be obtained and a plot of RU versus DBLα(M) concentration showed no sign of a plateau implying that the $K_D$ must be in excess of ~ 200 μM. These results seemed to imply that even with a properly folded DBLα domain, any interaction with function site 2 of CR1 is at least 50-100 fold weaker than that of C3b despite previous work showing that site 2 was critical for resetting.

### 5.6.3 Using a C1 chip

Given the poor interaction observed between CR1 15-17 and two versions of DBLα immobilised on the CM5 sensor chip, it was decided to repeat the measurements but deploy a C1 chip (*i.e.* lacking the carboxymethyl groups). As on the CM5 chip, two FCs were used for C3b immobilization (500 RU and 750 RU) while a third was use for DBLα(M) immobilisation (750 RU), and finally FC 1 was used as a reference blank. In this set of experiments, the four polymorphic variants of CR1 15-25 were also tested.

Contrasting results were obtained when 1 μM samples of the CR1 fragments were flowed over these flow cells (see Figs 5.7A and B**)**. While on the C3b surface, the CR1 constructs - including the four CR1 15-25 variants - interacted equally strongly on the DBLα surface, the results were difficult to interpret. As shown in Figure 5.7A a variety of negative curves were generated for three of the variants while sCR1 produced only a small response and CR1 15-17 (consistently with results obtained on the CM5 chip)

yielded an even smaller response. The exception was CR1 15-25EG (Fig. 5.7B) that gave rise to a much greater response than sCR1. Subsequently, a concentration series of sCR1 was flowed over the DBLα(M) surface. Figure 5.7D shows the results for the sCR1: DBLα(M) interaction that may be compared to the measurement of the sCR1: C3b interaction in Figure 5.7C. Consistently with the comparison in Figure 5.7B, DBLα(M) generates much smaller responses and unfortunately it was not possible to extract a reliable $K_D$ from these measurements due to poor fitting of the response versus concentration curve (see insert in Fig. 5.7D). CR1 15-25 EG was also flowed over both the C3b- and DBLα(M)-bearing surfaces; these experiments resulted in $K_D$ estimates of 0.26 μM and 2.38 μM respectively (see Figs. 5.7 E and F). Finally, CR1 15-17 was similarly flowed over these two surfaces. The results are fully consistent with those in Figure 5.7B in that responses on the DBLα(M)-bearing surface are very small. The derived $K_D$ value (1.45 μM) for DBLα (M) must therefore be regarded as dubious.

Thus the sCR1, CR1-15-17 and all four variants of CR1 15-25 bound better to C3b on the C1 chip than to the same protein on the CM5 chip and response-versus-concentration curves for these experiments fitted well to a hyperbolic plot allowing reliable $K_D$ estimates. In the case of immobilised DBLα(M) the responses obtained with CR1 15-17 and sCR1 were smaller (with respect to C3b) and the response-versus-concentration plots did not fit as well to hyperbolic (two-state and 1:1 stoichiometry) functions such that no meaningful $K_D$s could be extracted. The result for CR1 15-25EG suggest that, although weaker than observed on the C3b surface, there is some form of interaction between DBLα and this CR1 15-25 variant. It was therefore worthwhile to re-

investigate the behaviour of this variant on the CM5 chip as well as the C1 chip used here. Hence, for subsequent experiments, focus returned to the CM5 chip.

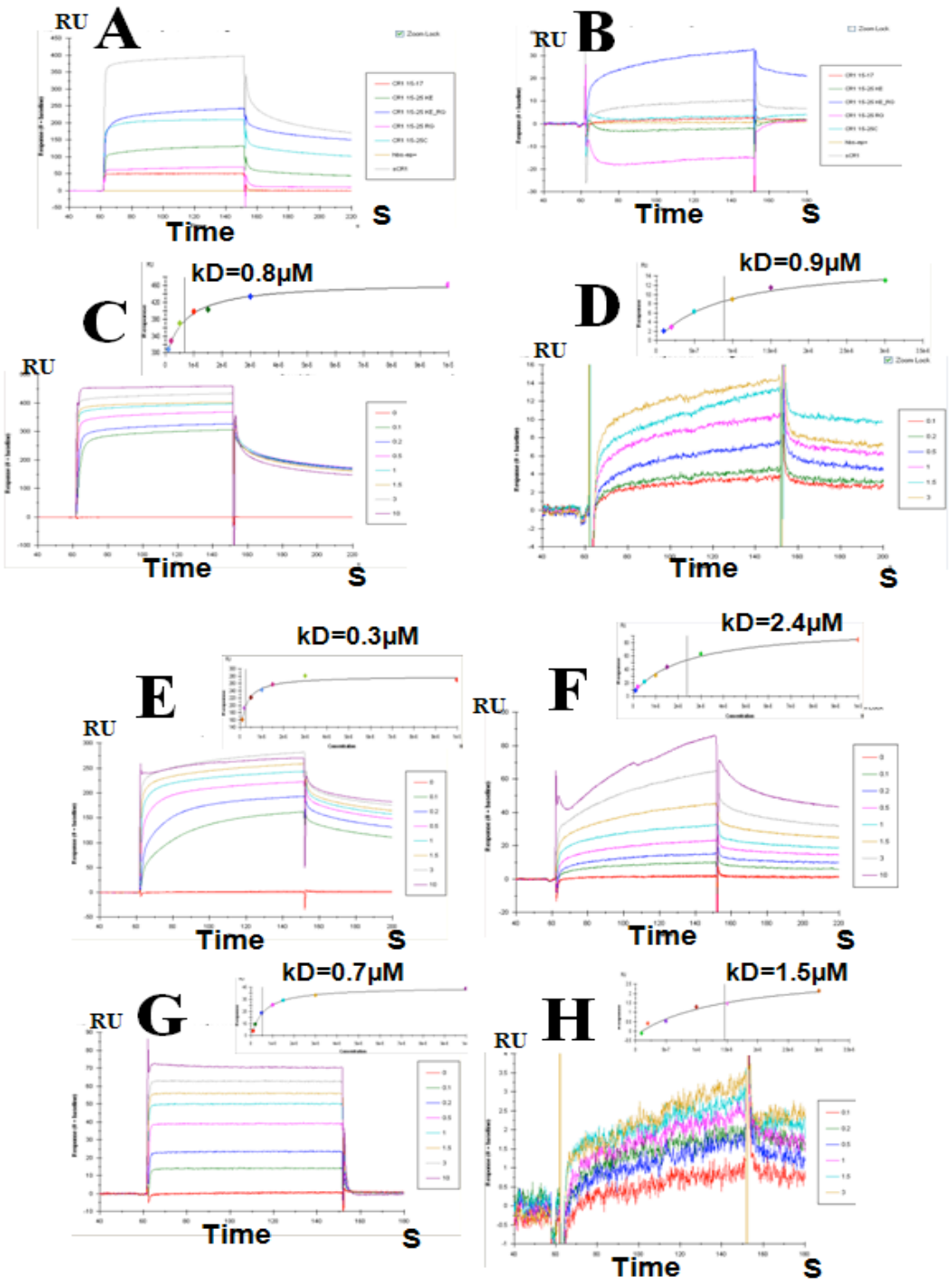



*Figure 5.7 Binding of potential ligands to C3b and DBLα(M) immobilised on a C1 sensor chip*

***(A)*** Solutions, as indicated, of 1 μM of the protein constructs were flowed over the flow cell surface bearing (750 RU) C3b. (***B)*** As in panel (A) except solutions were flowed over the flow cell surface loaded with 500 RU of DBLα(M). (***C)*** A concentration series of sCR1 flowed over the C3b-loaded surface. Insert shows response-versus-concentration data fitted to a hyperbolic function along with the calculated $K_D$ value. (**D)** As in (C) except the sCR1 solutions were flowed over the DBLα(M)-loaded surface. **(*E)*** As in (C) except a concentration series of CR1 15-25EG was used. (***F)*** As in (D) except a concentration series of CR1 15-25EG was flowed across the DBLα(M) surface**. (*G)*** and (***H)*** As in (E) & (F) except these curves are for CR1 15-17.

The set of experiments described above represented a first attempt at quantification of affinities between CR1 fragments and their potential ligand. They served as a basis for a further phase of SPR-based experiments, described below, incorporating some refinements. For example, the negative curves observed in Figure 5.7B are usually a result of strong non-specific binding of proteins to the blank/reference flow cell surface compared to that of the "working" flow cell. This was improved upon subsequently by performing a double-blank immobilization (see Materials and Methods, section 2.4.2) Also, duplicate injections were not performed in the initial SPR studies, which made it difficult to check reproducibility; a possible implication was that injected proteins might not have been completely removed from the chip surface in the wash step and might therefore accumulate during injections of a concentration series.

### 5.6.4 Further SPR-based investigations of CR1 15-25 binding to parasite proteins

At this stage of the work, the four polymorphic forms of CR1 15-25 had been successfully produced and purified, and two further DBLα constructs from PfEMP1 had kindly been gifted by Dr. Matthew Higgins. The two new DBLαs provided were NTS-DBLα and NTS–DBLα–CIDR (See Fig. 5.5D for appropriate lanes) Dr Higgins designed

these constructs on the basis that neighbouring regions might be needed for the putative binding of the DBLα domain to CR1.

The Cowman group (Melbourne) provided Rh4.9, corresponding to a recombinant version of the N-terminal domain of the parasite protein Rh4, the protein that interacts with CR1 during the sialic acid independent invasion pathway used by *P. falciparum* merozoites. A CM5 chip was chosen for the SPR experiments (see reason for choice explained in section 5.6.3 above) and FC 1 served as the blank or reference.

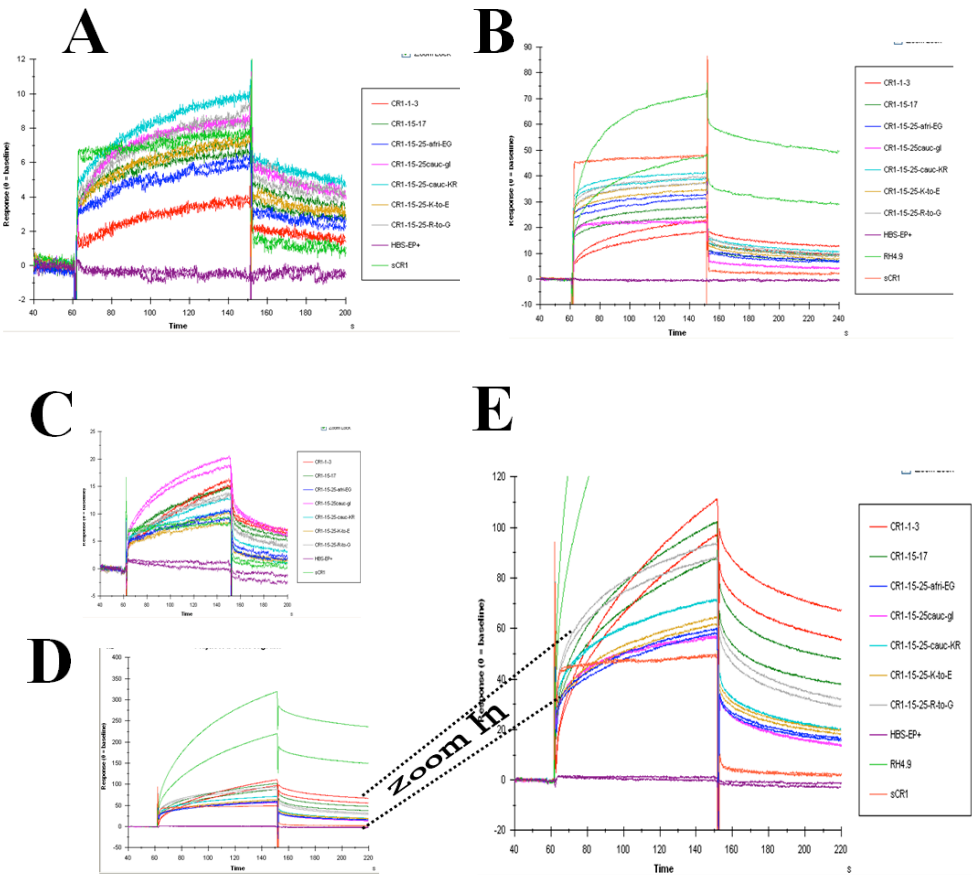



***Figure 5.8 Binding of CR1 constructs to DBL domains monitored by SPR***

***(A)*** Overlay of sensorgrams obtained by flowing (individually) 1 μM solutions of the indicated CR1 constructs over the NTS-DBLα loaded surface of FC 2. (***B)*** As in panel (A), except 10 μM solutions of CR1 constructs were flowed over the surface of FC 2. (***C)*** Solutions of 1 μM of the four CR1 15-25 polymorphic variants were flowed over the flow cell surface bearing immobilised NTS-DBLα-CIDR. (***D)*** Solutions of 10 μM of the four CR1 15-25 polymorphic variants were flowed over the flow cell (FC 4) surface bearing immobilised NTS-DBL-CIDR . (***E)*** Zoom in on panel (D**)** to highlight weaker binders on NTS-DBL-CIDR flow cell.

The three parasite DNA-encoded recombinant proteins were amine-coupled onto the surfaces of FCs 2 to 4: NTS-DBLα on FC 2; NTS-DBLα-CIDR on FC 3 and Rh4.9 on FC 4. Panels (A) and (B) of Figure 5.8 show the overlaid sensorgrams obtained after flowing (one at a time) various CR1 constructs (at 1 μM and 10 μM) across the surface of FC 2 (NTS-DBLα). All four polymorphic forms of CR1 15-25 exhibited similar very weak (probably non-specific) interactions with immobilised NTS-DBLα as did CR1 1-3, CR1 15-17 and sCR1. Strangely though, Rh4.9 which was also flowed over the chip, seemed to interact with NTS-DBLα better than any of the CR1 fragments. Figures 4.8C, D and E illustrate the sensorgrams obtained for interactions of the same set of injected protein samples with NTS-DBL-CIDRα on the chip surface As before the CR1 constructs interact only very weakly with the surface; there are insignificant differences between the binding curves obtained for CR1 1-3 (not reported to be important for rosetting) and CR1 15-17 (the purported rosetting-critical functional site of CR1) or full-length sCR1, while all four variants of CR1 15-25 bind equally poorly. It is thus very difficult to attach any significance to these sensorgrams in terms of physiologically relevant protein:protein interactions between CR1 and the PfEMP1 constructs. As was observed with NTS-DBLα, Rh4.9 interacts with immobilised NTS-DBL-CIDRα much better than do the CR1 constructs. This intriguing observation, however, is not backed up by any biological experiments to date and was not further investigated due to time constraints.

## 5.7 Surface plasmon resonance-based affinity measurements II (optimised experiments)

While the SPR results obtained up to this point had been very promising, they were not regarded as publishable in quality and in any case it was important that key observations be replicated. For the final set of SPR-based experiments, designed to be the definitive ones, two CM5 chips were loaded with proteins. Commercially available C3b, C4b and C1q were immobilised on FCs 2 to 4 of the 'complement chip' while NTS-DBLa-CIDR and Rh4.9 (two different loadings) were coupled to FCs 2 to 4 of the 'malaria chip' (summarised in Table 5.1**)**.

### 5.7.1 CR1 proteins flown over C3b (Complement Chip)

Figure 5.9A shows the sensorgrams obtained from flowing 10 μM solutions of the CR1 constructs over FC 2 of the complement chip loaded with 1500 RU of C3b. The quality of these data was regarded as very high ("publishable") on the basis of the excellent reproducibility of duplicate injections and the lack of negative data points, glitches or other artefacts. Moreover binding of the positive controls sCR1 and CR1 15-17, and the negative control CR1 10-11, are all as expected. Finally CR1 1-3 (*i.e.* functional site 1) binds to C3b significantly more weakly than CR1 15-17, which is consistent with several studies published by the Atkinson laboratory.

The four polymorphic forms of CR1 15-25 all had very similar binding curves and it may be inferred that all bind equally well to amine-coupled C3b. A form of CR1 15-25KR that runs as a larger entity in the size-exclusion chromatography (peak A in Figs. 3.12 and 4.13) binds less well than its smaller, non-glycosylated counterpart.

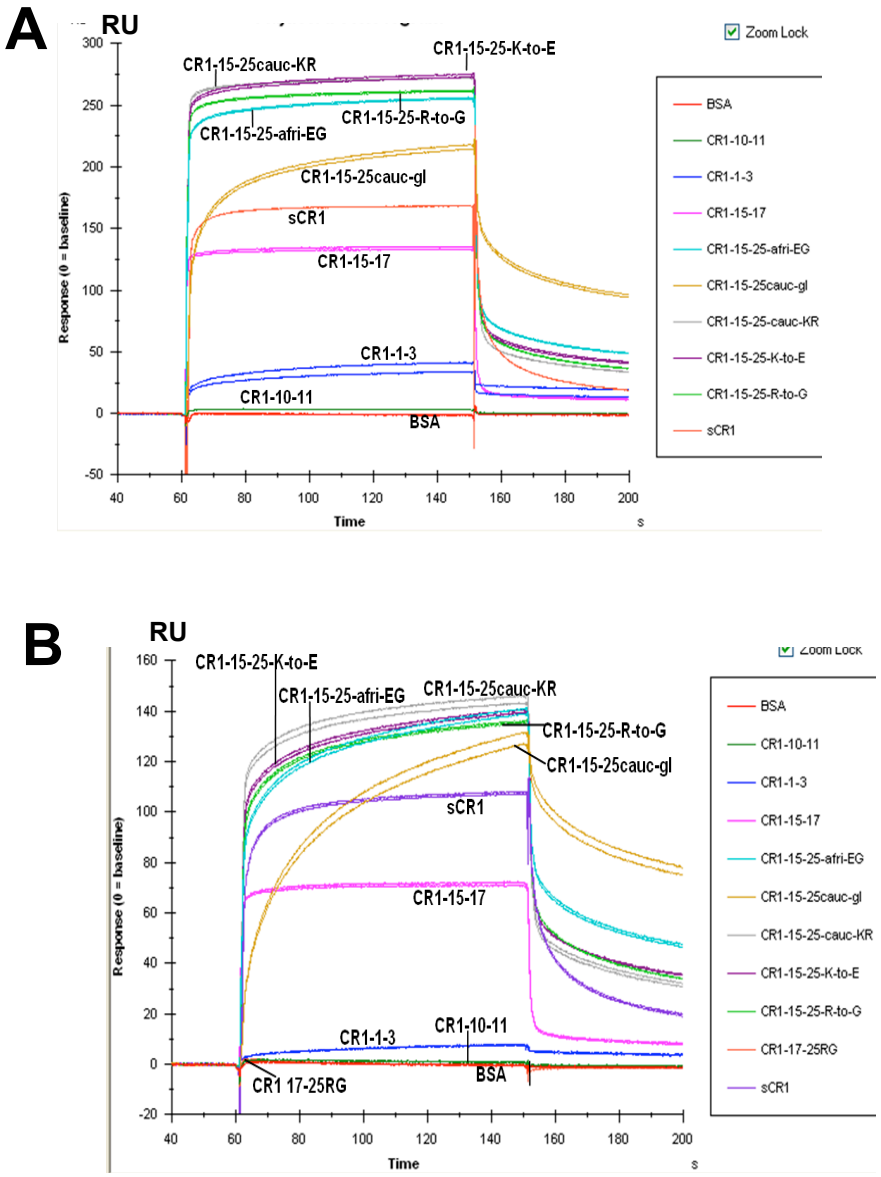



***Figure 5.9 "Definitive" study by SPR of binding to C3b by a set of CR1 constructs***

***(A)*** Solutions of 10 μM CR1 constructs flowed over 1500 RUs of immobilised C3b on the surface of FC 2 of the "complement chip". ***(B)*** As in panel (A) except 1-μM solutions of CR1 constructs were flowed over the same C3b-loaded surface.

Similar observations (see Fig. 5.9B**)** were made when 1-μM solutions of the CR1 constructs were flowed over the C3b-loaded chip surface. In this set of experiments, CR1 17-25KG was included as an additional negative control, confirming that the binding displayed by the CR1 15-25 constructs resides entirely in the N-terminal modules.

### 5.7.2 Series of CR1 concentrations on C3b Flow Cell

To obtain $K_D$ values for complexes formed between C3b and the CR1 15-25 constructs, a concentration series (0-10 μM) of each variant was flowed over C3b immobilised on FC 2 of the complement chip (see Figs. 5.10A-D). These data were of good quality and allowed plots of response-versus-concentration suitable for fitting and calculation of the $K_D$ values summarised in Table 5.2. The four values obtained are similar - 1.5 μM for CR1-15-25KR, 2.4 μM for CR1 15-25ER; 2.1 μM for CR1 15-25KG; and 1.8 μM for CR1 15-25EG - and there is no trend of affinities in the series: KR to (ER, KG) to EG; implying that there are no significant differences in affinities for C3b arising from these sequence variations.

For comparison, a concentration series of sCR1was deployed (using the same FC as above) to provide a $K_D$ value of 1.3 μM for the sCR1:C3b complex (Fig. 5.10G). Using the same approach, a $K_D$ value of 2.2 μM was obtained for the CR1 15-17:C3b complex (Fig. 5.10E). A higher concentration range (0 – 62 μM) was chosen with which to characterise the complex between CR1 1-3 and C3b, based on the results of Figure 5.9 and the literature. Even so, saturation was not achieved and hence the $K_D$ value obtained

at ~80 µM (Fig. 5.10F) must be regarded as an estimate only. These values are summarised, along with those for the CR1 15-25 variants, in Table 5.2.

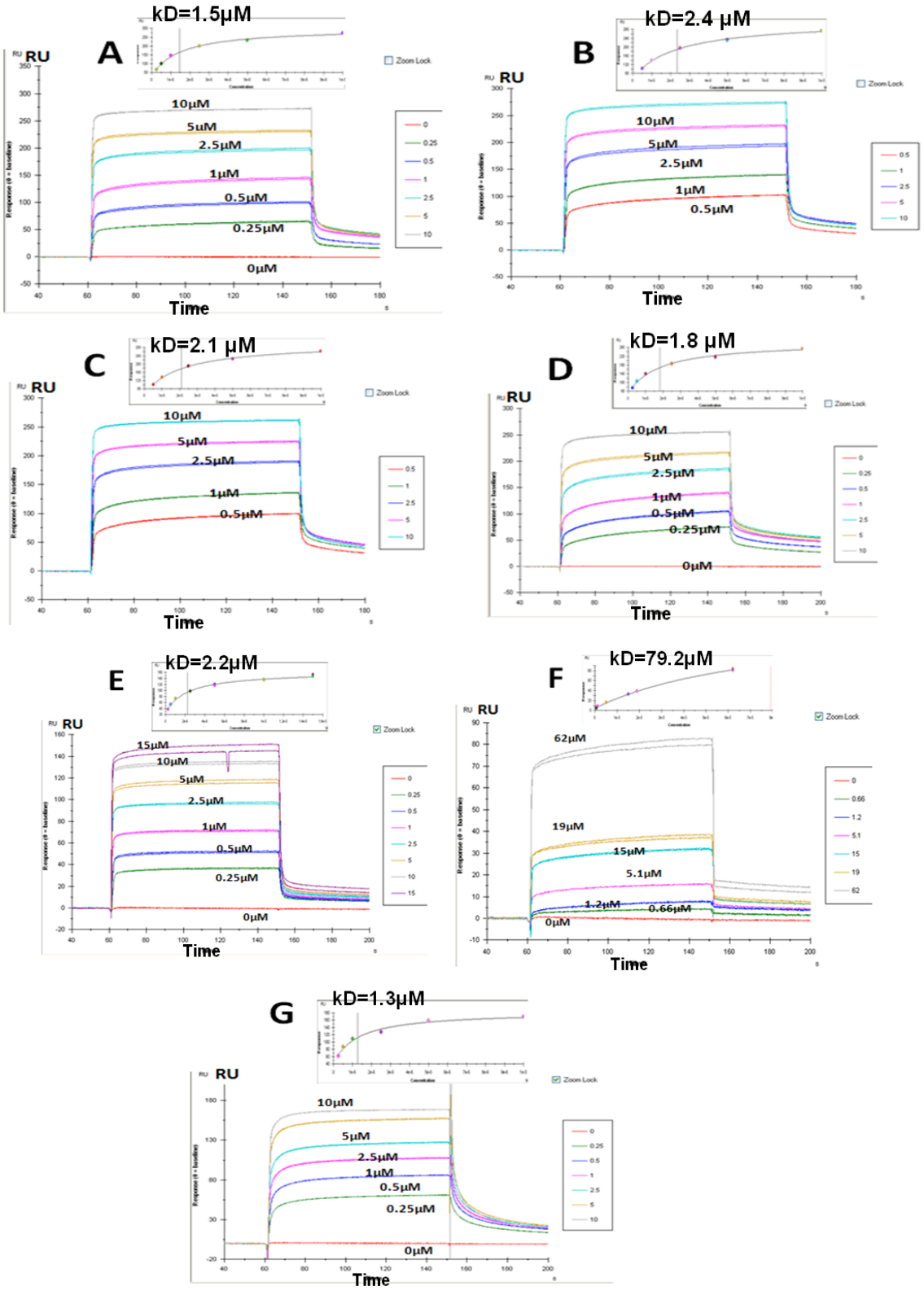


***Figure 5.10 Estimation of $K_D$ values for CR1 constructs binding to immobilised C3b***

***(A)*** Overlay of binding curves for a concentration series (0 – 10 μM) of CR1 15-25KR flowed over immobilised C3b (inset – response-versus-concentration plot fitted to obtain a $K_D$ value). (***B)*** As in panel (A) but for CR1 15-25ER. (***C)*** As in panel (A) but for CR1 15-25KG. (***D)*** As in panel (A) but for CR1 15-25EG. (***E)*** As in panel (A) but for CR1 15-17. (***F)*** As in panel (A) except a concentration series of 0 – 60 μM of CR1 1-3 was deployed. (***G)*** As in panel (A) but for sCR1.

### 5.7.3 CR1 constructs injected onto the C4b-loaded surface

As noted in Table 5.1, FC 3 of the "complement chip" was loaded (amine coupling) with 1500 RU of complement protein C4b (from Complement Technology). Figure 5.11 shows a series of SPR-based experiments on C4b, analogous to those (for C3b) shown in Figure 5.9.

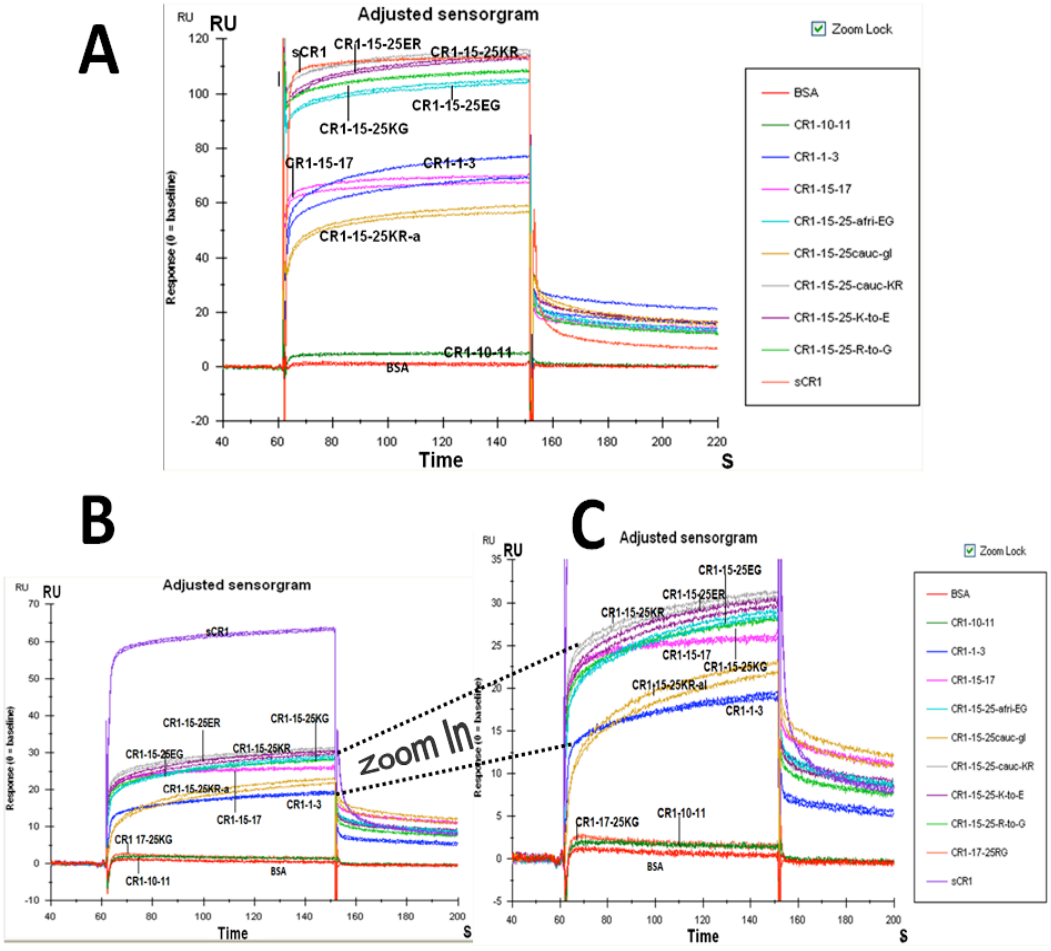



***Figure 5.11 "Definitive" study by SPR of binding to C4b by a set of CR1 constructs***

***(A)*** Sensogram-overlay for 10-μM solutions of the various CR1 constructs flowed over 1500 RUs of C4b immobilised on FC 3 of the complement chip. (***B)*** As in panel (A) but deploying 1-μM solutions of the CR1 constructs. (***C)*** An expanded region of panel (B) to highlight the weaker binders.

The positive control, sCR1 bound the most tightly out of this series as expected, while CR1 10-11 and CR1 17-25 served as good negative controls. CR1 1-3 bound less well than CR1 15-17 and neither bound to C4b as well as they did to C3b - this is largely consistent with the literature that reports CR1 functional site 2 to be the main site for C3b binding. Not surprisingly, the variants of CR1 15-25 also bound less well to C4b than they did to C3b. Importantly from the perspective of the current study, all four produced

very similar binding curves, implying there is no difference amongst these variants in their affinity for C4b.

### 5.7.4 Determination of $K_D$ values for complexes of CR1 constructs with C4b

The same concentration-series experiments that were used to determine $K_D$ values for C3b also yielded equivalent data (Fig. 5.12) for C4b since both C3b and C4b were loaded onto different flow cells of the same CM5 "complement chip". The resultant $K_D$ values are summarised in Table 5.2. For the CR1 15-25 variants the values are all comparable: 6.9 μM for CR1 15-25KR; 8.1 μM for CR1 15-25ER, 6.9 μM for CR1 15-25KG; 8.0 μM for CR1 15-25EG. These longer constructs bind slightly better to C4b than CR1 15-17 ($K_D$ = 11.5 μM) and significantly better than CR1 1-3 (21.4 μM). The strongest binder to C4b was, as expected, sCR1 with its multiple interaction sites ($K_D$ =1.6 μM).

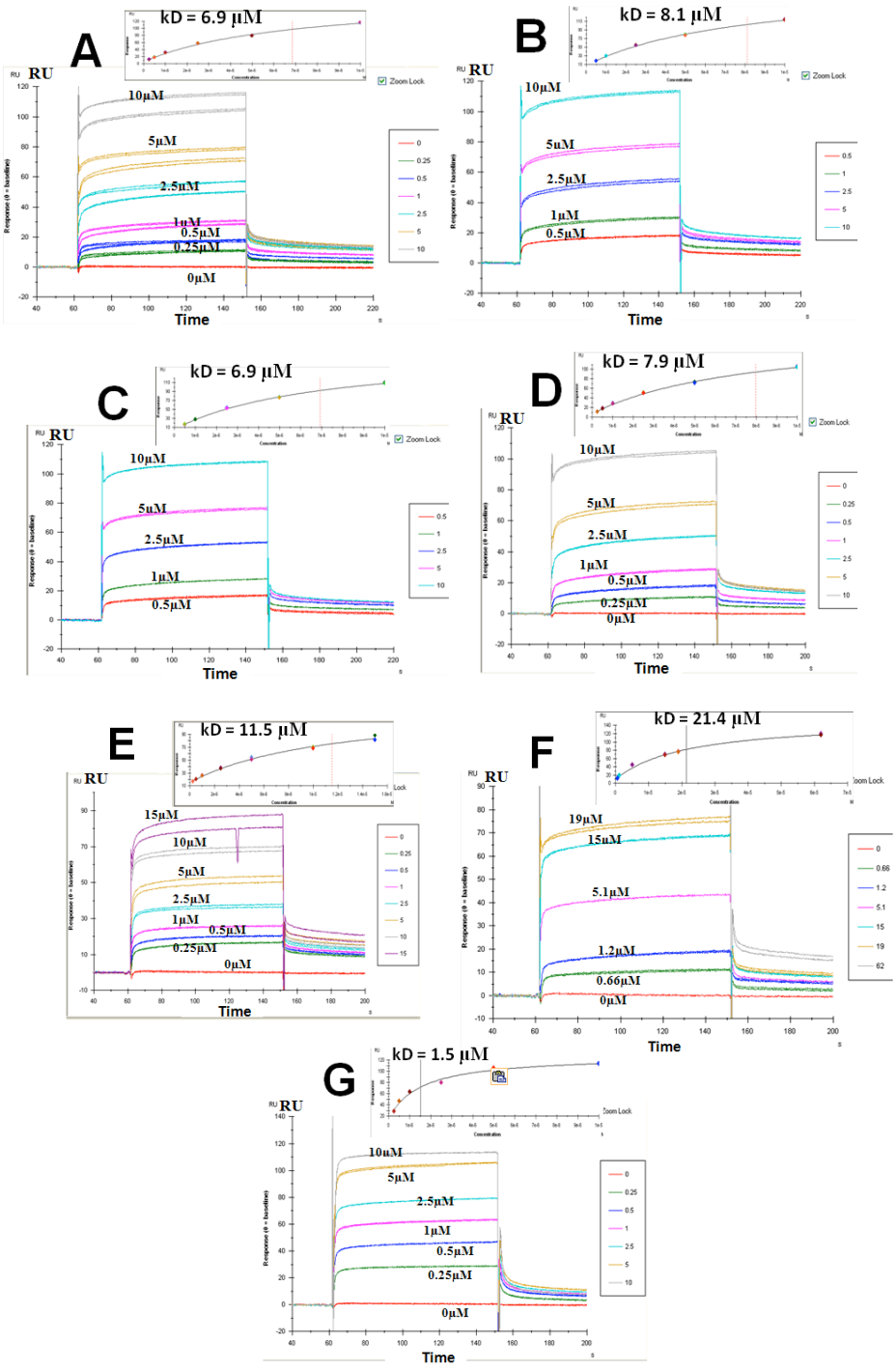



***Figure 5.12 Estimation of $K_D$ values for CR1 constructs binding to immobilised C4b***

***(A)*** Overlay of binding curves for a concentration series (0 – 10 µM) of CR1 15-25KR flowed over immobilised C4b (inset – response-versus-concentration plot fitted to obtain a $K_D$ value). ***(B)*** As in panel A) but for CR1 15-25ER. ***(C)*** As in panel (A) but for CR1 15-25KG. ***(D)*** As in panel (A) but for CR1 15-25EG. ***(E)*** As in panel (A) but for CR1 15-17. ***(F)*** As in panel (A) except a concentration series of 0 – 60 µM of CR1 1-3 was deployed. ***(G)*** As in panel (A) but for sCR1.

### 5.7.5 CR1 constructs injected onto a C1q-loaded surface

Since C1q had been loaded onto FC 4 of the complement chip, the experiments outlined in the previous sections also provided data (see Fig. 5.13) that could be used to estimate the affinities of the CR1 constructs for immobilised C1q. In general interactions appeared less strong than those with C3b (or C4b) as judged by the smaller responses elicited.

From the overlaid sensograms in Figures 4.13A and B, it appears that full-length sCR1 surprizingly bound less well to C1q than most of the CR1 fragments and it was also unexpected that CR1 15-17 (although the curves for duplicate injections do not overlay well) produces a larger response than the CR1 15-25 variants given that the C1q-binding site was previously localised (Klickstein et al. 1997) to LHR-D (*i.e.* CCPs 22-28). The double module CR1 10-11 was a good negative control but CR1 1-3 appeared to interact weakly with C1q, which was not anticipated and may represent non-specific interactions. None of the variants interacted strongly with C1q and there was only minor differences between the strengths of the responses. Further information was gleaned from the results of the concentration-series experiments (discussed below).

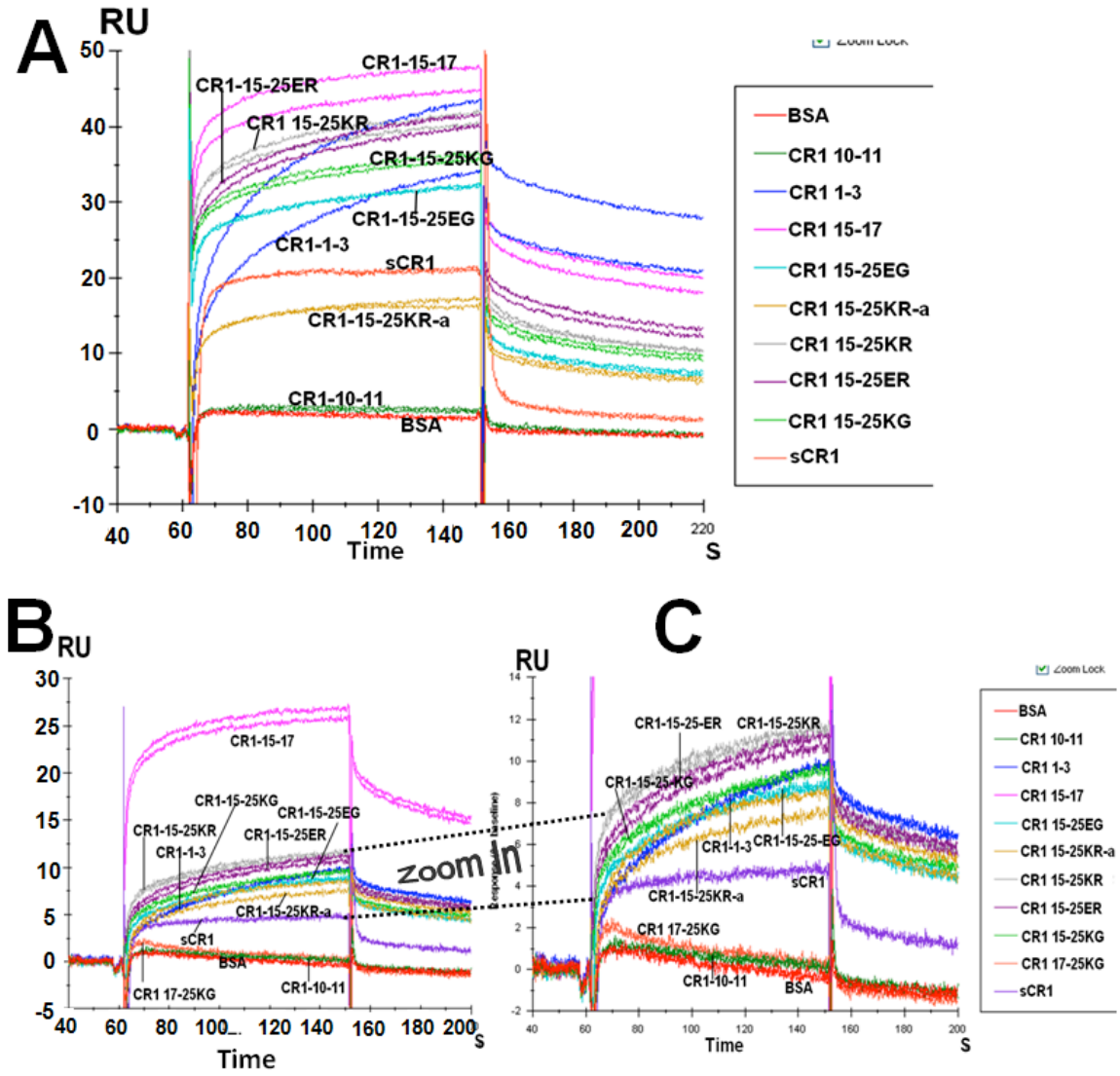



***Figure 5.13 "Definitive" study by SPR of binding to C1q by a set of CR1 constructs***

***(A)*** Sensogram-overlay for 10-μM solutions of the various CR1 constructs flowed over 2125 RUs of C1q immobilised on FC 4 of the complement chip. (***B)*** As in panel (A) but deploying 1-μM solutions of the CR1 constructs. (***C)*** An expanded region of panel (B) to highlight the weaker binders.

### 5.7.6 Determination of $K_D$ values for complexes of CR1 constructs with C1q

The $K_D$ values extracted (from the same experiments used to study interactions with C3b and C4b that were loaded onto different flow cells of the same chip) for complexes of CR1 and C1q (Fig. 5.14) are summarized in Table 5.2. The positive control, sCR1, was the tightest binder in the series tested with $K_D$ = 6.4 μM (notwithstanding the small response obtained when this was injected onto the C1q-loaded surface). Despite the lower responses obtained when CR1 15-25 constructs were flowed over C1q compared to C4b, (compare Figs 4.13 and 4.11) the calculated dissociation constants for C1q come out only slightly weaker, and with little evidence of differences between the four variants: 8.8 μM for CR1 15-25KR; 9.8 μM for CR1 CR1 15-25ER; 8.3 μM for CR1 15-25KG; 7.6 μM for CR1 15-25EG. Consistent with the relatively large response obtained when CR1 15-17 was injected (Fig. 5.13) a $K_D$ value (11 μM) comparable to those for the CR1 15-25 constructs was obtained. In the case of CR1 1-3, however, saturation was not reached, and a dubious value of 200 μM was obtained. Yet again, the size of the response obtained for injection of 10 μM CR1 1-3 (see Fig. 5.13) was out of proportion to its very weak interaction as revealed by these calculations of $K_D$.

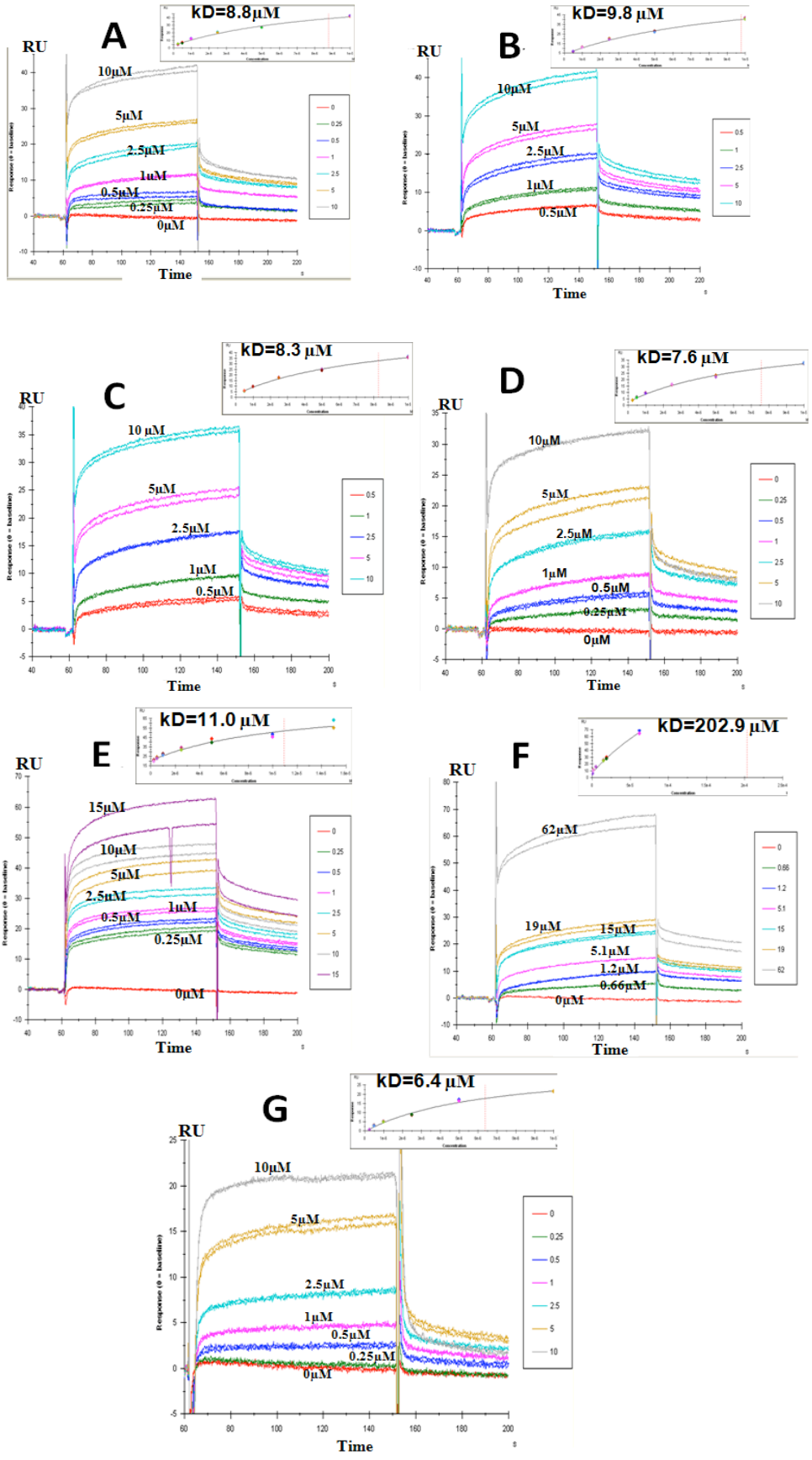

A
kD=8.8µM
RU
10µM
5µM
2.5µM
1uM
0.5µM
0.25µM
0µM
Time
B
kD=9.8 µM
RU
10µM
5µM
2.5µM
1µM
0.5µM
Time
C
kD=8.3 µM
RU
10 µM
5µM
2.5µM
1µM
0.5µM
Time
D
kD=7.6 µM
RU
10µM
5µM
2.5µM
1µM
0.5µM
0.25µM
0µM
Time
E
kD=11.0 µM
RU
15µM
10µM
5µM
2.5µM
1µM
0.5µM
0.25µM
0µM
Time
F
kD=202.9 µM
RU
62µM
19µM
15µM
5.1µM
1.2µM
0.66µM
0µM
Time
G
kD=6.4 µM
RU
10µM
5µM
2.5µM
1µM
0.5µM
0.25µM
0µM
Time

***Figure 5.14 Estimation of $K_D$ values for CR1 constructs binding to immobilised C1q***

***(A)*** Overlay of binding curves for a concentration series (0 – 10 μM) of CR1 15-25KR flowed over immobilised C1q (inset – response-versus-concentration plot fitted to obtain a $K_D$ value). (***B)*** As in panel A) but for CR1 15-25ER. (***C)*** As in panel (A) but for CR1 15-25KG. (***D)*** As in panel (A) but for CR1 15-25EG. (***E)*** As in panel (A) but for CR1 15-17. ***F*** As in panel (A) except a concentration series of 0 – 60 μM of CR1 1-3 was deployed. ***G*** As in panel (A) but for sCR1.

## 5.8 Malaria Chip

The 'malaria chip', as detailed in Table 4.1, had 1500 RU of NTS-DBLα-CIDR immobilised on FC 2, while on FCs 3 and 4 were loaded 408 RU and 1479 RU of Rh4.9, respectively. (FC1 served as a blank or reference). As was the case with the SPR experiments conducted on the "complement chip", duplicate injections were performed and the work was conducted at 25 °C in HBS-EP+ buffer (details in Materials and Methods).

### 5.8.1 Binding of CR1 constructs to immobilised NTS-DBLα-CIDR

Figure 5.15 shows the reference (*i.e.* FC 1)-subtracted binding curves obtained following injections of CR1 constructs (at 5 μM) over FC 2 (loaded with NTS-DBLa-CIDR). Only weak responses were recorded in all cases. There were some differences amongst the small signals obtained for the presumably weak binding to NTS-DBLa-CIDR of the four polymorphic forms of CR1 15-25. Injections of CR1 15-17 or sCR1 produced even smaller signals, only a little greater than the responses for the set of negative controls CR1 10-11, CR1 20-23 , CR1 21-22 and CR1 24-25C. Injection of CR1 1-3, however, yielded a relative large response (although poor duplicates) even though this construct had been added as a negative control.

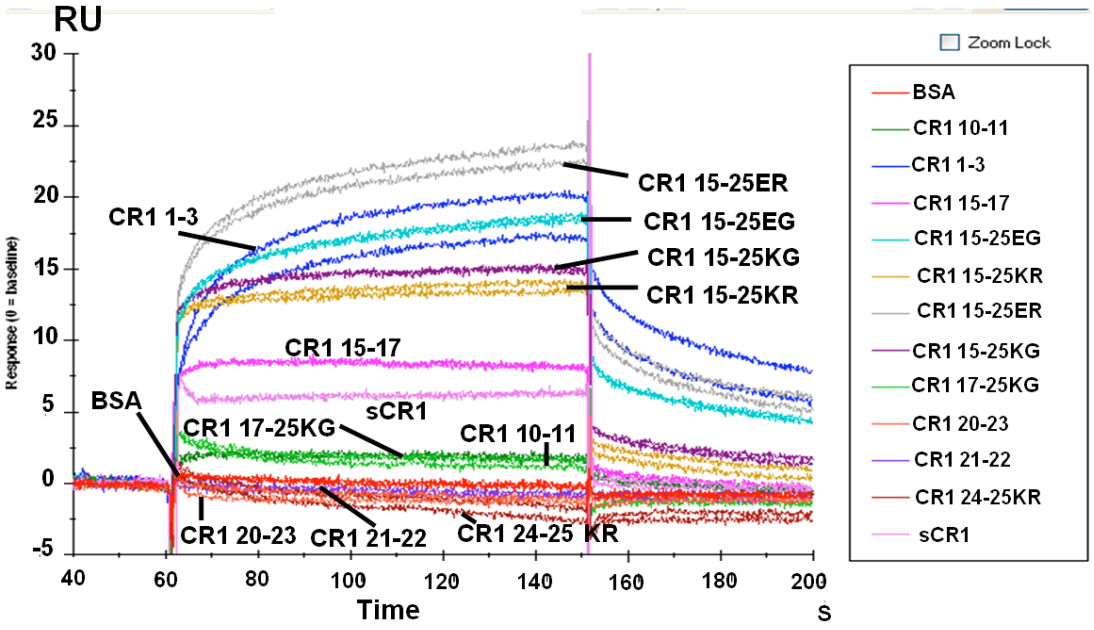



***Figure 5.15 SPR-derived data for CR1 constructs binding to surface loaded with NTS-DBLα-CIDR***

The constructs indicated were injected onto FC 2 of the malaria chip at 10 μM.

Although the binding curves suggested low interaction affinities it was decided to try to measure these by undertaking concentration-series experiments similar to those described above for C3b, C4b and C1q.

### 5.8.2 Attempted determination of $K_D$ values for CR1 constructs for DBL domain of PfEMP1

Given the similarity amongst the CR1 15-25 variants with respect to binding curves on the NTS-DBLα-CIDR, only CR1 15-25KG was chosen for $K_D$ determination. A concentration–series from 1 μM to 50 μM was injected over FC 2 of the "malaria chip", resulting in the sensograms shown in panel (A) of Figure 5.16. Increasing concentrations gave increasing responses but did not arrive at saturation, even at 50 μM. Thus a dubious

$K_D$ value of ~100 μM was calculated for the CR1 15-25KG interaction with NTS-DBLα-CIDR. The equivalent concentration series was used in the case of CR1 15-17 but again there appeared to be a linear correlation between increasing concentration and responses with no saturation at 50 μM and therefore an unreliable calculated $K_D$ value in the range of 140 μM (Fig. 5.16B). Similarly, a $K_D$ of ~100 μM was estimated for CR1 1-3 (Fig. 5.16D). For sCR1 a series of lower-concentration solutions were injected ranging from 0.5 to 10 μM (based on the expectation that full-length sCR1 would interact better than the fragments); strangely (see Fig. 5.16C), it displayed a fast on-rate and fast off-rate; saturation did not occur, although some evidence of levelling out is visible, and the $K_D$ came out at very roughly ~15 μM.

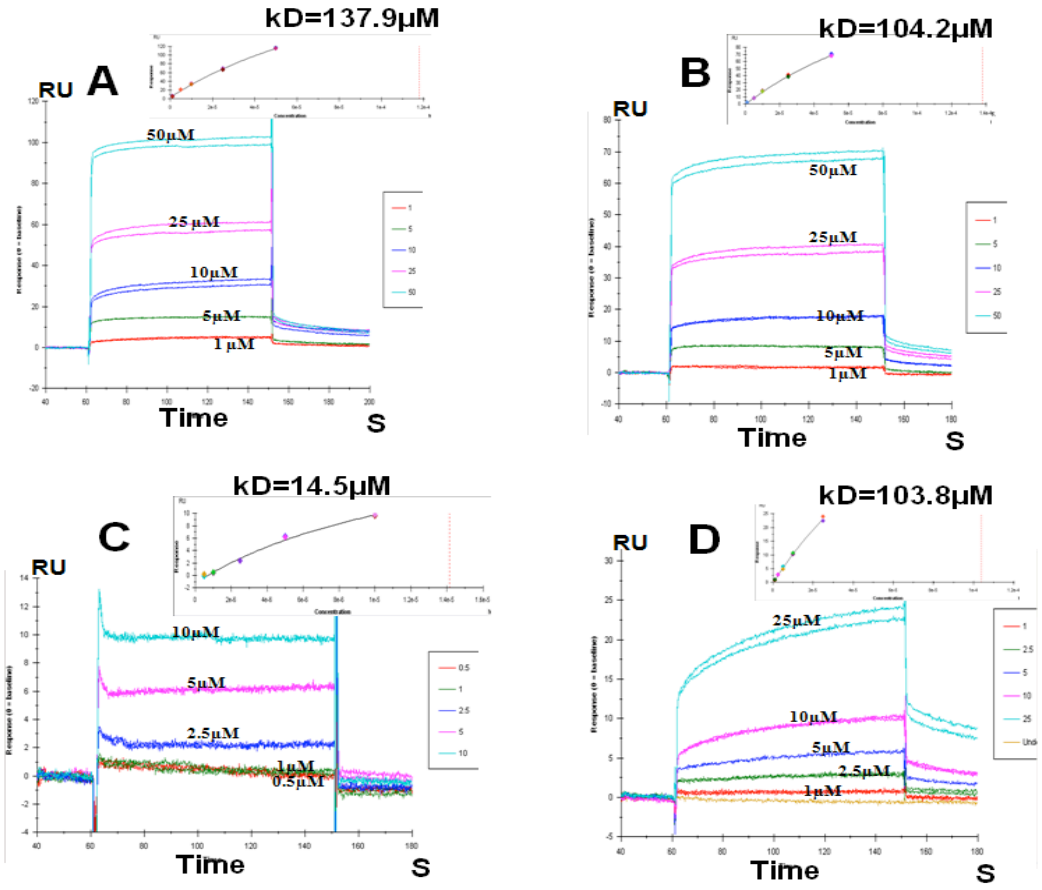



***Figure 5.16 Concentration series of CR1 constructs flowed over the NTS-DBLα-CIDR surface***

***(A)*** Overlaid sensorgrams obtained upon injection of CR1 15-25KG concentration series (0 – 50 µM) onto NTS-DBLα-CIDR that had been immobilised onto FC 2 of the complement chip (inset: response-verus-concentration curve to estimate $K_D$). ***(B)*** As in panel (A) but for CR1 15-17 ***(C)*** As in panel (A) except for a concentration series of 0 – 10 µM sCR1. ***(D)*** As in panel (A) but for CR1 1-3.

### 5.8.3 "Definitive" studies of CR1 constructs binding Rh4.9 on "malaria chip"

Figure 5.17 shows data obtained from the same sets of CR1 constructs as deployed in Figure 5.15 but in this case responses from FCs 3 and 4, loaded with Rh4.9 (see Table 5.1), are displayed.

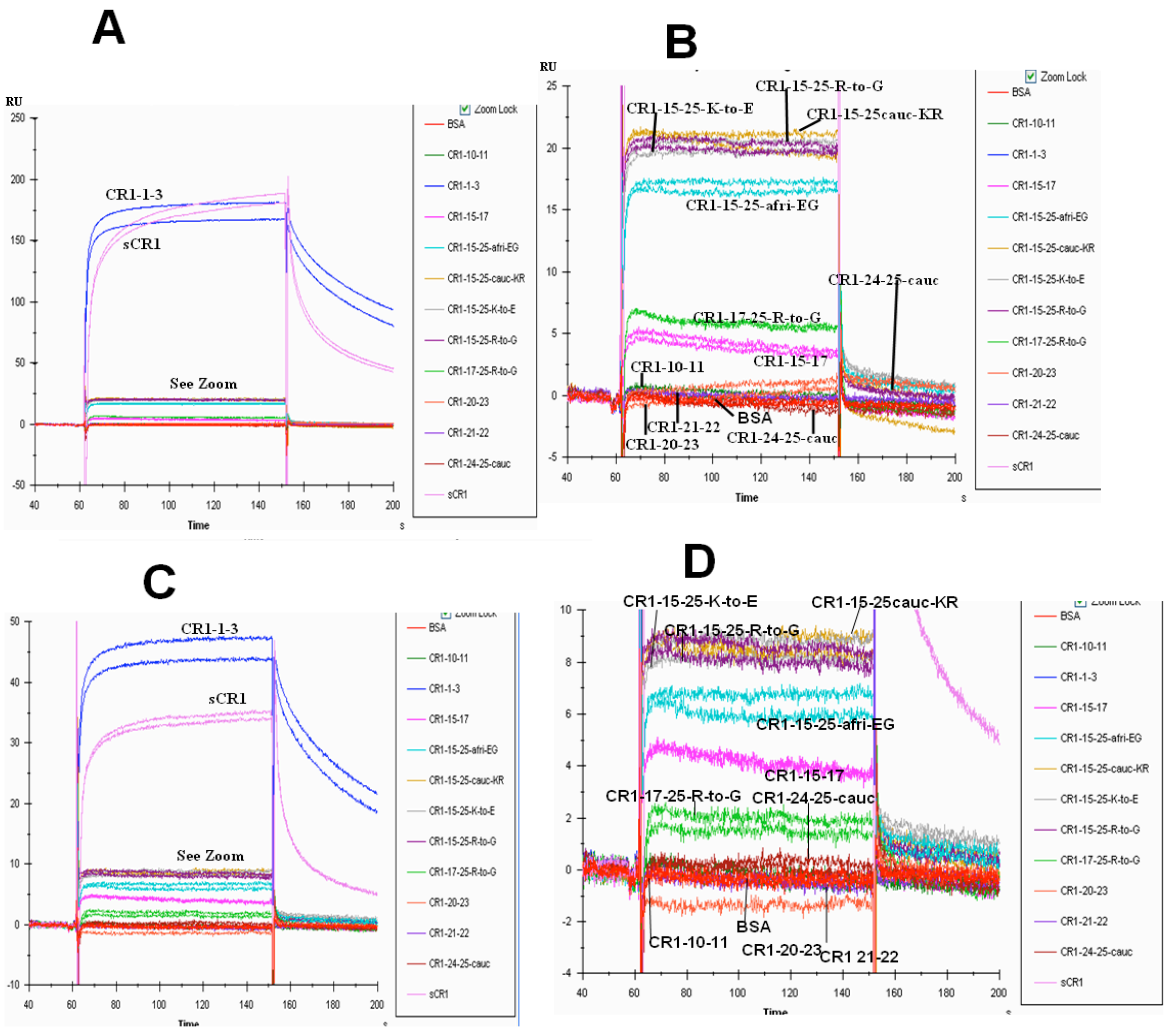



***Figure 5.17 Rh4.9 interactions with CR1***

***(A)*** sCR1 and CR1 fragments injected (separately) at 5 μM over 408 RUs of immobilised Rh4.9. (***B)*** Expansion of a part of panel (A) to highlight the data for the more weakly binding proteins. (***C)*** As in panel (A) but for 1476 RUs loaded onto FC 4. (***D)*** Expansion of a part of panel (C) to highlight the data for the more weakly binding proteins.

The quality of these SPR data was regarded as very high with strong, reproducible responses for some proteins and a lack of glitches or other artefacts. As is evident from Figures 4.17A and C, sCR1 and CR1 1-3 interact strongly with immobilised Rh4.9, while all other constructs bind only weakly (Figs. 5.17B and D). The CR1 15-25 constructs

produced slightly larger responses (similar amongst the four variants) than CR1 15-17 that in turn yielded a larger response than CR1 17-25 and the smaller constructs such as CR1 10-11, CR1 20-23 *etc.* These results pinpoint the unique CR1 functional site 1 (CR1 1-3) as the binding site for Rh4.9 and they are wholly consistent with the results of the invasion-inhibition assays. Importantly, none of the CR1 15-25 variants displayed significant affinity for Rh4.9 – thus ruling out invasion as a selective pressure on the Knops-blood group antigens. It was decided to proceed with $K_D$ measurements based on concentration-series experiments.

### 5.8.4 Determination of $K_D$ values for complexes of CR1 constructs with Rh4.9

From the data displayed in Figure 5.18, collected on FC 3 and 4, the following $K_D$ values were determined (summarised in Table 5.3; note that very similar results were obtained for FC 3 and 4, therefore duplicate $K_D$ values were not shown in the Table). For CR1 15-17 saturation was not achieved at the top end of the 0–50 μM range and a highly dubious "theoretical" $K_D$ of ~300 μM was estimated based on a near-linear plot of response versus ligand concentration; For CR1 15-25RG (chosen as a representative for all four variants that gave similar sensorgrams when injected onto the Rh4.9 surfaces) saturation was again not attained at 50 μM although some flattening of the response-versus-concentration plot was perceptible and a $K_D$ in the range of 120 μM was estimated. For sCR1 a reliable $K_D$ of 3.0 μM was calculated that was somewhat tighter than the value of 11 μM derived from the sensorgrams for CR1 1-3. These quantitative data back up the earlier statement that the primary binding site for Rh4.9 on CR1 lies in CR1 1-3 and that there is considerable discrimination in binding to Site 1 compared to the very similar (in

sequence) Site 2. That sCR1 binds more tightly than CR1 1-3, however, does suggest a contribution from other CCPs within CR1 – this and other implications of these findings, are discussed later.

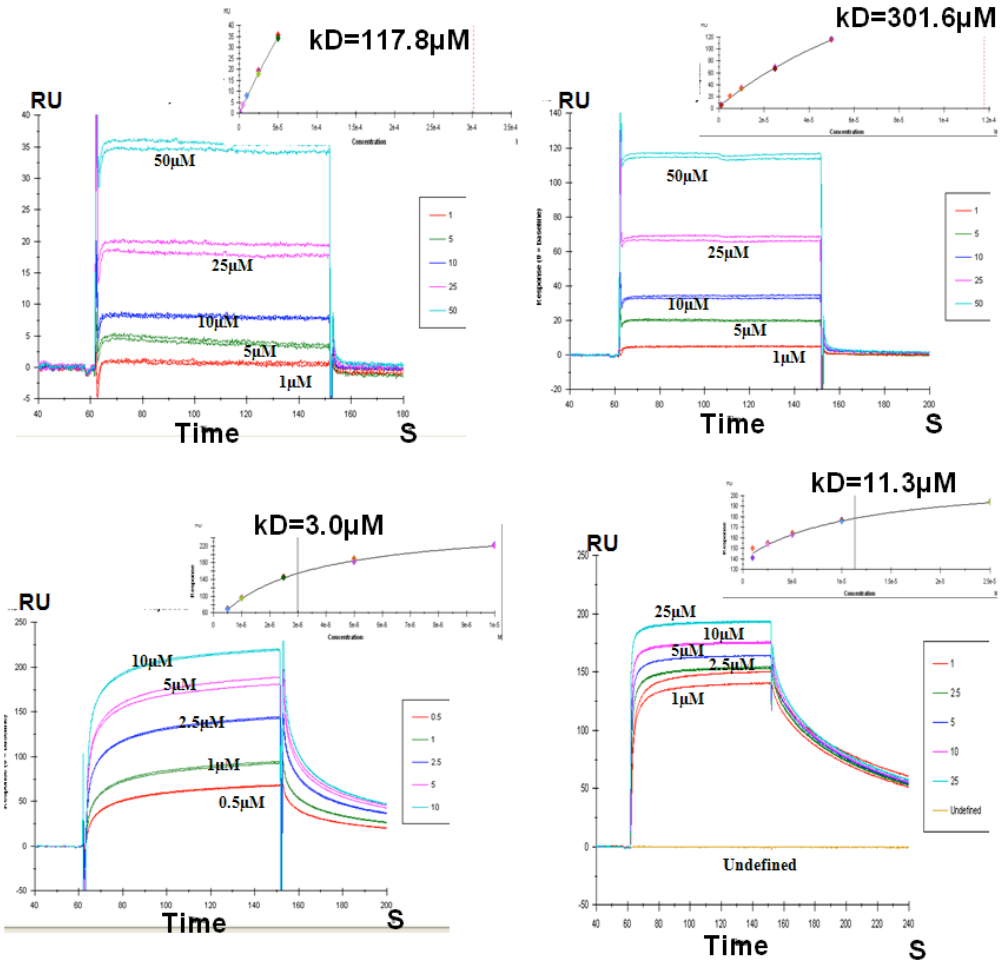



***Figure 5.18 "Definitive" $K_D$ measurements from SPR data for CR1 constructs flowed over Rh4.9***

***(A)*** A concentation series (0 – 50 mM) of CR1 15-17 was injected over FC 3 and 4) that had been loaded with 408 and 1476 RUs of Rh4.9. Insert – a plot of response versus concentration of CR1 15-17 used to estimate a $K_D$ value for this interaction. ***(B)*** As in panel (A) but for (R,G)CR1 15-25. (***C)*** As in panel (A) but for a 0-10 mM concentration series of sCR1. (***D)*** As in panel (A) but for
CR1 1-3.

| CONSTRUCT | COMPLEMENT CHIP *(μM)* | | | MALARIA CHIP *(μM)* | |
|---|---|---|---|---|---|
| | C3B | C4B | C1Q | NTS-CDBL-CIDR | Rh4.9 |
| CR1 15-25KR | 1.5 | 6.9 | 8.8 | - | - |
| CR1 15-25ER | 2.4 | 8.1 | 9.8 | - | - |
| CR1 15-25KG | 2.1 | 6.9 | 8.2 | 137.9 | 117.8 |
| CR1 15-25EG | 1.8 | 8.0 | 7.6 | - | - |
| CR1 15-17 | 2.2 | 11.5 | 11.0 | 104.2 | 301.6 |
| CR1 1-3 | 79.2 | 21.4 | 202.9 | 103.8 | 2.98 |
| sCR1 | 1.3 | 1.5 | 6.4 | 14.15 | 11.33 |

*Table 5.2 Summary of kD values*

### 5.8.5 ELISA experiments on binding of CR1 constructs to complement proteins

In addition to the SPR experiments described earlier, Enzyme-Linked Immunosorbent Assay (ELISA) was employed to further investigate the possible effect of the knob polymorphism on the binding sites of the CR1 protein. Detail amount of proteins and conditions used for the experiment are described in section 2.4.3 of chapter 2. Results obtaines from this experiment support the observation in the SPR described above, that there is no evidence in the influence of polymorphism on their interactions. With C3b interaction, sCR1 (with two high affinity C3b binding sites) should bind more avidly to C3b than CCPs 15-25 (one C3b binding site) but this can't be concluded form this experiment since the detection system employs a rabbit poly anti-CR1 – thus, more potential binding sites on full length sCR1 vs CCPs 15-25. Also, African mutations do not influence CCP 15-25 binding to C3b. (Figure 5.19A).C4b also has same conclusion as with C3b binding, but despite CR1 having three C4b binding sites, only one is in CCPs 15-25 and this might be a likely reason for overall lower binding. (Figure 5.19B). C1q binds to CR1 and also binds to CCPs 15-25 but with a lower avidity compared to full

length CR1. African mutations in repeat CCP24 do not appear to influence C1q binding. These outcomes suggest that there is more than one binding site on CR1 for C1q. (Figure 5.19C). In MBL, binding to CR1 is weak; also, African mutations do not influence MBL binding. (Figure 5.19D). The implications of these together with the SPR data are discussed in section 6.3.4 and 6.3.5 of chapter six.

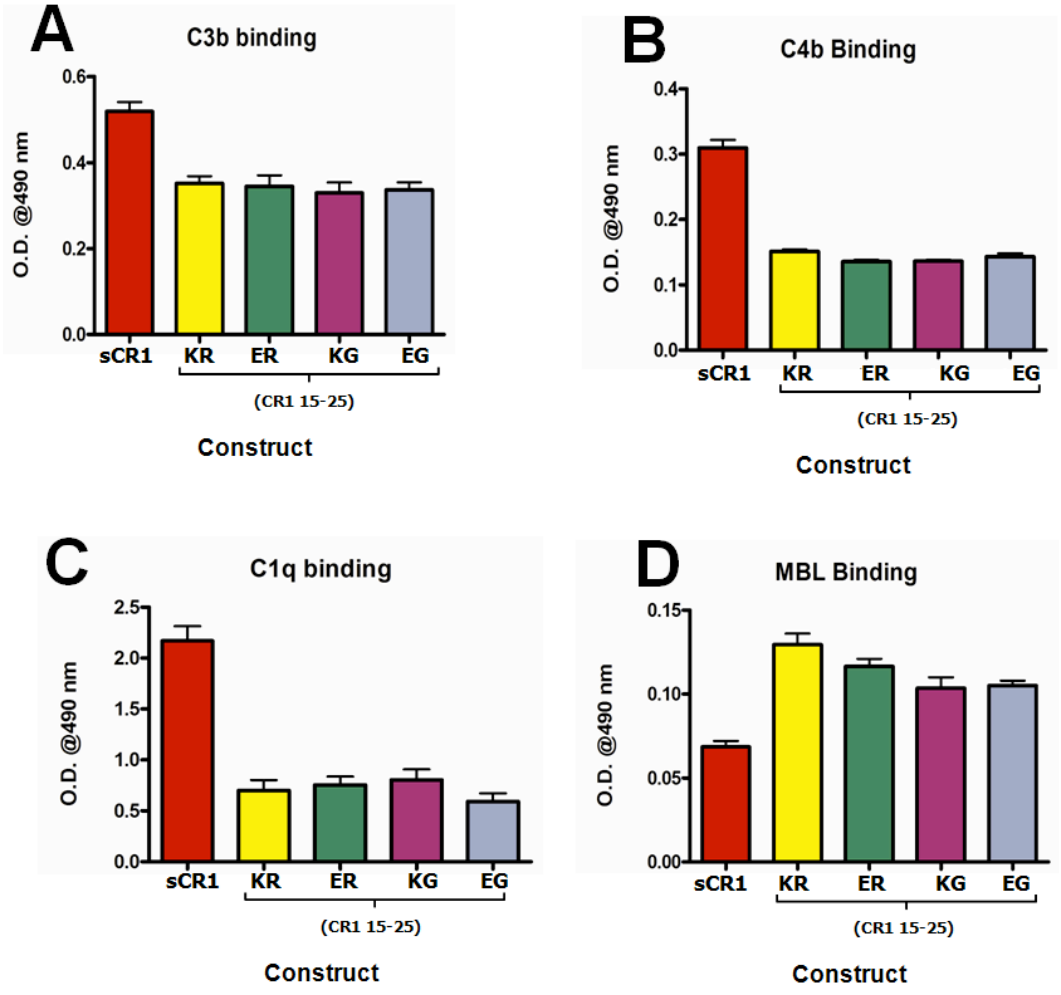



***Figure 5.19 Binding of CR1 polymorhpic forms to complement proteins by ELISA.***

(A) sCR1 and CR1 15-25 variants binding to C3b. (B) Same as in panel (A) but plate coated with C4b (C) Assessing C1q interaction on CR1 coated surface (D) Behaviour of the variants on MBL coated surface. (In B, C, D,and E) Labels below indicated in respective colour coded bars.

# CHAPTER SIX

# DISCUSSION, CONCLUSION AND RECOMMENDATIONS

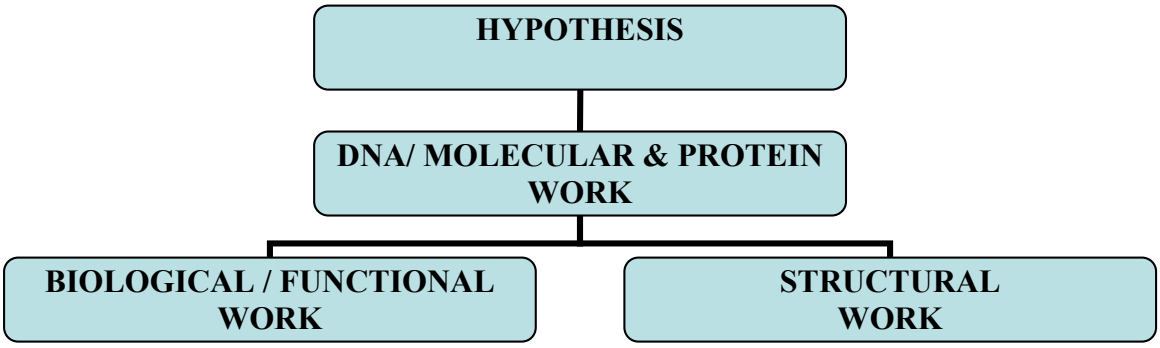
# CHAPTER 6 DISCUSSION, CONCLUSION AND RECOMMENDATIONS

## 6.1 Discussions

This chapter will discuss the results presented in Chapters 2, 3 and 4 with respect to the limitations of the techniques used and the implications of the data obtained. The overall strategy of the current work is summarised in the flowchart below.

The overarching aim of this project was to test the hypothesis that the McC(a/b) and Sl(1/2) Knops blood-group polymorphisms in LHR-D of CR1 modulate the capacity of the protein to act as one or more of the following: a complement regulator; the immune adherence receptor; a mediator of rosetting of *P. falciparum*-infected erythrocytes; or a ligand for invasion of erythrocytes by *P. falciparum* merozoites.

Towards this end, *CR1* constructs were overexpressed and recombinant versions of CR1 "fragments" (corresponding to individual CCP modules or series of modules from its exodomain) were produced in milligram yields. The resultant proteins were purified, structurally characterised and then tested in biological settings and in biochemical assays. Each aspect of the work will be discussed in turn below.




***Flow chart 6.1 Overview of the current study***

# CHAPTER 6 DISCUSSION, CONCLUSION AND RECOMMENDATIONS

## 6.2 Protein Production and Purification

First the DNA manipulation steps and the cloning work will be reviewed and then the production and purification of the recombinant proteins will be considered with an emphasis on yields.

### 6.2.1 DNA manipulation and cloning

The eukaryotic organism *P. pastoris* was the expression host of choice in this work. No attempt was made to use *E. coli* in the current project due to the presence of multiple disulfide bonds within the target proteins. Mammalian and insect cells might also have been explored in this respect but they often produce rather low yields and are expensive to culture. Cloning in this study typically involved insertion of the desired PCR-amplified DNA sequence (from human cDNA) into the plasmid expression vector (pPicZ αB), via a route that involved Invitrogen's TOPO cloning technology followed by double digestion with restriction endonculeases. After amplification in *E. coli* cells, the recombinant plasmid was linearised and used to transform *P. pastoris* cells (KM71 H strain). This work followed a route well established in the Barlow group but involved overcoming some challenges posed by the requirement for relatively long stretches of DNA to be amplified and inserted into plasmids.

The cloning of DNA encoding varying numbers of CCP modules (from one to eleven) was successfully accomplished in this study, entailing manipulation of DNA segments containing between ~150 and ~3000 base pairs. Increasing degrees of difficulty

were encountered in amplification, ligation and transformation steps, as the size of the construct increased. For instance, amplification of the DNA (initial PCR) for CR1 15-25 yielded low and inconsistent quantities of nucleic acid (2133 bp); only one of three attempts resulted in detactable bands on an agarose gel (see Fig. 3.10A). Amplification of the DNA encoding CR1 17-25 was also problematic partly because of sequence similarities among the LHRs, and the target DNA band was less intense compared to the shorter, unwanted, one (see Fig. 3.8A).

To increase the chances of successful transformation, and in the hope of increased copy numbers in the *P. pastoris* clones (especially in the case of the longer construct), the amounts of DNA used were increased by between two- and three-fold. The positive effect of this strategy could be seen in the eventual production of useful quantities of all four polymorphic forms of the 15-25 fragment (see Figs. 3.11 to 3.14).

### 6.2.2 Production of CCP module-containing protein fragments

The variously sized fragments of the CR1 exodomain seemed to be produced at different levels in *P. pastoris*. In general, the smaller constructs *(e.g.* CR1 21 and CR1 10-11) were obtained in greater yields than the longer ones (*e.g.* CR1 17-25 and CR1 15-25). To compensate for this, the longer DNA constructs were expressed in larger-scale cell cultures, even during the initial expression trials. The reasons for the observed lower yields were not fully investigated. For example it was not established whether the problem lay at the level of transcription, translation, post-translational modifications, secretion or susceptibility to proteolysis.

# CHAPTER 6 DISCUSSION, CONCLUSION AND RECOMMENDATIONS

The longer the mRNA transcript, the higher the chances of stem-loop formation and the consequent potential for early termination of translation on the ribosome. Likewise the presence of rare codons is more likely in longer constructs (given that codon usage in yeast differs from that in man), thus translation is slower due to shortages of cognate tRNA. In general, sequence-specific translation and transcrition problems should be minimized by recourse to codon optimization and thus the case for procuring a synthetic gene is enhanced for these longer fragments. This strategy was subsequently adopted in the case of the recombinant *P. falciparum* protein targets (see below).

At the protein level, it has been noticed previously (in the Barlow group, unpublished) that the presence within various recombinant fragments of specific CCP modules (*e.g.* modules 4 and 14 of factor H) appear (for unknown reasons) to correlate with poor production yields, so it is possible that some of the longer CR1 constructs also harbour "problem" CCP modules. It may also be that some CCP modules in CR1 require neighbours for stability and that when they are exposed at the N or C terminus of a recombinant protein fragment they are unstable and prone to proteolysis. It is also possible that the number of disulfides (two in each module) could pose a problem for the post-translational modification machinery of the yeast cell. Presumably, the more disulfides the more chances for incorrect disulfide formation leading to rejection by chaperones in the endoplasmic reticulum.

Another (potentially related) possibility is that the recombinant proteins are being secreted but are being proteolysed during culture, during the harvesting of the secreted protein from the cell culture supernatent, or during the enzymatic deglycosylation incubation step that is required in most cases. Barlow-group members have observed this

phenomenon on several occasions in the case of *P. pastoris*-expressed proteins. The addition of peptides such as found in tryptone or peptone to the growth medium might help to avoid proteolysis during cell culture. It might also be beneficial to engineer out all N-glycosylation sites and thus avoid the necessity for the deglycosylation step.

### 6.2.3 Production of DBLα and DBLα-CIDR-DBLγ

In this study, the challenges discussed in sections 6.2.1 and 6.2.2 became even more apparent when attemting to produce PfEMP1 fragments, *i.e.* the DBLα and DBLα-CIDR-DBLγ constructs (see Fig. 5.5A, B and C). Anticipating that the length of the DNA would make cloning and expression more problematic, and given the known "high-AT content" issues related to the heterologous expression of *P. falciparum* genes, genes for both DBLα (1,137 bp) and DBLα-CIDR-DBLγ (3,183 bp) were designed and synthesised by GeneArt to optimise codon usage (and ordered in a form already ligated into the pPicZ αB vector). Not only were rare (in yeast) codons replaced but the proprietary optimization algorithm also minimized sequences likely to form mRNA secondary structure, repeating codons that might introduce frameshift errors and cryptic translation-start sites *etc.* Furthermore the opportunity was also taken in the design process to engineer out the N-glycosylation sites by replacing N with Q in instances of the sequence NXS/T. During DBLα production trials, peptone was sometimes added in an attempt to reduce proteolytic activities. Despite all of these measures, the protein yield was such that it could only be detected by Western blotting (see Fig. 3.17D).

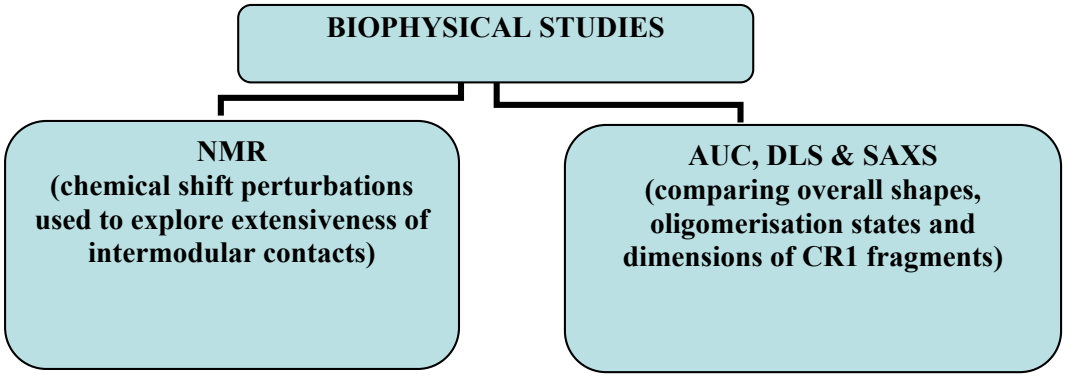
### 6.2.4 Summing up

To sum up, recombinant protein production in yeast of the CR1 fragments (ranging from the single module CCP21 (Fig. 3.2) to as long as 11-modules (Fig. 3.10) yielded milligram quantities of pure material. On the other hand, attempt to produce the *P. falciparum* proteins yielded disappointing and inconclusive results (Fig. 3.17). One difference not noted so far is that domain boundaries were very obvious in the case of most CCP modules (excluding the aforementioned cases where one module may require a neighbour for stability). On the other hand, the domain boundaries of DBLα domains are more difficult to establish and more appropriate choices, or more extensive trials of different boundaries might have led to better production yields. Of course, not all possibilities for improving the yield were exhausted in the case of the DBLα–containing fragments due to time limitations. For example, it would be worth testing a wide range of culture conditions for optimal expression and seeking further ways of reducing proteolytic degradation. It would also be worthwhile exploring ways of purifying and characterising the potentially useful trace amounts of protein actually produced.

# CHAPTER 6 DISCUSSION, CONCLUSION AND RECOMMENDATIONS

## 6.3 Biophysical Discussion

The structural work discussed here includes NMR, AUC and DLS as summarised below.




***Flow chart 6.2 Biophysical studies***

### 6.3.1 NMR

NMR chemical shifts are highly sensitive to context and this phenomenon was exploited in a simplified approach to structural investigations that did not demand the major work that would have been required in a full structure determination. The very few changes in chemical shifts for the amides of CR1 21 upon attachment of CCP 22 (to form CR1 21-22) provided strong evidence that these modules do not share extensive side-by-side interactions despite the atypically long linking sequence between them that would allow or even promote such an arrangement. This is not to say that these modules do not adopt a "bent" end-to-end arrangement (*i.e.* one with a sizable angle of tilt between the modules) as was observed for CCPs 12-13 of factor H that are also joined by an eight-residue linker (Schmidt *et al.*, 2010). $^{15}$N Relaxation experiment (coupled with the almost complete backbone assignment already carried out) will help to clarify whether the 21-22 pair forms a rigid or a flexible structure.

Based on similar arguments, relatively few changes in chemical shifts of modules 21 and 22 (in CR1 21-22) accompany attachment of modules 20 and 23 to create the four-module construct CR1 20-23 and, very importantly, none of these perturbed resonances were assigned to residues in the eight-residue linker between CCPs 21 and 22. Thus whatever end-to-end arrangement of modules existed in CR1 21-22 must be retained in the CR1 20-23 context. It was noticed, however, that most of the changes that did occur were in CCP 22 (Appendix D and Fig. 4.11) indicating a more intimate 22-23 interface (these modules joined by a typical four-residue linker) compared to the 20-21 interface.

### 6.3.2 AUC, DLS and SAXS

The conclusions that were drawn from the NMR data were confirmed from the AUC experiments and SAXS-derived data for CR1 20-23. The SAXS scattering curve for CR1 20-23 fitted very well to a model (Fig. 4.16) of the structure that is extended, with no evidence for a 180-degree bend in the middle. In fact the model reveals that none of the modules is very tilted compared to its neighbours and that there is no overall curvature within this region of CR1 that spans the LHR-C:LHR-D boundary. The DLS results obtained for CR1 15-25 were, overall, consistent with a non-aggregated protein that is neither fully extended nor globular in the case of each of the four variants despite some variation in “contamination” of preparations by aggregates. The AUC data for these four constructs also indicated that they were all nearly identical in apparent molecular weight, association state (monomeric) and overall shape and dimensions. In future work, SAXS studies of CR1 15-25 and CR1 17-25 would be valuable if the data were analysed taking

into account the available high-resolution structure for CR1 15-17, the aforementioned model of CR1 20-23 and the AUC results already obtained for CR1 15-25. This exercise could yield a reliable model for the CR1 15-25 structure and could also shed light on the important issue of inter-module flexibility within this fragment.

### 6.3.3 Summing up

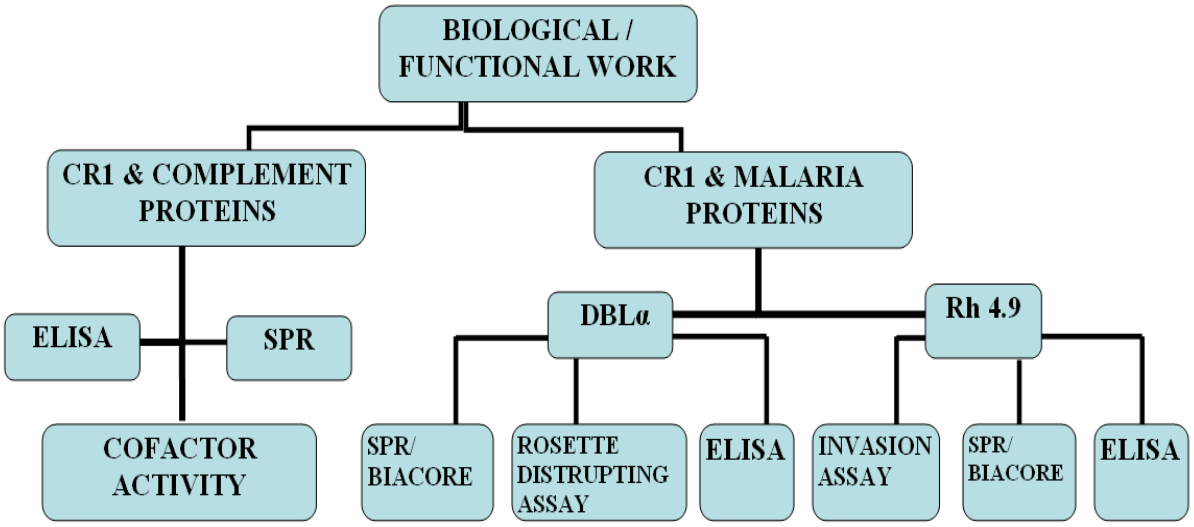
This work has effectively falsified our hypotheses concerning the architectural role of the unique linker between CCP modules 21 and 22. A limitation of the study, however, is that it was not possible or feasible to work with longer constructs that would have been more representative of the full-length exodomain. Clearly (based on the earlier discussion) these would have been too difficult to produce in a recombinant form. Moreover they would have become increasingly difficult to meaningfully characterise using currently available biophysical techniques. This is why even though soluble full-length CR1 was available it was not investigated from a structural perspective in this study. Indeed sCR1 has already been studied by SAXS (Furtado *et al.,* 2008) and by negative-staining transmission electron microscopy (Weisman *et al.*, 1990) among others **(**Kirkitadze *et al*., 1999**,** Kirkitadze *et al*., 1999**)** but these studies did not shed light on the module-specific issues addressed in this project.

# CHAPTER 6 DISCUSSION, CONCLUSION AND RECOMMENDATIONS

## 6.4 Biological work

To restate the starting hypothesis, polymorphisms in CR1 appear to be under selective pressures in different geographical regions suggesting that they have major although unknown phenotypic consequences. The aim of the current study was to explore the functional repercussions of some key Knops blood-group antigens by studying in a range of assays the recombinant fragments of CR1 produced as discussed above. Constructs had been designed based on the site of polymorphic variation (CCPs 24-25) being separated by a potentially hinge-like region (putatively CCPs 20-23) from one of the three key functional/ligand-binding sites (CCPs 15-17) of CR1. Thus with availability of CR1 15-25 it became possible to investigate whether any of the activities of functional site 2 were modulated by the presence of the different variants. A range of techniques were utilised to measure interactions with both host and *P. falciparum* proteins; these have been summarized in the flowchart below, and will be discussed in the following sections. Note that the finding that CCPs 1-3 (and not 15-17) are exploited for invasion of erythrocytes by merozoites was made after the design of the CR1 constructs used in this study.

# CHAPTER 6 DISCUSSION, CONCLUSION AND RECOMMENDATIONS




***Flow chart 6.3 Biological studies***

### 6.3.1 Factor I-cofactor activity *versus* C3b

The sequential proteolytic cleavages at two positions within the CUB domain of the C3b α-chain to produce iC3b (and the tiny C3f fragment) by factor I is strictly dependent on the presence of protein cofactors that belong to the regulators of complement activation family. As detailed in the Introduction, two homologous membrane-bound proteins – CR1 and membrane cofactor protein (MCP, CD46) (along with the soluble protein factor H) - each carry out this role via their respective C3b-binding sites consisting of three or four contiguous CCP modules. The cofactor activity of CR1 is almost certainly one of its main biological tasks since the initial cleavage product, iC3b, is no longer capable of binding to factor B (to form a C3b convertase) hence CR1 protects erythrocytes from uncontrolled deposition of C3b and subsequent complement-mediated hemolysis. The first product, iC3b is an important opsonin, like C3b, but is also a ligand for complement receptors 2 and 3; in the presence of CR1 (but not of MCP or factor H) factor I carries out

a further, third, cleavage of the α-chain creating C3c (of no known function) and C3dg that is non-specifically cleaved by other proteases to produce C3d, which remains tethered to membranes or surfaces. C3d is a major ligand for CR2 on B-cell surfaces where it sensitizes cells for antibody production by ~10,000-fold.

It was reasoned that if the cofactor activity of CR1 were modulated by the Knops blood-group variations in LHR-D then this could have major implications for the individuals carrying such variants, hence providing the much sought link between genotype and phenotype for this blood group. Therefore measurment of cofactor activity represented a good starting point in terms of attempting to explain why specific CR1 variants are disproportionately represented in certain populations.

A simple fluid-phase assay was carried out in which each of the four variants of CR1 15-25 were added individually to a mixture of factor I and C3b. After an interval of time, C3b and its proteolytic fragments were resolved by SDS-PAGE and visualized using Coomassie staining. Like the sCR1 and factor H positive controls, all four CR1 15-25 polymorphic variants were shown to have cofactor activity towards C3b. The assay also confirmed previous work showing that CCPs 15-17 appear necessary and sufficient for this activity while CCPs 21-22 (negative control) had no more cofactor activity than the buffer. It also became clear that the CR1 constructs that include modules 15-17 are cofactors both for cleavage of C3b to iC3b and (unlike factor H) further cleavage to C3c and C3dg; this is, again, in line with expectations. Thus together these experiments confirm the utility of these four (CR1 15-25) constructs for monitoring any modulating influences on functional site 2 of the variations in CCPs 24-25.

On the other hand, the assays were not carried out in a way that would have revealed subtle functional disruption since no proper time course or dose-dependency of activity was recorded. Ideally, the amount of protein required to achieve 50% cleavage of C3b over a fixed time interval would have been compared between variants; alternatively using limiting amounts of protein, the times required to reach 50% C3b cleavage could have been compared.

The proteins produced for this study were also sent to the laboratory of our collaborator, John Atkinson (Washington University Medical School, St. Louis) who also carried out cofactor assays on the four variants. These investigators employed a more sensitive methods based on detection of biotinylated proteins and they looked at both C3b and C4b cleavage. In these assays, only partial C3b/C4b cleavage occured *i.e.* the reaction had not neared its end-point when terminated. Thus it is possible to compare the extent of the four reactions mediated by the four variants. Although it would have been possible (and preferable) to carry out a densitometric analysis to quantify the strengths of the bands, visual inspection is sufficient to conclude that all four lanes (*i.e.* each of which contains a variant) look very similar, both in the case of C3b and of C4b (see Figs 5.5E and F).

In summary, while more rigorous assays should be carried out to reveal any subtle effects, there are no major differences between the McC(a/b) and Sl(1/2) Knops blood-group variants of CR1 in terms of the ability of the second copy of functional site 2 to act as a cofactor for factor I. It seems highly unlikely that the other copy of site 2 (in CCPs 8-10) would be influenced by variations in CCPs 24-25, while functional site 1 (modules 1-3) is probably not a major contributor to cofactor activity anyway. Hence we can

conclude there will be little differences in protection of erythrocytes from C3b/C4b amplification, via the cofactor route, between individuals in different Knops blood groups. Note that this finding is entirely consistent with the structural studies that did not provide any evidence for physical proximity between LHR-C and LHR-D.

It should also be mentioned that CR1 has decay accelerating activity as well as cofactor activity. Thus CR1 accelerates the irreversible decay of the C3 convertases, C3bBb and C4b2a, and also catalyses disassembly of the trimolecular C5 convertases. Decay-accelerating activity directed at the C3 convertases mainly resides in functional site 1 so it did not seem worthwhile to measure this in CR1 15-25. It sems unlikely that the sequence variations in CCPs 24-25 will have any influence on activities resident in CCPs 1-3. The C5 convertase decay acceleration also requires functional site 2 (Goldberg *et al.*, 1991,1998) but due to time limitations (and the lack of suitable constructs), this activity of CR1 was not investigated.

### 6.3.2 Rosette disruption assays

The formation of rosettes involving infected and non-infected red blood cells has been correlated, in some populations, with severe malarial anaemia and cerebral malaria, regarded as the most severe and life-threatening forms of the disease **(**Doumbo *et al*., 2009**).** It has also been observed that red cells of the knobs $Sl_2$ phenotype form fewer rosettes (Rowe *et al.,* 1997). It was previously shown that both sCR1 and CR1 15-17 can disrupt preformed rosettes, presumably via a competition mechanism given that functional site 2 of CR1 reportedly represents a key binding site for PfEMP1 (Rowe *et al.,* 1997, 2000 ). Rosette-disruption assays therefore represent a potentially insightful

way in which to investigate whether the four variants of CR1 15-25 differ in their affinity for PfEMP1.

It was established in the current work that all four variants of this recombinant protein disrupted rosettes, which is consistent with the presence in these fragments of correctly folded modules 15-17 (and hence also with the positive cofactor assays, see Fig. 5.1). There were, however, no significant differences in their respective rosette-disrupting activities. In the current experiments, the concentrations of CR1 fragments used were 20 μM (incubated over a time period of 30 mins). A limitation of this assay, which is quite complicated to carry out, is that no full investigation of dose-dependency was possible and so we cannot be sure that the extent of observed rosette disruption is a linear function of the amount of active CR1 present. Furthermore the errors are inevitably rather large and so small differences could not have been detected. On the other hand, these functional results are totally consistent with the structural ones (above) that did not support the case for physical proximity between CCPs 24-25 and CCPs 15-17 of CR1.

Any effects on rosette formation, in the plasma of individuals, of the Sl and McC polymorphic variations of CR1 may be indirect ones. It is for example possible, since CR1 molecules on cell surfaces form clusters and such clusters could influence rosette formation, that it is the ability of CR1 to cluster that is influenced by the Knops blood-group polymorphisms. Note however that our biophysical studies did not detect differences in self-association between the variants. It should also be borne in mind that in a population-based study subjects might have been exposed to the malaria parasite previously and might have adapted to infection by changing their rosetting behaviour in some unknown way.

### 6.3.3 Invasion Experiments

During the current project it emerged that CR1 is the major sialic acid-independent receptor for invasion of erythrocytes by merozoites (Tham *et al.,* 2010). This immediately raised the intriguing and exciting possibility that different Knops blood group variants of CR1 might interact differentially with the *P. falciparum* "invasion" protein Rh4.

The proteins (CR1 15-25 and others) created in the current project were therefore sent to our collaborator, Alan Cowman (Melbourne) for testing in invasion-competition assays. None of the CR1 15-25 constructs were able to significantly inhibit invasion. Although this was a negative finding, it was a very useful piece of evidence (taken along with assays on the other Edinburgh-produced fragments tested by the Melbourne group) for the finding that the invasion-critical Rh4-binding site lies in CCPs 1-3 (Tham W H *et al.,* 2011). In particular it is intriguing that the very similar CCPs 15-17 are not a region of interaction during utilization of this invasion route.

It was not really surprising therefore that all four polymorphic forms of CR1 15-25 had the same or identical properties in this assay. In summary, these results imply that the Knops blood-group polymorphism do not affect the sialic acid-independent invasion pathway. A limitation of these experiments is that they did not directly test whether the variants in CCPs 24-25 could modulate CCPs 1-3 since we did not have the requisite constructs, but this does seem unlikely.

### 6.3.4 Interactions between CR1 and C3b/C4b

Previous work (Krych *et al.,* 1991) had identified Sites 1 (CCPs 1-3) and 2 (CCPs 8-10 and 15-17) in LHRs A-C as being critical for C3b and C4b binding and hence to the key biological role of CR1 as the immune-adherence receptor (as well as its ability to prevent C3b amplification on the erythrocyte surface *via* its cofactor and decay accelerating activities). These published reports based on ELISAs had suggested that C3b bound primarily to Site 2, while previous unpublished SPR-based work (in Rosie Mallin's PhD thesis, University of Edinburgh) indicated a $K_D$ in the range of 2 μM for the interaction between CR1 15-17 and immobilised C3b. In the current work all four polymorphic forms of the protein construct CCPs 15-25 (*i.e.* encompassing modules 15-17 corresponding to the second copy of functional site 2) bound to immobilised C3b (in SPR experiments) with similar "sensorgram" profiles and $K_D$ values in the narrow range of 1.4-2.4 μM. These values were similar to that obtained in the current study (on the same chip) for CR1 15-17, *i.e.* ~ 2 μM but somewhat weaker than the ~1 μM $K_D$ obtained for sCR1 with its multiple interaction sites for C3b. Furthermore, all four variants bound with similar affinities to C4b (in the range of 6.9 to 8.1 compared with of 11 μM for minimal construct CR1 15-17 and 1.5 μM for sCR1 with its three C4b-binding sites). Note that the site 2-containing CR1-fragments bound C3b somewhat more tightly than FH fragments containing CCPs 1-4 ($K_D$ = 10 μM) or FH CCPs 19-20 ($K_D$ = 4 μM).

Similar results were obtained using the ELISAs performed in the Atkinson laboratory (See figure 5.19 of section 5.8.5). These results were less quantitative than the SPR-derived data but nonetheless provide valuable, orthogonal, validation for a lack of any measurable differences between the CR1 15-25 variants in terms of their ability to

bind C3b or C4b. In particular note that ELISA does not involve the chemical modification of C3b/C4b (*i.e.* amine coupling) that was carried out when performing the SPR studies.

Thus the Knops blood group polymorphisms in LHR-D have no or negligible effect on the interactions of the LHR-C copy of functional site 2 with C3b and C4b, as the immune adherence receptor (Krych *et al.,* 1992) and its potential cytolytic consepuences (Fearon *et al.,* 1989). It seems therefore very unlikely that they will influence binding at the LHR-B copy of functional site 2 or indeed site 1 in LHR-A. Therefore these results seem to rule out the hypothesis that the immune adherence functionality of CR1 is significantly modulated by the McC (a/b) and Sl (1/2) polymorphisms.

These functional results – showing a lack of effect of variations in CCP 25 on interactions of functional site 2, with either parasite or host ligands - are entirely consistent with conclusions from the structural work (in Chapter 4 and discussed above) that LHR-D is not bent back on to LHR-C via the long (CCP 21-22) linker between these LHRs (between). Taken together the structural and functional studies suggest LHRs-C and D are structurally and functionally independent units and that LHR A, B and C presumptuously bind C3b and C4b in a similar manner to factor H modules 1-3 **(**Wu *et al.,* 2009**).**

### 6.3.5 LHR-D specific interactions (ie with C1q and MBP)

The putative role of LHR-D in ligand binding is much less well explored than that of functional site 1 or site 2. Reports (Klickstein *et al,*1997*,* Ghiran *et al*. 2000, Tas *et al*., 1999) have suggested that LHR-D is the main interaction region for binding of CR1 to

# CHAPTER 6 DISCUSSION, CONCLUSION AND RECOMMENDATIONS

C1q, a multimeric protein that triggers (upon binding multivalently to antibody-antigen complexes) the classical pathway of complement activation. The biological function of a C1q-CR1 interaction is, however, unclear. In the current study it was established that CR1 15-25 – which incorporates part of LHR-D (*i.e.* modules 22-25 but not modules 26-28) – exhibits detectable binding (observed via SPR and ELISA) to C1q although this is weaker than binding to C3b. As observed in the case of CR1 binding to C3b and C4b, however, there was no significant difference in binding to C1q between the four polymorphic forms. Of interest is that in our study CR1 15-17 (but not CR1 20-23) bound to immobilised C1q (by SPR). Studies based on ELISAs conducted in the Atkinson group also showed that the four CR1 15-25 variants bound equally to C1q although not as well as sCR1, consistent with the notion that modules 26-28 (in addition to CCPs 22-25) could be involved in C1q binding.

LHR-D was also previously reported to bind to MBL of the lectin-initiated complement activation pathway and this was explored by ELISA (although not by SPR due to time constraints). Binding of MBL to CR1 appeared to be weak in this study, while no variation between the polymorphic versions of CR1 15-25 could be discerned. It is however unclear why the binding to CR1 15-25 appeared stronger than binding to sCR1. It may be that the very poor signals obtained from these ELISAs were too low to represent significant interactions.

In conclusion, all four variants bind equally well to C1q immobilised by amine coupling on an SPR chip or absorbed onto an ELISA plate, and poorly (if at all) to MBL in ELISA. Although our constructs lack a complete copy of LHR-D they would nonetheless be expected to highlight differences between the variants if these existed.

### 6.3.6 Possible effect of glycosylation

As has been speculated, the presence of N-linked glycans could influence the interaction of CR1 with its binding partners (Krych-Goldberg *et al.,* 1998). A setback of our protein production strategy is that *P. pastoris* does not incorporate mammalian-type glycans but rather "hypermannosylates", producing a heterogeneous mixture of improperly glycsosylated recombinant proteins that require treatment with endoglycosidases prior to purification and biophysical characterisation. The resultant proteins are expected to carry a GlcNAc residue at each potential N-glycosylation site (for the locations of these sites in CCPs 15-25 see Table 3.2) so in no sense can they said to be physiologically glycsosylated. Thus a limitation of the present work is that it effectively ignores any effects of glycosylation on structure and function of CR1.

Furthermore, the bigger of the two peaks observed (out of Peak A and B) in the size-exclusion purification of CR1 15-25 KR (Fig. 3.11) could either be due to conformational change or incomplete deglycosylation (this requires testing by mass spectrometry). In any case, the protein from peak A (of Fig. 3.11C, later described in the SPR experiment as 'CR1 15-25Cauc glyc') exhibited different binding pattern on all the three surfaces tested (loaded with C3b, C4b or C1q- Figs. 5.9, 5.11 and 5.13). This observation is worth further investigation.

### 6.3.7 Interactions of CR1 with parasite-encoded protein domains DBLα and Rh4.9

The observed (by SPR) binding of sCR1, or the other recombinant CR1 constructs, to recombinant DBLα-containing constructs (*i.e.* representing the rosette-mediating protein

PfEMP1) was generally very weak and possibly corresponded to non-specific interactions. For example the CR1 1-3 sensorgrams were very similar to the CR1 15-17 sensorgrams. Moreover, the various polymorphisms in CCPs 24-25 did not seem to have any effect on the putative weak binding between CR1 15-25 and DBLα. While none of the measured interactions were wholly convincing due to low responses and failure to saturate, the interaction with sCR1 was the only one characterised by fast on and off rates, and a very low signal considering its higher MWt (see Fig. 4.16C), implying that it was if anything weaker than the interactions with the fragments.

There are a number of possible explanations for our inability to observe physiologically meaningful interactions between the recombinant DBLα-containing constructs and any of the CR1 constructs. First, one of the recombinant versions of DBLα (P) behaved anomalously during purification and might not have been fully folded due, for example, to incorrect choices of domain boundaries during the expression work. Second, the binding site in PfEMP1 for CR1 might in fact include regions beyond the DBLα–CIDR domain used in the "definitive" SPR measurements, and attempts in Edinburgh to construct triple domain DBLα-CIDR-DBLγ unfortunately failed (Figure 3.17D). Third, interactions between CR1 and PfEMP1 may be inherently very weak but in the physiological setting it could be that multiple CR1 molecules interact with multiple PfEMP1 molecules thus achieving a strong cell-cell adherence via extensive multivalency; note that 20 μM concentrations of CR1 15-17 were sufficient to disrupt rosettes – see Fig. 5.3. In summary, attempts to test directly whether the McC(a,b) and Sl(1,2) polymorphisms modulate interaction with PfEMP1 were inconclusive but the lack of effects observed in the rosette-disruption assays (discussed above) combined with our

inference that LHRs-C and D are structurally and functionally independent units, would indicate that this is, in any case, very unlikely.

In contrast to the studies aimed at establishing dissociation constants for the interaction between PfEMP1 and CR1, quantitative studies of the Rh4-CR1 interaction by SPR were straightforward. These studies confirmed that CCPs 1-3 were necessary and sufficient for binding to the recombinant Rh4.9 domain. They further showed that CCPs 15-25 do not harbour a region that can interact with Rh4.9 and that this is true for all four variants thus excluding the possibility that one or other of the Knops blood group SNPs might create a new binding site elsewhere than CCPs 1-3. The CR1 15-25 "negative" control was also important in that it helped make the case for non-involvement of other CCPs (that were not tested directly) in Rh4.9 binding, based on the very high levels of internal sequence similarity within CR1.

### 6.3.8 Summing up

Our CR1 15-25 proteins behaved as expected in that they bind C3b, C4b, and (less convincingly) C1q but not Rh4.9 (and by extension *P. falciparum* Rh4). Moreover they disrupt rosettes formed by parasitized erythrocytes and have co-factor activity for cleavage (*via* factor I recruitment) of both C3b and C4b **(**Roversi *et al*., 2011**)**

In nearly all cases the activities of CR1 15-25 relative to those of CR1 15-17 and full-length sCR1 were in line with expectations and with the previous assertions that CCPs 15-17 corresponds to a complete copy of C3b/C4b/PfEMP1-binding functional site 2, while sCR1 contains multiple binding sites for C3b and C4b. These internally consistent results that are also in agreement with the literature thus provide a strong basis upon which valid comparisons between the variants may be made. None of the studies,

however, detected significant differences between the variants of CR1 in any of its functional aspects.

Thus, taking together all of the findings regarding interactions of CR1 with the *P. falciparum* proteins we can now conclude that: sialic acid-independent invasion is mediated by functional site 1 in LHR-A of CR1 and does not involve the highly similar site 2; rosetting is mediated by functional sites 2 in LHRs-B and C but does not involve functional site 1 (although a direct interaction between PfEMP1 and any region of CR1 remains unproven); the Knops blood-group antigens that lie in LHR-D, despite their supposed link to resistance to severe malaria in certain populations, do not modulate the rosette-promoting interaction at site 2 (nor, presumably, the Rh4 interaction with site 1).

In sum, the above studies serve to: authenticate the biological activity of our recombinant products; confirm much previous published work regarding the location of functional sites in CR1; agree with structural studies implying that there are minimal interactions between CCPs in LHR-C and LHR-D; and very substantially weaken – or even falsify - the hypothesis that Knops blood group variants differ in terms of their interactions with known host and parasitic ligands.

### 6.3.9 Some recommendations for future studies and directions

**1)** A comparison of deglycosylated and glycosylated proteins would be useful to assess the impact of N-glycans on function both within the *in vitro* assays and the cell-based studies. This would require the use of strains of *P. pastoris* that have been engineered (as reviewed by Wildt .S and Gernogross T.U) to incorporate mammalian style N-glycans.

**2)** While the rosette-disruption assays did not reveal any differences between variants, it would be very useful to compare in a quantitative manner the affinities between variants and relevant domains of PfEMP1. As discussed, these may simply be too weak to measure but further efforts to explore other immobilisation conditions/tactics and the use of different recombinant constructs of PfEMP1 fragments would be justified. These could also lead to co-crystallisation based structural studies that could help in design of rosette-disrupting therapeutics.

3**)** To test these polymorphic constructs with newly found binding partners of CR1 such as bacterial, viral and parastic proteins will possibly help identify the possible effect or advantage of this distinctive polymorphism on particular population with which they are associated.

**4)** Given the high similarity between CCPs 1-3 and CCPs 15-17, a "substitution by homology" mutagenesis approach could be used (in which residues from functional ite 2 are individually replaced with equivalents from functional site 1) in CR1 15-25 constructs to obtain gain-of-function mutations and hence find critical residues for Rh4 binding.

# BIBLIOGRAPHY

## Bibliography


1. Adams J H, Sim B K, Dolan S A, Fang X, Kaslow D C, Miller L H.(1992)."A family of erythrocyte-binding proteins of malaria parasites." *Proc Natl Acad Sci USA 89*, 7085–7089.

2. Ahearn J M and Fearon D T.(1989)."Structure and function of the complement receptors, CR1 (CD35) and CR2 (CD21)". *Adv. Immunol. 46*: 183.

3. Aidoo M, Terlouw D J, Kolczak M S, McElroy P D, ter Kuille F O and Kariuki S *et al*., (2002)."Protective effects of the sickle cell gene against malaria morbidity and mortality." *Lancet* Vol. 359, 1311-1342.

4. Allison AC.(1956)."The sickle-cell and Haemoglobin C genes in some African populations". *Ann. Human Genet*. 21: 67-89.

5. Arnot D E, Cavanagh D R, Remarque E J, Creasey A M, Sowa M P, Morgan W D, Holder A A, Longacre S and Thomas A W, (2001). "Comparative testing of six antigen-based malaria vaccine candidates directed toward merozoite-stage Plasmodium falciparum." *Parasitology* 123(Pt 2), 113-123.

6. Awandare G A, Spadafora C, Moch J K, Dutta S, Haynes J D and Stoute J A (2011). "Plasmodium falciparum field isolates use complement receptor 1 (CR1) as a receptor for invasion of erythrocytes." *Mol Biochem Parasitol* 177, 57-60.

7. Barlow P N, Baron M, Norman D G, Day A J, Willis A C, Sim R B and Campbell I D (1991). "Secondary structure of a complement control protein module by two-dimensional 1H NMR." *Biochemistry* 30, 997-1004.

8. Barragan A, Kremsner PG Wahlgren M and Carson J. (2000). "Bloob group A antigen is a co receptor in Plasmodium falciparum rosetting".*Infect Immun* 68: 2971-3599.

9. Baruch D L, Pasloske B L, Singh H B, Bi X, Ma X C and Feldman M *et al*, (1995). "Cloning the P. falciparum gene encoding PfEMP1, a malaria variant antigen and adherence receptor on the surface of parasitized human erythrocytes." *Cell* Vol. 82, 77–87.

10. Bellamy R, Kwiatkowski D and Hill A V. (1998). "Absence of an association between intercellular adhesion molecule 1, complement receptor 1 and interleukin 1 receptor antagonist gene polymorphisms and severe malaria in a West African population." *Trans R Soc Trop Med Hyg* Vol. 92, 312-316.

# BIBLIOGRAPHY


11. Bergfors T M (1999). "Protein Crystallization". International University Line, La Jolla, CA.

12. Berne B J and Pecora R (2000). "DynamicLight Scattering with Applications to Biology, Chemistry, and Physics". Mineola, NY: Dover Publications.

13. BeutlerE (1994). "G6PD deficiency". *Blood 84*: 3613–36.

14. Biggs B A and Brown G V (2001). "Malaria In: Principles and Practice of Clinical Parasitology", 53-98. Edited by Gillespie S. and Pearson, R.D. John Wiley & Sons Ltd, USA.

15. Birmingham D J, (1995). "Erythrocyte complement receptors." *Crit Rev Immunol* 15, 133-154.

16. Birmingham D J and Hebert L A (2001). "CR1 and CR1-like: the primate immune adherence receptors". *Immunological Reviews*, Vol. 180: 100-111.

17. Blein S, Ginham R, Uhrin D, Smith B O, Soares D C, Veltel S, McIlhinney R A, White J H, Barlow P N. (2004). "Structural analysis of the complement control protein (CCP) modules of GABA(B) receptor 1a: only one of the two CCP modules is compactly folded". *J Biol Chem*, 48292-306.

18. Blom A M, Villoutreix B O, Dahlback B (2004). *Molecular Immunolgy* 40: 1333-1346.

19. Bodenhausen G and Ruben D J. (1980). "Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy". *Chem Phys Lett*. 69:185.

20. Breman J G (2001). "The ears of the hippopotamus: manifestations, determinants and estimates of the malaria burden". *Am J Trop Med. 64 (Suppl 1-2)*: 1-11.

21. Camus D and Hadley T J. (1985). "A *Plasmodium falciparum* antigen that binds to host erythrocytes and merozoites". *Science. Nov* 1; 230(4725):553-6.

22. Carlson J, Helmby H, Hill A V S, Brewster D, Greenwood B M and Wahlgren M. (1990). "Human cerebral malaria: association with erythrocyte resetting and lack of anti-rosetting antibodies". *Lancet 336*:1457-1460.

23. Carlson J and Wahgren M (1992). "*Plasmodium falciparum* erythrocyte resetting is mediated by promiscuous lectin- like interactions". *J Exp Med 176*: 1311-1317.

24. Cereghino J L and Cregg J M (2000). "Heterologous protein expression in the methylotrophic yeast Pichia pastoris". *FEMS Microbiol Rev*. Jan;24 (1):45-66.

25. Chen Q, Schlichtherle M and Wahlgren M. (2000). "Molecular aspect of severe malaria". *Clin. Microbiol. Rev. 13*: 439-450.

26. Chen Q, Fernandez V, Sundstrom A, Schlichtherle M, Datta S and Hagblom P *et al.*, (1998). "Developmental selection of var gene expression in Plasmodium falciparum." *Nature* Vol. 394, 392–395

# BIBLIOGRAPHY


27. Clare J J, Rayment F B, Ballantine S P, Sreekrishna K, Romanos M A.(1991). "High-level expression of tetanus toxin fragment C in Pichia pastoris strains containing multiple tandem integrations of the gene". *Biotechnology* (N Y). May;9(5):455-60.

28. Clare J J, Romanos M A, Rayment F B, Rowedder J E, Smith M A, Payne M M, Sreekrishna K, Henwood C A. (1991)."Production of mouse epidermal growth factor in yeast: high-level secretion using Pichia pastoris strains containing multiple gene copies". *Gene*. Sep 15;105(2):205-12.

29. Clarke and Cyril A. (1964). "Genetics for the Clinician". Oxford: Blackwell

30. Clough B, Atilola F A, Black J and Pasvol G.(1998). "Plasmodium. Falciparum : importance of IgM in the rosetting of parasite - infected erythrocytes". *Exp Parasitol 89*,: 129-132.

31. Cockburn I A, Mackinnon M J, O'Donnell A, Allen S J, Moulds J M and Baisor M *et al*., (2004). "A human complement receptor 1 polymorphism that reduces Plasmodium falciparum rosetting confers protection against severe malaria." *Proc Natl Acad Sci (USA)* Vol. 101, No. 1, 272-277.

32. Cohen J H, Caudwell V, Levi-Strauss M, Bourgeois P and Kazatchkine M D, (1989). "Genetic analysis of CR1 expression on erythrocytes of patients with systemic lupus erythematosus." *Arthritis Rheum* Vol. 32, 393-397.

33. Cooke G S and Hill A V S (2001).Genetics of susceptibility to human infectious disease. *Nat. Rev. Genet. 2*: 967–977 .

34. Cooper N R (1998). Complement and Viruses. In the Human complement in health and diseases ( Volanakis J E F and Frank M M, eds). 393 -408.

35. Corrigan R A and Rowe J A (2010). "Strain variation in early innate cytokine induction by Plasmodium falciparum." *Parasite Immunol* 32, 512-527.

36. Cowman A F and B S Crabb. 2006. "Invasion of red blood cells by malaria parasites". *Cell 124*: 755–766.

37. Cregg J M, Cereghino J L, Shi J, Higgins D R (2000). "Recombinant protein expression in Pichia pastoris". *Mol Biotechnol. Sep*;16(1):23-52.

38. Cregg J M, Vedvick T S, Raschke W C. (1993). "Recent advances in the expression of foreign genes in Pichia pastoris". *Biotechnology* (N Y). Aug;11(8):905-10.

39. Crow J. (1993). "Felix Bernstein and the first human marker locus". *Genetics* 133 1, 4-7.

# BIBLIOGRAPHY


40. Cserti C M and Dzik W H. (2007). "The ABO blood group system and plasmodium falciparum malaria." *Blood* Vol.110, No. 7, 2250-2258.

41. da Silva R P, Hall B F, Joiner K A and Sacks D L (1989). *J. Immunol.*, 143, 617-622.

42. D'arcy A. (1994). "Crystallizing proteins: a rational approach?" *Acta Cryst.* D50, 469-471.

43. Daniels GL, Cartron JP, Fletcher A, Garratty G, Henry S, Jørgensen J, Judd WJ, Levene C, Lin M, Lomas-Francis C, Moulds JJ, Moulds JM, Moulds M, Overbeeke M, Reid ME, Rouger P, Scott M, Sistonen P, Smart E, Tani Y, Wendel S, Zelinski T. (2003). "International Society of Blood Transfusion Committee on terminology for red cell surface antigens". Vancouver Report. *Vox Sang. Apr*;84(3):244-7.

44. Dantelsson C, Pascual M, French L, Steiger G and Schifferit J A, (1994). "Soluble complement receptor type 1 (CD35) is released from leucocytes by surface cleavage." *Eur J Immunol* Vol. 24, 2725-2731

45. Dahlbäck M, Jørgensen L M, Nielsen M A, Clausen T M, Ditlev S B, Resende M, Pinto V V, Arnot D E, Theander T G and Salanti A, (2011). "The chondroitin sulfate A-binding site of the VAR2CSA protein involves multiple N-terminal domains." *J Infect Dis* 203(11), 1679-1685.

46. Dervillez X, Oudin S and Libyh M T *et al*, (1997). "Catabolism of the human erythrocyte C3b/C4b receptor (CR1/CD35): vesiculation and /or proteolysis?" *Immunopharmacology* Vol. 38, 129-140.

47. Dobson N J, Lambris J D and Ross G D, (1981). "Characteristics of isolated erythrocyte complement receptor type one (CR1, C4b-C3b receptor) and CR1-specific antibodies." *J Immunol* Vol. 126, 693-698.

48. Duraisingh M T, Maier A G, Triglia T and Cowman A F. (2003a). "Erythrocyte-binding antigen 175 mediates invasion in *Plasmodium falciparum* utilizing sialic acid-dependent and -independent pathways". *Proc. Natl. Acad. Sci. USA 100*, 4796–4801.

49. Doumbo O K, Thera M A, Kone A K, Raza A, Tempest L J, Lyke K E, Plowe C V and Rowe J A (2009). "High levels of Plasmodium falciparum rosetting in all clinical forms of severe malaria in African children." *Am J Trop Med Hyg* 81, 987-993.

50. Dykman T R, Cole J L, Iida K and Atkinson J P, (1983). "Polymorphism of human erythrocyte C3b/C4b receptor." *Proc Natl Acad Sci (USA)* Vol. 80, 1698-1702.

# BIBLIOGRAPHY


51. Fearon D T. (1980) "Identification of the membrane glycoprotein that is the C3b receptor of the human erythrocyte, polymorphonuclear leukocyte, B lymphocyte and monocyte." *J. Exp. Med. 152*: 20-30.

52. Fearon D T, Klickstein L B, Wong W W, Wilson J G, Moore F D Jr, Weis J J, Weis J H, Jack R M, Carter R H and Ahearn J A (1989). "Immunoregulatory functions of complement: structural and functional studies of complement receptor type 1 (CR1; CD35) and type 2 (CR2; CD21)." *Prog Clin Biol Res* 297, 211-220.

53. Fearon D T (1979). "Regulation of the amplification C3 convertase of human complement by an inhibitory protein isolated from human erythrocyte membrane." *Proc Natl Acad Sci (USA)* Vol. 76, 5867-5871.

54. Farries T C, Seya T, Harrison R A and Atkinson J P (1990). "Competition for binding sites on C3b by CR1, CR2, MCP, factor B and factor H." *Complement Inflamm* Vol. 7, 30-41

55. Ferre-D'Amare A and Burley S (1994). "Use of dynamic light scattering to assesscrystallizability of macromolecules and macromolecular assemblies." *Structure* 2, 357-359.

56. Ford E B.( 1973) "Genetics for Medical Students (7th ed.)" *London: Chapman & Hall.*(1942).

57. Franke D and Svergun D I. (2009). "DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering." *J. Appl. Cryst.* 42, 342-346.

58. Fried M and Duffy P E. (1996). "Adherence of Plasmodium falciparum to chondroitin sulfate A in the human placenta." *Science* 272, 1502–1504.

59. Furtado P B, Huang C Y, Ihyembe D, Hammond R A, Marsh H C, Perkins S J. (2008). "The partly folded back solution structure arrangement of the 30 SCR domains in humancomplement receptor type 1 (CR1) permits access to its C3b and C4b ligands." *J Mol Biol*. 375:102-18.

60. Ghiran I, Barbashov S F, Klickstein L B, Tas S W, Jensenius J C, Nicholson-Weller A. (2000) "Complement Receptor 1/CD35 is a receptor for mannan-binding lectin." *J. Exp. Med. 192*: 1797-1808.

61. Ghumra A, Khunrae P, Ataide R, Raza A, Rogerson S J, Higgins M K and Rowe J A, (2011). "Immunisation with recombinant PfEMP1 domains elicits functional rosette-inhibiting and phagocytosis-inducing antibodies to Plasmodium falciparum." *PLoS One* 6, e16414.

# BIBLIOGRAPHY


62. Gibson N C and Waxman F J, (1994). "Relationship between immune complex binding and release and the quantitative expression of the complement receptor type 1 (CR1, CD35) on human erythrocytes." *Clin Immunol* Vol. 70, 104.

63. Gilles H M. (1993). Historical Outline In: Bruce-Chwatt's Essential Malariology. 3rd Edition, The Bath Press, UK.

64. Grzesiek S and Bax A. (1992). "Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR". *J. Am. Chem. Soc.* 114:6291.

65. Grzesiek S and Bax A. (1992). "An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins". *J. Magn. Reson*. 99:201.

66. Grzesiek S and Bax A. (1993). "Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins." *J Biomol Nmr* 3, 185-204.

67. Guinier A (1939) "*Annals of Physics (Paris)* " 12, 161-237

68. Hamer I, Paccaud J P, Belin D, Maeder C and Carpenter J L, (1998). "Soluble form of complement C3b/C4b receptor (CR1) results from a proteolytic cleavage in the C-terminal region of CR1 transmembrane domain." *Biochem J* Vol. 329, 183-190.

69. Handunnetti S M, Van Schravendijk M R, Hasler T,Barnwell J W, Greenwalt D E and Howard R J.(1992) "Involvement of CD36 on erythrocytes as a rosetting receptor in *Plasmodium falciparum* -infected erythrocytes. ." *Blood 80*: 2097-2104.

70. Hanscheid T, (1999). "Diagnosis of malaria: a review of alternatives to conventional microscopy." *Clin Lab Haematol* Vol. 21, 235-245.

71. Harding S E, Settelle D B and Bloomfield V A. (1992). "Laser Light Scattering in Biochemistry". Royal Society of Chemistry, Cambridge.

72. Heddini A, Pettersson F, Kai O, Shafi J, Obiero J, Chen Q, Barragan A, Wahlgren M, Marsh K. (2001)."Fresh isolates from children with severe *Plasmodium falciparum* malaria bind to multiple receptors." *Infect Immun.*,69(9):5849-56.

73. Helgeson M, Swanson J and Polesky H F. (1970). "Knops-Helgeson (Kna), a high frequency erythrocyte antigen". *Transfusion, Vol. 10: 137-8.*

74. Herrera A H, Xiang L, Martin S G, Lewis J, Wilson J G. (1998) "Analysis of complement receptor type 1 (CR1) expression on erythrocytes and of CR1 allelic markers in Caucasian and African American populations." *Clin Immunol Immunopathol*,176-183.

# BIBLIOGRAPHY


75. Ho M and White N J, (1999). "Molecular mechanisms of Cytoadherence in malaria." *Am J Physiol Cell Physiol* Vol. 276, 1231-1242.

76. Holding P A and Snow R W., (2001). "Impact of Plasmodium falciparum malaria on performance and learning: review of the evidence." *Am J Trop Med Hyg* Vol. 64, (Suppl): 68-75.

77. Holers V M, Chaplin D D, LeyKam J F, Gruner B A, Kumar V and Atkinson J P, (1987). "Human complement C3b/C4b receptor (CR1) mRNA polymorphism that correlates with the CR1 allelic molecular weight polymorphism." *Proc Natl Acad Sci (USA)* Vol. 84, 2459-2463.

78. Hviid L (1998). "Clinical disease, immunity and protection against Plasmodium falciparum malaria in populations living in endemic areas." *Molecular Medicine*, 1-10.

79. Iida K and Nussenzweig V (1981). "Complement receptor is an inhibitor of the complement cascade." *J Exp Med* Vol. 153, 1138-1150.

80. Jallow M, Teo Y Y, Small K S, Rockett K A, Deloukas P and Clark T G *et al*, (2009). "Genome-wide and fine-resolution association analysis of malaria in West Africa." *Nat Genet* 41, 657-665.

81. Jha P and Kotwal G J. (2003). "Vaccinia complement control protein: multi-functional protein and a potential wonder drug". *J. Biosci*. 28, 265-271.

82. Joergensen L M, Salanti A, Dobrilovic T, Barfod L, Hassenkam T, Theander T G, Hviid L and Arnot DE, (2009 ). "The kinetics of antibody binding to Plasmodium falciparum VAR2CSA PfEMP1 antigen and modelling of PfEMP1 antigen packing on the membrane knobs." *PLoS One* 19;4(8), e6667.

83. Joergensen L, Bengtsson D C, Bengtsson A, Ronander E, Berger S S, Turner L, Dalgaard M B, Cham G K, Victor M E, Lavstsen T, Theander T G, Arnot D E and Jensen A T, (2010 ). "Surface co-expression of two different PfEMP1 antigens on single plasmodium falciparum-infected erythrocytes facilitates binding to ICAM1 and PECAM1." *PLoS Pathog* Sep 2;6(9), e1001083.

84. Kalli K R, Hsu P, Bartow T J, Ahearn J M, Matsumoto A K and Klickstein L B *et al*, (1991). "Mapping of the C3b-binding site CR1 and construction of a (CR1)2-F(ab')2 chimeric complement inhibitor." *J Exp Med* Vol. 174, 1451-1460.

85. Kazatchkine M D, Jouvin M H, Wilson J G, Fischer E and Fischer A, (1987). "Human diseases associated with C3 receptor deficiencies." *Immunol Lett.* Vol. 14, 191.

86. Kaul D K, Roth-Jr E F, Nagel R L, Howard R J and Handunnetti S M.(1991)" Rosetting of *Plasmodium falciparum*-infected red blood cells with uninfected red

blood cells enhances microvascular obstruction under flow conditions." *Blood 78*: 812-819.

87. Kirkitadze M D, Dryden D T, Kelly S M, Price N C, Wang X, Krych M, Atkinson J P and Barlow P N, (1999). "Co-operativity between modules within a C3b-binding site of complement receptor type 1." *FEBS Lett* 459, 133-138.

88. Kirkitadze M D, Krych M, Uhrin D, Dryden D T, Smith B O, Cooper A, Wang X, Hauhart R, Atkinson J P and Barlow P N, (1999). "Independently melting modules and highly structured intermodular junctions within complement receptor type 1." *Biochemistry* 38, 7019-7031.

89. Kirkitadze M D and Barlow P N. (2001) "Structure and flexibility of the multiple domain proteins that regulate complement activation." *Immunol Rev.180*: 146-61.

90. Klickstein L B, Barbashov S F, Liu T, Jack R M, Nicholson-Weller A. (1997) "Complement receptor type 1 (CR1, CD35) is a receptor for C1q." Immunity: 345-55.

91. Klickstein L B, Wong W W, Smith J A, Weis J H, Wilson J G and Fearon DT. (1987). "Human C3b/C4b receptor (CR1). Demonstration of long homologous repeating domains that are composed of the short consensus repeats characteristics of C3/C4 binding proteins". J Exp Med, Vol. 165:1095–1112.

92. Klotz F W, Orlandi P A, Reuter G, Cohen S J, Haynes J D, Schauer R, Howard R J, Palese P and Miller L H. (1992). "Bind-ing of *Plasmodium falciparum* 175-kilodalton erythrocyte binding antigen and invasion of murine erythrocytes requires N-acetylneu-raminic acid but not its O-acetylated form". *Mol. Biochem. Parasitol.* 51, 49–54.

93. Koch M H J and Bordas J. (1983) *Nucl Instrum Methods* 208, 461-469

94. Konarev P V, Volkov V V, Sokolova A V, Koch M H J and Svergun D I, (2003). "PRIMUS: a Windows PC-based system for small-angle scattering data analysis." *Appl Crystallogr* 36, 1277-1282.

95. Koolman J and Rohm K-H, (1998)."Colour Atlas of Biochemistry", Thieme, Stuttgart,

96. Korb L C and Ahearn J M. (1997). "C1q binds directly and specifically to surface blebs of apoptotic human keratinocytes." *J. Immunol.158:4525*.

97. Kozin M B and Svergun DI, (2001). "Automated matching of high- and low-resolution structural models." *J. Appl. Crystallogr* 34, 33-41.

98. Krych M, Atkinson J P and Holers V M, (1992). "Complement receptors." *Curr Opin Immunol* 4, 8-13

# BIBLIOGRAPHY


99. Krych M, Hauhart R and Atkinson JP (1998). "Structure-function analysis of the active sites of complement receptor type 1". *J Biol Chem, Vol. 273*: 8623-8629.

100. Krych M , Hourcade D and Atkinson J P (1991) "Sites within the complement C3b/C4b receptor important for the specificity of ligand binding." *Proc Natl. Acad. Sci. USA 88*: 4353.

101. Krych-Goldberg M, Moulds J M and Atkinson J P. (2002). "Human complement receptor type 1 (CR1) binds to a major malarial adhesin". *Trends Mol Med.* Nov;8(11):531-7.

102. Krych M, Clemenza L, Howdeshell D, Hauhart R, Hourcade D and Atkinson J P, (1994). "Analysis of the functional domains of complement receptor type 1 (C3b/C4b receptor; CD35) by substitution mutagenesis." *J Biol Chem* 269, 13273-13278.

103. Krych-Goldberg M, Hauhart R E, Subramanian V B, Yurcisin B M, Crimmins DL and Hourcade D E *et a*l, (1999). "Decay accelerating activity of complement receptor type 1 (CD35). Two active sites are required for dissociating C5 convertases." *J Biol Chem* Vol. 274, 31160-31168

104. Krych-Goldberg M, Hauhart R E, Porzukowiak T and Atkinson J P, (2005). "Synergy between two active sites of human complement receptor type 1 (CD35) in complement regulation: implications for the structure of the classical pathway C3 convertase and generation of more potent inhibitors." *J Immunol* 175, 4528-4535.

105. Krych-Goldberg M and Atkinson J P. (2001) " Structure-function relation of complement receptor type 1." Immunol. Rev. 180: 112-122.

106. Lambros C and Vanderberg J P. (1979). "Synchronisation of *Plasmodium falciparum* erythrocytic stages in culture". Journal of Parasitology 65, 418-420

107. Laue T M, Shah B D, Ridgeway T M and Pelletier SL (1992). "Computer-aided interpretation of analytical sedimentation data for proteins". *In Analytical ultracentrifugation in biochemistry and polymer science*. (eds. S.E. Harding et al.)pp. 90–125,The Royal Society of Chemistry,Cambridge, UK.

108. Lavstsen T, Salanti A, Jensen A T, Arnot D E and Theander T G, (2003). "Sub-grouping of Plasmodium falciparum 3D7 var genes based on sequence analysis of coding and non-coding regions." *Malar J* Sep 10;2, 27. Epub 2003 Sep 2010

109. Law S K A. and Reid K B M. (1995). “Complement” second edition. Textbook published by Oxford University Press.

110. Leninger A L, Nelson D L and Cox M M. (2004). "Leninger–Principles of Biochemistry". W.H. Freeman and Co. Inc. NY.

# BIBLIOGRAPHY


111. Lindahl G, Sjobring U and Johnsson E, (2000). "Human complement regulators: a major target for pathogenic microorganisms." *Curr Opin Immunol* Vol., 12, 44-51.

112. Liszewski M K, Farries T C, Luslin D M, Rooney I A and Atkinson J P, (1996). "Control of the complement system." *Adv Immunol* Vol. 61, 201-283.

113. Ma Y G, Cho M Y, Zhao M, Park J W, Matsushita M, Fujita T, Lee B L.(2004). "Human mannose-binding lectin and L-ficolin function as specific pattern recognition proteins in the lectin activation pathway of complement". *J Biol Chem.* Jun 11;279(24):25307-12.

114. Mackintosh C L, Beeson J G and Marsh K, (2004). "Clinical features and pathogenesis of severe malaria." *Trends in Parasitology* Vol. 20, 597–603.

115. Maier A G, Duraisingh M T, Reeder J C, Patel S S, Kazura J W, Zimmerman P A and Cowman A F. (2003) "*Plasmodium falciparum* erythrocyte invasion through glycophorin C and selection for Gerbich negativity in human populations." *Nat Med 9*: 87–92.

116. Mallin Rosie L (2003). PhD Thesis; "A Structural Study of the C3b-Binding Site of Complement Receptor Type I (CD 35)". University of Edinburgh, Edinburgh.

117. Walter H. Manning (2010). "Clinical Decision Making in Fluency Disorders". 3rd Edit .Cengage Learning. p 125.

118. Marchesi V T, Tillack T W, Jackson R L, Segrest J P, Scott R E. (1972) " Chemical characterization and surface orientation of the major glycoprotein of the human erythrocyte membrane." *Proc Natl Acad Sci U S A. Jun;69(6)*: 1445-9.

119. Mayer D C, Cofie J, Jiang L., Hartl D. L., Tracy E, Kabat J, Mendoza L. H. and Miller L H (2009). "Glycophorin B is the erythrocyte receptor of *Plasmodium falciparum* erythrocyte-binding ligand, EBL-1 ." *Proc Natl Acad Sci USA 106*: 5348–5352.

120. Meade S M and Earickson R J (2005). Medical Geography. Guilford

121. Medof M E and Nussenzweig V, (1984). "Control of the function of substrate-bound C4b-C3b by the complement receptor." *J Exp Med* Vol. 159, 1669-1685.

122. Medof M E, Prince G M and Mold C, (1982). "Release of soluble immune complexes from immune adherence receptors on human erythrocytes is mediated by C3b inactivator independently of Beta 1H and is accompanied by generation of C3c." *Proc Natl Acad Sci USA* Vol. 79, 5047-5051

123. Miller L H, Baruch D I, Marsh K and Doumbo K O. (2002) "The pathogenic basis of malaria." *Nature 415*: 673-679.

# BIBLIOGRAPHY


124. Miller L H. (1994). "Impact of malaria on genetic polymorphism and genetic diseases in Africans and African Americans". *Proc Natl Acad. Sci., USA, Vol. 91: 2415-2419.*

125. Miller L H, Good M F and Milon G, (1994). "Malaria Pathogenesis." *Science* Vol. 264, 1878-1883

126. Molthan L. (1983). "Expansion of the York, Cost, Mccoy, Knops blood group system: the new Mccoy antigens McCc and McCd". *Med Lab Sci. Apr*;40(2):113-21.

127. Molthan L and Giles C M. (1975). A new antigen Yk-a (York), and its relationship to Cs-a (Cost). *Vox Sang.*;29(2):145-53.

128. Molthan L and Moulds J. (1978). "A new antigen, McCa (McCoy), and its relationship to Kna (Knops). " *Transfusion*, Vol. 18: 566-8.

129. Morgan P B and Harris C L (1999). "Complement Regulatory Proteins", Textbook published by Academic Press.

130. Moulds J M, (2002). "A review of the Knops blood group: separating fact from fallacy." *Immunohematology* 18, 1-8.

131. Moulds J M, (2010). "A The Knops blood-group system." *Immunohematology* 18, 1-8.

132. Moulds J M, Thomas BJ, Doumbo O, Diallo D A, Lyke K E and Plowe C V, (2004). "Identification of the $Kn^a/Kn^b$ polymorphism and a method for Knops genotyping."*Tranfusion, Vol. 44: 164-9.*

133. Moulds J M, Zimmerman P A, Doumbo O K, Kassambara L, Sagara I, Diallo D A, Atkinson J P and Krych-Goldberg M. (2001). "Molecular identification of Knops blood group polymorphisms found in long homologous region D of complement receptor 1". *Blood, Vol. 97*: 2879-2885.

134. Moulds J M, Kassambara L, Middleton J J, Baby M, Sagara I, Guindo A, Coulibaly S, Yalcouye D, Diallo D A, Miller L and Doumbo O, (2000). "Identification of complement receptor one (CR1) polymorphisms in West Africa." *Genes and Immunity* Vol. 1, 325-329.

135. Moulds J M, Nickells M W, Moulds J J, Brown M C and Atkinson J P. (1991). "The C3b/C4b receptor is recognized by the Knops, McCoy, Swain-langley, and York blood group antisera." *J Exp Med* : 1159-1163.

136. Munson L G, Scott M E, Landay A L and Spear G T. (1995). "Decreased levels of complement receptor 1 (CD35) on B lymphocytes in persons with HIV infection." *Clin Immunol Immunopathol* Vol. 75, 20-25.

# BIBLIOGRAPHY


137. Newbold C I, Warn P, Black G, Berendt A, Craig A and Snow B *et al*, (1997). "Receptor-specific adhesion and clinical disease in Plasmodium falciparum." *Am. J. Trop. Med. Hyg* Vol. 57, 389-398.

138. Newbold C, Craig A, Kyes S, Rowe A, Fernandez-Reyes D and Fagan. (1999). " Cytoadherence, pathogenesis and the infected red cell surface in *Plasmodium falciparum*." *Int J Parasitol 29*: 927-937.

139. Noedl H, Se Y, Schaecher K, Socheat D and Fukuda M M, (2008). "Evidence of Artemisinin-Resistant Malaria in Western Cambodia." *New Eng. J. Med.* Vol. 359, No. 24 (Correspondence), 2619-2620.

140. Norman D G, Barlow P N, Baron M, Day A J, Sim R B and Campbell I D, (1991). "Three-dimensional structure of a complement control protein module in solution." *J Mol Biol* 219, 717-725.

141. Noumsi G T, Tounkara A, Diallo H, Billingsley K, Moulds J J and Moulds J M, (2011). "Knops blood group polymorphism and susceptibility to Mycobacterium tuberculosis infection." *Transfusion* 51, 2462-2469.

142. Orlandi P A, Klotz F W, Haynes J D. (1992) "A malaria invasion receptor, the 175-kilodalton erythrocyte binding antigen of *Plasmodium falciparum* recognizes the terminal Neu5Ac(alpha 2-3)Gal- sequences of glycophorin A." *J Cell Biol. Feb;116(4)* : 901-9.

143. Pascual M, Duchosal M A, Steiger G, Giostra E, Pechere A and Paccaud J P *et al*., (1993). "Circulating soluble CR1 (CD35). Serum levels in diseases and evidence for its release by human leukocytes." *J Immunol* Vol. 151, 1702-1711.

144. Pasvol G, Wilson R J M, Smalley M E and Brown J. (1978). "Separation of viable schizont-infected red cells of Plasmodium falciparum from human blood "*.Annals of Tropical Medicine and Parasitology* 72, 87-88.

145. Pecora, R. 1985. Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy. New York, NY: Plenum Press.

146. Persson K E, McCallum F J, Reiling L, Lister N A, Stubbs J, Cowman AF, Marsh K and Beeson J G. (2008). "Variation in use of erythrocyte pathways by *Plasmodium falciparum* mediates evasion of human inhibitory antibodies". *J. Clin. Investig.* 118:342-351.

147. Peterson D S and Wellems T E. (2000). "EBL-1, a putative erythrocyte-binding protein of *Plasmodium falciparum*, maps within a favored linkage group in two genetic crosses." *Mol Biochem Parasitol 105*: 105–113.

# BIBLIOGRAPHY


148. Pinto V V, Salanti A, Joergensen L M, Dahlbäck M, Resende M, Ditlev S B, Agger E M, Arnot D E, Theander T G and Nielsen M A, (2010 ). "The effect of adjuvants on the immune response induced by a DBL4ε-ID4 VAR2CSA based Plasmodium falciparum vaccine against placental malaria." *Malar J* 19;9, 100.

149. Ramasamy R, (1998). "Molecular basis for evasion of host immunity and pathogenesis in malaria." *Biochimica et Biophysica Acta* Vol. 1406, 10-27.

150. Rao N, Ferguson D J, Lee S F and Telen M J. (1991). " Identification of human erythrocyte blood group antigens on the C3b/C4b receptor." *J Immunol* : 3502-3507.

151. Rayner J C, Vargas-Serrato E, Huber C S, Galinski M R, Barnwell J W. (2001) "A *Plasmodium falciparum* homologue of Plasmodium vivax reticulocyte binding protein (PvRBP1) defines a trypsin-resistant erythrocyte invasion pathway." *J Exp Med 194*: 1571–1581.

152. Rayner J C, Galinski M R, Ingravallo P, Barnwell J W. (2000). "Two *Plasmodium falciparum* genes express merozoite proteins that are related to Plasmodium vivax and Plasmodium yoelii adhesive proteins involved in host cell selection and invasion". *Proc Natl Acad Sci U S A.* Aug 15;97(17):9648-53.

153. Reid M E. (2004). "The blood group antigen".Facts Book. London, Elsevier Academic Press.

154. Ricklin D, Hajishengallis G, Yang K and Lambris J D, (2010). "Complement: a key system for immune surveillance and homeostasis." *Nat Immunol* 11, 785-797.

155. Roessle M W, Klaering R, Ristau U, Robrahn B, Jahn D, Gehrmann T, Konarev P, Round A, Fiedler S, Hermes C and Svergun D, (2007). "Upgrade of the small-angle X-ray scattering beamline X33 at the European Molecular Biology Laboratory, Hamburg." *J Appl Crystallogr* 40, S190-S194

156. Roitt I, Brostoff J. and Male D. (1998). Immunology fifth edition. Textbook published by Mosby International Limited.

157. Ross G D, Lambris J D, Cain J A and Newman S L, (1982). "Generation of three different fragments of bound C3 with purified factor I or serum. Requirements for factor H vs CR1 cofactor activity." *J Immunol* Vol. 129, 2051-2060.

158. Ross R (1897). "On some peculiar pigmented cells found in two mosquitoes fed on malarial blood." *Br Med* Vol. J 2, 1786–1788.

159. Roversi P, Johnson S, Caesar J J, McLean F, Leath K J, Tsiftsoglou S A, Morgan B P, Harris C L, Sim R B and Lea S M, (2011). "Structural basis for

complement factor I control and its disease-associated sequence polymorphisms." *Proc Natl Acad Sci U S A* 108, 12839-12844.

160. Rowe J A, Raza A, Diallo D A, Baby M, Poudiougo B, Coulibaly D, Cockburn I A, Middleton J, Lyke K E, Plowe C V, Doumbo O K, Moulds J M. (2002). "Erythrocyte CR1 expression level does not correlate with a HindIII restriction fragment length polymorphism in Africans; implications for studies on malaria susceptibility." *Genes Immun*: 497-500.

161. Rowe J A, Rogerson S J, Raza A, Moulds J M, Kazatchkine M D, Marsh Kevin, (2000). "Mapping of the region of complement receptor (CR) 1 required for *Plasmodium falciparum* rosetting and demonstration of the importance CR1 in rosetting in field isolates." *J.Immunol. 165*: 6341-6346.

162. Rowe A, Moulds J M, Newbold C L and Miller L H. (1997). "*P. falciparum* rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1." *Nature 388*: 292-295.

163. Rowe J A, Obeiro J, Newbold C I and Marsh K, (1995). " *Plasmodium falciparum* rosetting is associated with malaria severity in Kenya." *Infect. Immun. 63*: 2323-2326.

164. Rowe J A, Opi D H and Williams T N, (2009). "Blood groups and malaria: fresh insights into pathogenesis and identification of targets for intervention." *Curr Opin Hematol* 16, 480-487.

165. Ruuska P E, Ikaheimo L, Silvennoinen-Kassinen S and T. A. Kaar M L, . (1992). "Normal C3b receptor (CR1) genomic polymorphism in patients with insulin-dependent diabetes mellitus (IDDM): is the erythrocyte CR1 expression an acquired phenomenon?" *Clin Exp Immunol* Vol. 89, 18.

166. Sachs J and Malaney P, (2002). "The economic and social burden of malaria." *Nature* Vol. 415, 680-685.

167. Salanti A, Jensen A T, Zornig H D, Staalsoe T, Joergensen L, Nielsen M A, Khattab A, Arnot DE, Klinkert M Q, Hviid L and Theander T G, (2002). "A sub-family of common and highly conserved Plasmodium falciparum var genes." *Mol Biochem Parasitol* Jun;122(1), 111-115.

168. Salanti A, Staalsoe T, Lavstsen T, Jensen A T, Sowa M P, Arnot D E, Hviid L and Theander T G, (2011). "Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering Plasmodium falciparum involved in pregnancy-associated malaria." *Vaccine*, Nov 25. [Epub ahead of print].

169. Sander W. Tas, Lloyd B. Klickstein, Sergei F. Barbashov and Anne Nicholson-Weller. (1999). "C1q and C4b Bind Simultaneously to CR1 and Additively Support Erythrocyte Adhesion". *The Journal of Immunology*, 163: 5056-5063

# BIBLIOGRAPHY


170. Sander A F, Salanti A, Lavstsen T, Nielsen M A, Magistrado P, Lusingu J, Ndam N T and Arnot D E, (2008). "Multiple var2csa-type PfEMP1 genes located at different chromosomal loci occur in many Plasmodium falciparum isolates." *Malar J* 5;7, 101.

171. Sander A F, Salanti A, Lavstsen T, Nielsen M A, Theander T G, Leke R G, Lo Y Y, Bobbili N, Arnot D E and Taylor D W, (2011). "Positive selection of Plasmodium falciparum parasites with multiple var2csa-type PfEMP1 genes during the course of infection in pregnant women." *Int J Parasitol* 41(1), 71-80. Epub 2010 Sep 2017.

172. Schlesinger L S and Horwitz M A (1990). "Phagocytosis of leprosy bacilli is mediated by complement receptors CR1 and CR3 on human monocytes and complement component C3 in serum". *J. Clin. Invest.,* Vol. 85: 1304.

173. Schmidt C Q, Herbert A P,Kavanagh D, Gandy C, Fenton C J, Blaum B S, Lyon M, Uhr´ın D and Barlow P N (2008). "A new map of glycosaminoglycan and c3b binding sites on factor h". *Journal of Immunology* (Baltimore, Md.: 1950) 181: 2610–9. PMID: 18684951.

174. Schmidt C Q, Herbert A P, Mertens H D, Guariento M, Soares D C, Uhrin D, Rowe A J, Svergun D I and Barlow P N, (2010). "The central portion of factor H (modules 10-15) is compact and contains a structurally deviant CCP module." *J Mol Biol* 395, 105-122

175. Schmidt C Q, Slingsby F C, Richards A and Barlow P N, (2011). "Production of biologically active complement factor H in therapeutically useful quantities." *Protein Expr Purif* 76, 254-263.

176. Scholander C, Treutiger C T, Huultenby K and Walgren M. (1996). " Novel fibrillar structure confers adhesive properties to malaria infected erythrocytes." *Nat Med 2*: 204-208.

177. Schuck P, (2000). "Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling." *Biophys J* 78, 1606-1619

178. Sim B K, Orlandi P A, Haynes J D, Klotz F W, Carter J M, Camus D, Zegans M E, Chulay J D. (1990). "Primary structure of the 175K *Plasmodium falciparum* erythrocyte binding antigen and identification of a peptide which elicits antibodies that inhibit malaria merozoite invasion." *J Cell Biol. Nov;111 (5 Pt 1)*: 1877-84.

179. Sim B K L, Chitnis C E, Wasniowska K, Hadley T J, Miller L H. (1994). " Receptor and ligand domains for invasion of erythrocytes by *Plasmodium falciparum*." *Science 264*: 1941–1944.

# BIBLIOGRAPHY


180. Smith B O, Mallin R L, Krych-Goldberg M, Xuefeng W, Hauhart R E and Bromek K *et al.,* (2002). "Structure of the C3b Binding site of CR1 (CD35), the Immune Adherence Receptor." *Cell* Vol. 108, 769-780.

181. Soares D C, Gerloff D L, Syme N R, Coulson A F, Parkinson J and Barlow P N, (2005). "Large-scale modelling as a route to multiple surface comparisons of the CCP module family." *Protein Eng Des Sel* 18, 379-388, Epub 2005 Jun 2023.

182. Spadafora C, Awandare G A, Kopydlowski K M, Czege J, Moch J K, Finberg R W, Tsokos G C and Stoute J A. (2010). "Complement receptor 1 is a sialic acid–independent erythrocyte receptor of *Plasmodium falciparum*." *PLoS Pathog 6*: 1000968.

183. Staalsoe T, Hamad A A, Hviid L, Elhassan I M, Arnot D E and Theander T G, (2002). "In vivo switching between variant surface antigens in human Plasmodium falciparum infection." *J Infect Dis*. Sep 1;186(5), 719-722. Epub 2002 Aug 2005.

184. Stefan Wildt and Tilllman U. Gerngross. (2005). "The humanization of N-glycosylation pathways in yeast". Nature Reviews Microbiology 3, 119-128.

185. Stoute J A, Odindo A O, Owuor B O, Mibei E K, Opollo M O and Waitumbi J N, (2003). "Loss of Red Blood Cell-Complement Regulatory Proteins and Increased Levels of Circulating Immune Complexes are associated with Severe Malarial Anaemia." *JID* Vol. 187, 522-525.

186. Stoute JA, (2005). "Complement-regulatory proteins in severe malaria: too little or too much of a good thing?" *Trends Parasitol* 21, 218-223.

187. Stoute JA, (2011). "Complement receptor 1 and malaria." *Cell Microbiol*.

188. Stubbs J, Simpson K M, Triglia T, Plouffe D, Tonkin C J, Duraisingh M T, Maier A G, Winzeler E A and Cowman A F, (2005). "Molecular mechanism for switching of P. falciparum invasion pathways into human erythrocytes." *Science* 309, 1384-1387.

189. Subramanian V B, Clemenza L, Krych M and Atkinson J P, (1996). "Substitution of two amino acids confers C3b binding to the C4b binding site of CR1 (CD35). Analysis based on ligand binding by chimpanzee erythrocyte complement receptor." *J Immunol* 157, 1242-1247.

190. Su X Z, Heatwole V M., Wertheimer S P, Guinet F, Herrfeldt J A, Peterson D S, Ravetch JA and Wellems T E. (1995). " The large diverse gene family var

encodes proteins involved in cytoadherence and antigenic variation of plasmodium falciparun-infected erythrocyte." *Cell, 82*: 89-100.

191. Svergun D I, (1992). "Determination of the regularization parameter in indirect-transform methods using perceptual criteria." *J. Appl. Crystallogr* 25, 495-503.

192. Sykes, B. (1999). "The human inheritance: genes, language and evolution ." *Oxford: Oxford U. Pr.*

193. Tas S W, Klickstein L B, Barbashov S F and Nicholson-Weller A, (1999). "C1q and C4b bind simultaneously to CR1 and additively support erythrocyte adhesion." *J Immunol* 1631, 5056-5063.

194. Tedder T F, Fearon D T, Gartland G L and Cooper M D, (1983). "Expression of C3b receptors on human B cells and myelomonocytic cells but not natural killer cells". *J Immunol, 130: 370-375.*

195. Tham W H, Schmidt C Q, Hauhart R E, Guariento M, Tetteh-Quarcoo P B, Lopaticki S, Atkinson J P, Barlow P N and Cowman A F, (2011). "Plasmodium falciparum uses a key functional site in complement receptor type-1 for invasion of human erythrocytes." *Blood* 118, 1923-1933.

196. Tham W H, Wilson D W, Lopaticki S, Schmidt C Q, Tetteh-Quarcoo P B, Barlow P N, Richard D, Corbin J E, Beeson J G, and Cowman A F. (2010). "Complement receptor 1 is the host erythrocyte receptor for *Plasmodium falciparum* PfRh4 invasion ligand". *Proc Natl Acad Sci U S A.*

197. Tham W H, Wilson D W, Reiling L, Chen L, Beeson JG, and Cowman A F. (2009). "Antibodies to reticulocyte-binding protein-like homologue 4 inhibit invasion of *Plasmodium falciparum* into human erythrocytes." *Infect Immun 77*: 2427–2435.

198. Thathy V, Moulds J M, Guyah B, Otieno W and Stoute J A. (2005). "Complement receptor 1 polymorphisms associated with resistance to severe malaria in Kenya". *Malaria Journal, Vol. 4:54-61.*

199. Thieblemont N, Haeffner-Cavaillon N, Ledur A, L'Age-Stehr J, Ziegler-Heitbrock H W and Kazatchkine M D. (1993)". CR1 (CD35) and CR3(CD11b/CD 18) mediates infections of human monocytes and monocytic cell lines with complement-opsonised HIV independently of CD4." *Clin. Exp. Immunol 92*: 106-113.

200. Thomas B N, Donvito B, Cockburn I, Fandeur T, Rowe J A, Cohen J H M and Moulds J M, (2005). "A complement receptor-1 polymorphism with high frequency in malaria endemic regions of Asia but not Africa." *Genes Immun* Vol. 6, 31-36.

# BIBLIOGRAPHY


201. Treutiger C J, Hedlund I, Helmy H, Carlson J, Jepson A, and Twumasi P. (1992). "Rosette formation in *Plasmodium falciparum* isolates and anti-rosette activity of sera from Gambians with cerebral or uncomplicated malaria." *Am. J. Trop. Med. Hyg. 46*: 503-510.

202. Victor M E, Bengtsson A, Andersen G, Bengtsson D, Lusingu J P, Vestergaard L S, Arnot D E, Theander T G, Joergensen L and Jensen A T, (2010). "Insect cells are superior to Escherichia coli in producing malaria proteins inducing IgG targeting PfEMP1 on infected erythrocytes." *Malar J* Nov 15;9, 325.

203. Volkov V V and Svergun D I, (2003). "Uniqueness of ab initio shape determination in small-angle scattering." *J Appl Crystallogr* 36, 860-864.

204. Verrelli B C, McDonald J H, Argyropoulos G, Destro-Bisol G, Froment A, Drousiotou A, Lefranc G, Helal A N, Loiselet J, Tishkoff S A. (2002). "Evidence for balancing selection from nucleotide sequence analyses of human G6PD." Am J Hum Genet.71:1112-28.

205. Vranken W F, Boucher W, Stevens T J, Fogh R H, Pajon A, Llinas M, Ulrich E L, Markley J L, Ionides J and Laue E D. (2005) "The CCPN data model for NMR spectroscopy: development of a software pipeline. "*Proteins.* 59, 687-696.

206. Vuister G W and Bax A. (1992). "Resolution enhancement and spectral editing of uniformly 13C-enriched proteins by homonuclear broadband 13C decoupling". J. Magn. Reson. 98:428-435.

207. Wahlgren M, Carlson J, Helmby H, Hedlund I and Treutiger CJ. (1992). "Molecular mechanisms and biological importance of *Plasmodium falciparum* erythrocyte rosetting*" Mem Inst Oswarldo*. 87 Suppl 3:323-9.

208. Wahlgren M, Carlson J, Udomsangpetch R and Perlman P, (1989). "Plasmodium falciparum-infected erythrocytes form spontaneous erythrocyte rosettes?" *Parasitology Today* Vol. 5, 183-185.

209. Wang C W, Hermsen C C, Sauerwein R W, Arnot D E, Theander T G and Lavstsen T, (2009). "The Plasmodium falciparum var gene transcription strategy at the onset of blood stage infection in a human volunteer." *Parasitol Int* Dec;58(4), 478-480. Epub 2009 Jul 2016.

210. Warhurst DC and Williams J E, (1996). "Laboratory diagnosis of malaria. ACP Broadsheet No 148." *J Clin Pathol* Vol. 49, 533–538.

211. Warrell D A (1993). Clinical features of malaria. In: Bruce-Chwatt's Essential Malariology. 3rd Edition, The Bath Press, UK.

# BIBLIOGRAPHY


212. Weatherall D J, Miller L H, Baruch D I, Marsh K, Doumbo O K and Casals-Pascual *et al*., (2002). "Malaria and the Red Cell." *Haematology*, 35-57.

213. Waitumbi J N, Opollo M O, Muga R O, Misore A O and Stoute J A, (2000). "Red cell surface changes and erythrophagocytosis in children with severe Plasmodium falciparum anaemia." *Blood* Vol. 95, 1481-1486.

214. Weisman H F, Bartow T, Leppo M K, Marsh H C Jr, Carson G R, Concino M F, Boyle M P, Roux K H, Weisfeldt M L, Fearon D T.(1990) "Soluble human complement receptor type 1: in vivo inhibitor of complement suppressing post-ischemic myocardial inflammation and necrosis." *Science*. 249:146-51.

215. Williams T N, Mwangi T W, Wambua S, Alexander N D, Kortok M, S. RW and Marsh K, (2005). "Sickle Cell Trait and the Risk of Plasmodium falciparum Malaria and other childhood Diseases." *Journal of Infectious Diseases* Vol. 192, 178-186.

216. Wilson J G, Murphy E E, Wong W W, Klickstein L B, Weis J H and Fearon D T (1986) " Identification of a restriction fragment length polymorphism by a CR1 CDNA that colorate waith the number of CR1 on erythrocytes." *J.Exp.Med* 164:50-59.

217. Wu J, Wu Y Q, Ricklin D, Janssen B J, Lambris J D and Gros P, (2009). "Structure of complement fragment C3b-factor H and implications for host protection by complement regulators." *Nat Immunol* 10, 728-733.

218. Wyler D J, Sypek J P and McDonald J A., (1985). "In vitro parasite-monocyte interactions in human Leishmaniasis: Possible role of fribronectin in parasite attachment." *Infect Immun* Vol. 49, 305-311

219. Xiang L, Rundles J R, Hamilton D R and Wilson J G. (1999). " Quantitative alleles of CR1;coding sequence analysis and comparison of heplotype in two ethinc grops." *J Immunol* 163:4939-4945.

220. Yoon S and Fearon D T. (1985). "Characterization of a soluble form of the C3b/C4b receptor (CR1) in human plasma". *J Immunol* 134:3332-3338.

221. Zimmerman P A J, Fitness J M, Moulds D T, McNamara L J, Kasehagen J, Rowe A and Hill A V S. (2003). "CR1 Knops blood group alleles are not associated with severe malaria in the Gambia." *Genes Immun.* 4:368–373

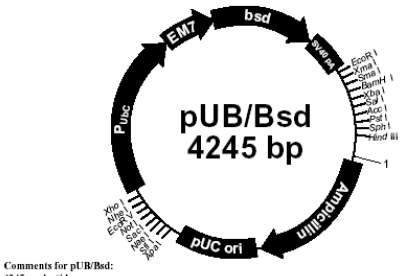
222. http://www.biacore.com/lifesciences/products/Consumables/guide/c1/index.html

223. http://microgen.ouhsc.edu/biacore.htm

# APPENDICES

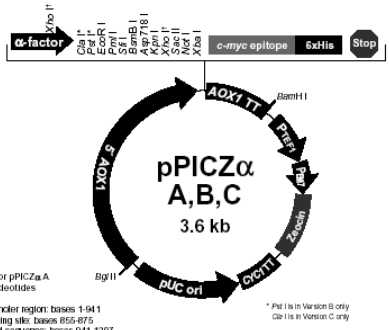
## Appendix A

### Material and method




**Comments for pUB/Bsd:**
**4245 nucleotides**

*bla* promoter: bases 102-200
Ampicillin (*bla*) resistance gene: bases 201-1061
pUC origin: bases 1206-1879
5´ polylinker: bases 1883-1933
Human ubiquitin C (*UbC*) promoter: bases 1941-3155
EM7 promoter: bases 3163-3229
Blasticidin (*bsd*) resistance gene: bases 3230-3628
SV40 early polyadenylation sequence: bases 3670-3800
3´ polylinker: bases 3800-3845




Comments for pPICZα A
3593 nucleotides

5´ *AOX1* promoter region: bases 1-941
5´ *AOX1* priming site: bases 855-875
α-factor signal sequence: bases 941-1207
Multiple cloning site: bases 1208-1276
*c-myc* epitope: bases 1275-1304
Polyhistidine (6xHis) tag: bases 1320-1337
3´ *AOX1* priming site: bases 1423-1443
*AOX1* transcription termination region: bases 1341-1682
*TEF1* promoter: bases 1683-2093
EM7 promoter: bases 2095-2162
*Sh ble* ORF: bases 2163-2537
*CYC1* transcription termination region: bases 2538-2855
pUC origin: bases 2866-3539 (complementary strand)

* *Pst* I is in Version B only
*Cla* I is in Version C only

†The two *Xho* I sites in the vector allow the user to clone their gene in frame with the Kex2 cleavage site, resulting in expression of their native gene without additional amino acids at the N-terminus.

Map of pUB/Bsd-TOPO® vector:

Map of pPICZαB vector

See www.invitrogen.com for full sequences

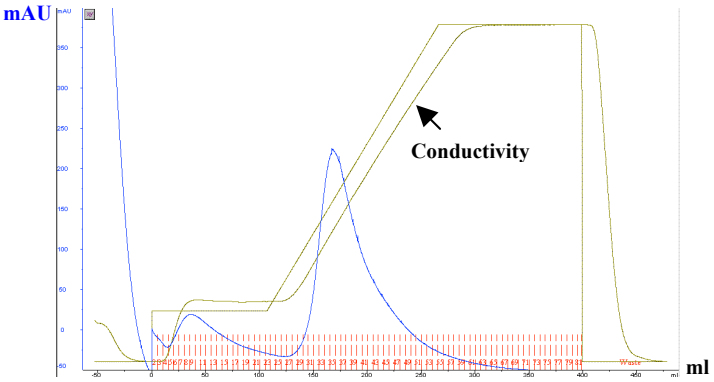
# APPENDICES

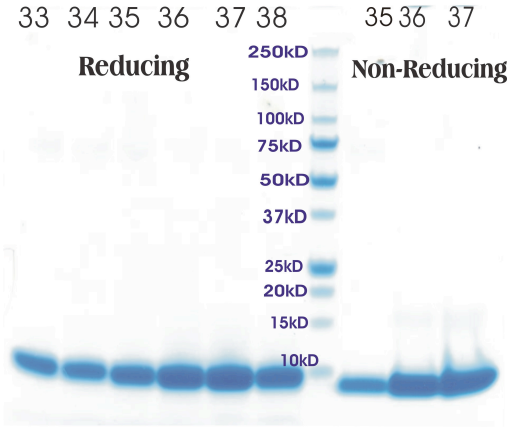
## Appendix B

| BUFFERS AND MEDIA | MATERIALS |
|---|---|
| **BMG (Buffered minimal glycerol)** | This medium contained 100 mM Potassium phosphate pH 6, 1.34% YNB (yeast nitrogen base with ammonium sulphate without amino acids), $4x10^{-5}$% biotin, 1% glycerol. |
| **BMM (Buffered minimal methanol)** | This media contained 100 mM Potassium phosphate pH 6, 1.34% YNB (yeast nitrogen base with ammonium sulphate without amino acids), $4x10^{-5}$ % biotin, 0.5% methanol. |
| **LB (Luria-Bertani)** | Tryptone (10 g), NaCl (10 g) and yeast extract (2 g) were dissolved in 1 L $dH_2O$, autoclaved and sealed. Prior to use LB Ampicillin (to a final concentration of 100 µg/ml) was added. Preparation of agar plates includes the addition of 15 g agar prior to autoclaving. |
| **LB (Luria-Bertani) Lennox** | Tryptone (10 g), NaCl (5 g) and yeast extract (2 g) were dissolved in 1 L $dH_2O$, autoclaved and sealed. Prior to use Zeocin™™ (to a final concentration of 25 µg/ml) was added. Preparation of agar plates includes the addition of 15 g agar prior to autoclaving. |
| **SDS Sample buffer** | The SDS sample buffer contained 50 mM Tris-HCl, 100 mM ßME, 2%S DS (w/v) , 0.1% bromophenol blue (w/v), and 10% glycerol (v/v) in distilled $H_2O$. |
| **SOC medium** | For this medium, 20 g Tryptone, 5 g Yeast Extract, 2 ml of 5 M NaCl, 2.5 ml of 1M KCl, 10 ml of 1M $MgCl_2$, 10 ml of 1M MgSO4 and 20ml of 1M glucose were mixed. The mixture was adjusted to to 1 L with distilled $H_2O$ (dH2O) and sterilize by autoclaving. |
| **TAE Agarose** | This consisted of agarose 1 % (w/v) in 1 x TAE buffer; 50X TAE buffer contained Tris-acetate (2 M) and 100 mM EDTA in distilled H2O. |
| **1X Tris-Glycine-SDS Buffer (TGS)** | The TGS buffer contained (in distilled $H_2O$) 0.025M Tris Base, 0.192 M Glycine, 0.1% SDS (w/v), pH 8.3. |
| **YPD (Yeast Extract Peptone Dextrose Medium)** | Tryptone (20 g) and yeast extract (10 g) were dissolved in 900 ml dH2O, autoclaved and sealed. The solution was cooled to approximately 50 °C before addition of sterile filtered glucose solution (20g in 100 ml dH2O). Prior to use Zeocin™™ (to a final concentration of 100 µg/ml) was added. Preparation of agar plates includes the addition of 15 g agar prior to autoclaving. |
| **YPDS (Yeast Extract Peptone Dextrose Medium with Sorbitol)** | Just as YPD (Yeast Extract Peptone Dextrose Medium) but 182 g sorbitol is added. Preparation of agar plates includes the addition of 15 g agar prior to autoclaving. |

# APPENDICES

## Appendix C




**Chromatography and SDS-PAGE gel of CR1 17**

# APPENDICES

## Appendix D

### D1

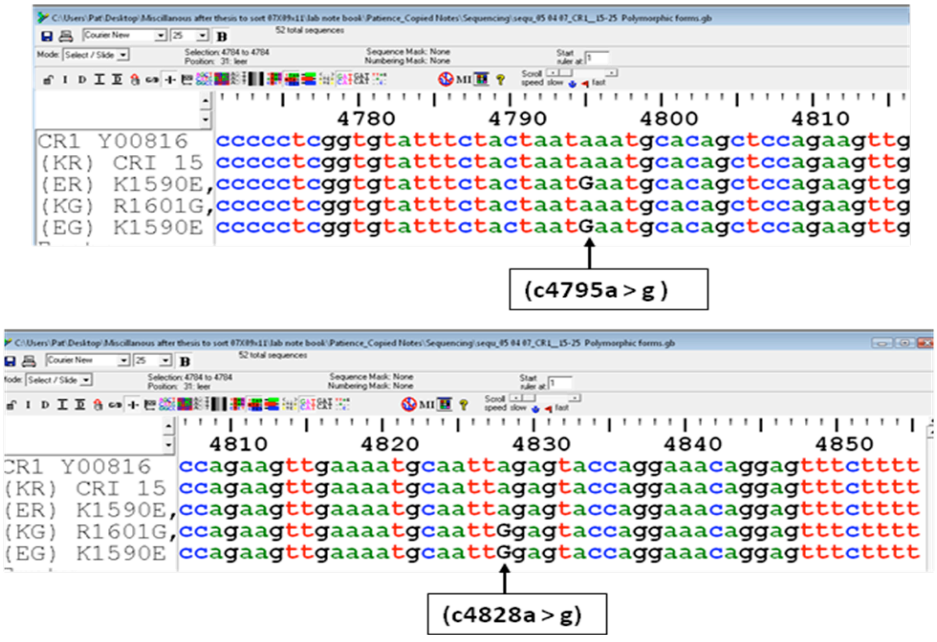
| # | Residue | Peak Number | # | Residue | Peak Number | # | Residue | Peak Number |
|---|---|---|---|---|---|---|---|---|
| 1314 | Glu - E | | 1363 | Phe - F | 30 | 1412 | Ile – I | 5 |
| 1315 | Ala - A | | 1364 | Asn - N | 115 | 1413 | Pro - P | - |
| 1316 | Glu - E | | 1365 | Leu - L | 78 | 1414 | Ile – I | 113 |
| 1317 | Ala - A | | 1366 | Ile - I | 13 | 1415 | Asn - N | 112 |
| 1318 | Ala - A | 18 | 1367 | Gly- G | 118 | 1416 | Asp - D | 27 |
| 1319 | Gly- G | 9 | 1368 | Glu - E | 22 | 1417 | Phe - F | 64 |
| 1320 | Glu - E | 94 | 1369 | Ser - S | 109 | 1418 | Glu - E | 26 |
| 1321 | His - H | 42 | 1370 | Thr - T | 44 | 1419 | Phe - F | 20 |
| 1322 | Ile – I | 20 | 1371 | Ile - I | 66 | 1420 | Pro - P | - |
| 1323 | Phe - F | | 1372 | Arg - R | 48 | 1421 | Val - V | 83 |
| 1324 | Cys - C | | 1373 | Cys - C | 99 | 1422 | Gly- G | 49 |
| 1325 | Pro - P | - | 1374 | Thr - T | 19 | 1423 | Thr - T | 14 |
| 1326 | Asn - N | 65 | 1375 | Ser - S | 54 | 1424 | Ser - S | 34 |
| 1327 | Pro - P | - | 1376 | Asp - D | 75 | 1425 | Leu - L | 69 |
| 1328 | Pro - P | - | 1377 | Pro - P | - | 1426 | Asn - N | 38 |
| 1329 | Ala - A | 73 | 1378 | His - H | 24 | 1427 | Tyr - Y | |
| 1330 | Ile - I | 45 | 1379 | Gly- G | 60 | 1428 | Glu - E | 26 |
| 1331 | Leu - L | 58 | 1380 | Asn - N | 36 | 1429 | Cys – C | 59 |
| 1332 | Asn - N | 31 | 1381 | Gly- G | 81 | 1430 | Arg – R | 79 |
| 1333 | Gly - G | 37 | 1382 | Val - V | 70 | 1431 | Pro – P | - |
| 1334 | Arg - R | 72 | 1383 | Trp - W | 86 | 1432 | Gly – G | 32 |
| 1335 | His - H | 53 | 1384 | Ser - S | 35 | 1433 | Tyr – Y | |
| 1336 | Thr - T | 116 | 1385 | Ser - S | 98 | 1434 | Phe – F | |
| 1337 | Gly- G | 5 | 1386 | Pro - P | - | 1435 | Gly- G | 59 |
| 1338 | Thr - T | 82 | 1387 | Ala - A | 11 | 1436 | Lys – K | 16 |
| 1339 | Pro - P | - | 1388 | Pro - P | - | 1437 | Met – M | 87 |
| 1340 | Ser - S | 25 | 1389 | Arg - R | 29 | 1438 | Phe – F | |
| 1341 | Gly- G | 117 | 1390 | Cys - C | 43 | 1439 | Ser – S | 3 |
| 1342 | Asp - D | 97 | 1391 | Glu - E | 10 | 1440 | Ile – I | 67 |
| 1343 | Ile - I | 93 | 1392 | Leu - L | 90 | 1441 | Ser – S | 39 |
| 1344 | Pro - P | - | 1393 | Ser - S | 63 | 1442 | Cys – C | 33 |
| 1345 | Tyr - Y | 28 | 1394 | Val - V | 84 | 1443 | Leu – L | 89 / 16 |
| 1346 | Gly- G | 105 | 1395 | Arg - R | 106 | 1444 | Glu – E | 116 |

# APPENDICES

| # | Residue | Peak Number | # | Residue | Peak Number | # | Residue | Peak Number |
|---|---|---|---|---|---|---|---|---|
| 1347 | Lys - K | 72 | 1396 | Ala - A | 62 | 1445 | Asn – N | 138 |
| 1348 | Glu - E | | 1397 | Gly- G | 2 | 1446 | Leu – L | 7 |
| 1349 | Ile - I | 1 | 1398 | His - H | 47 | 1447 | Val – V | 88 |
| 1350 | Ser - S | | 1399 | Cys - C | 85 | 1448 | Trp – W | 40 |
| 1351 | Tyr - Y | 22 | 1400 | Lys - K | 104 | 1449 | Ser – S | 114 |
| 1352 | Thr - T | | 1401 | Thr – T | 4 | 1450 | Ser – S | 101 |
| 1353 | Cys - C | 77 | 1402 | Pro - P | - | 1451 | Val – V | 6 |
| 1354 | Asp - D | 80 | 1403 | Glu - E | | 1452 | Glu – E | 103 |
| 1355 | Pro - P | - | 1404 | Gln - Q | | 1453 | Asp – D | |
| 1356 | His - H | Part of 42 | 1405 | Phe - F | | 1454 | Asn - N | |
| 1357 | Pro - P | - | 1406 | Pro - P | - | 1455 | Cys - C | |
| 1358 | Asp - D | 119 | 1407 | Phe - F | | 1456 | Arg - R | Part of 6 |
| 1359 | Arg - R | 111 | 1408 | Ala - A | 12 | | | |
| 1360 | Gly- G | 23 | 1409 | Ser - S | 57 | | | |
| 1361 | Met - M | 108 | 1410 | Pro - P | - | | | |
| 1362 | Thr - T | 76 | 1411 | Thr - T | 17 | | | |

Blue fonts are the eight amino acid residues in the linker region and Red labels represent the peaks that moved. Green fonts labelled are assignments not sure off or having two or three peaks on top of each other, hence rendered "Part of".

**D2**




**Single base pair changes that led to the knop polymorphism located on CR1 24-25**

# APPENDICES

## Appendix E

### Statistical Summery of Rosette Disruption Assay

| | 1way ANOVA | | | | | |
|---|---|---|---|---|---|---|
| 1 | Table Analyzed | Data 1 | | | | |
| 2 | | | | | | |
| 3 | One-way analysis of variance | | | | | |
| 4 | P value | < 0.0001 | | | | |
| 5 | P value summary | **** | | | | |
| 6 | Are means signif. different? (P < 0.05) | Yes | | | | |
| 7 | Number of groups | 8 | | | | |
| 8 | F | 15.45 | | | | |
| 9 | R square | 0.7300 | | | | |
| 10 | | | | | | |
| 11 | Bartlett's test for equal variances | | | | | |
| 12 | Bartlett's statistic (corrected) | 5.345 | | | | |
| 13 | P value | 0.6179 | | | | |
| 14 | P value summary | ns | | | | |
| 15 | Do the variances differ signif. (P < 0.05) | No | | | | |
| 16 | | | | | | |
| 17 | ANOVA Table | SS | df | MS | | |
| 18 | Treatment (between columns) | 4708 | 7 | 672.5 | | |
| 19 | Residual (within columns) | 1741 | 40 | 43.53 | | |
| 20 | Total | 6449 | 47 | | | |
| 21 | | | | | | |
| 22 | Tukey's Multiple Comparison Test | Mean Diff. | q | Significant? P < 0.05? | Summary | 95% CI of diff |
| 23 | PBS vs 21-22 | 13.00 | 4.826 | Yes | * | 0.8228 to 25.18 |
| 24 | PBS vs sCR1 | 27.33 | 10.15 | Yes | **** | 15.16 to 39.51 |
| 25 | PBS vs 15-17 | 27.50 | 10.21 | Yes | **** | 15.32 to 39.68 |
| 26 | PBS vs 15-25C | 28.83 | 10.70 | Yes | **** | 16.66 to 41.01 |
| 27 | PBS vs 15-25KERG | 27.17 | 10.09 | Yes | **** | 14.99 to 39.34 |
| 28 | PBS vs 15-25KE | 27.67 | 10.27 | Yes | **** | 15.49 to 39.84 |
| 29 | PBS vs 15-25RG | 29.67 | 11.01 | Yes | **** | 17.49 to 41.84 |
| 30 | 21-22 vs sCR1 | 14.33 | 5.321 | Yes | * | 2.156 to 26.51 |
| 31 | 21-22 vs 15-17 | 14.50 | 5.383 | Yes | * | 2.323 to 26.68 |
| 32 | 21-22 vs 15-25C | 15.83 | 5.878 | Yes | ** | 3.656 to 28.01 |
| 33 | 21-22 vs 15-25KERG | 14.17 | 5.260 | Yes | * | 1.989 to 26.34 |
| 34 | 21-22 vs 15-25KE | 14.67 | 5.445 | Yes | ** | 2.489 to 26.84 |
| 35 | 21-22 vs 15-25RG | 16.67 | 6.188 | Yes | ** | 4.489 to 28.84 |
| 36 | sCR1 vs 15-17 | 0.1667 | 0.06188 | No | ns | -12.01 to 12.34 |
| 37 | sCR1 vs 15-25C | 1.500 | 0.5569 | No | ns | -10.68 to 13.68 |
| 38 | sCR1 vs 15-25KERG | -0.1667 | 0.06188 | No | ns | -12.34 to 12.01 |
| 39 | sCR1 vs 15-25KE | 0.3333 | 0.1238 | No | ns | -11.84 to 12.51 |
| 40 | sCR1 vs 15-25RG | 2.333 | 0.8663 | No | ns | -9.844 to 14.51 |
| 41 | 15-17 vs 15-25C | 1.333 | 0.4950 | No | ns | -10.84 to 13.51 |
| 42 | 15-17 vs 15-25KERG | -0.3333 | 0.1238 | No | ns | -12.51 to 11.84 |
| 43 | 15-17 vs 15-25KE | 0.1667 | 0.06188 | No | ns | -12.01 to 12.34 |
| 44 | 15-17 vs 15-25RG | 2.167 | 0.8044 | No | ns | -10.01 to 14.34 |

# APPENDICES

| | 1way ANOVA<br>Tabular results | | | | | |
|---|---|---|---|---|---|---|
| 1 | Table Analyzed | Data 1 | | | | |
| 2 | | | | | | |
| 3 | One-way analysis of variance | | | | | |
| 4 | P value | < 0.0001 | | | | |
| 5 | P value summary | **** | | | | |
| 6 | Are means signif. different? (P < 0.05) | Yes | | | | |
| 7 | Number of groups | 7 | | | | |
| 8 | F | 16.25 | | | | |
| 9 | R square | 0.8228 | | | | |
| 10 | | | | | | |
| 11 | ANOVA Table | SS | df | MS | | |
| 12 | Treatment (between columns) | 2656 | 6 | 442.7 | | |
| 13 | Residual (within columns) | 572.0 | 21 | 27.24 | | |
| 14 | Total | 3228 | 27 | | | |
| 15 | | | | | | |
| 16 | Tukey's Multiple Comparison Test | Mean Diff. | q | Significant? P < 0.05? | Summary | 95% CI of diff |
| 17 | Binding medium vs KP buffer | 6.425 | 2.462 | No | ns | -5.580 to 18.43 |
| 18 | Binding medium vs CR1 mAb J3B11 | 33.00 | 12.65 | Yes | **** | 21.00 to 45.00 |
| 19 | Binding medium vs Factor H full | 15.75 | 6.035 | Yes | ** | 3.745 to 27.75 |
| 20 | Binding medium vs FH 8-15 | 7.500 | 2.874 | No | ns | -4.505 to 19.50 |
| 21 | Binding medium vs FH 10-15 | 9.600 | 3.679 | No | ns | -2.405 to 21.60 |
| 22 | Binding medium vs FH 12-13 | 9.000 | 3.449 | No | ns | -3.005 to 21.00 |
| 23 | KP buffer vs CR1 mAb J3B11 | 26.58 | 10.18 | Yes | **** | 14.57 to 38.58 |
| 24 | KP buffer vs Factor H full | 9.325 | 3.573 | No | ns | -2.680 to 21.33 |
| 25 | KP buffer vs FH 8-15 | 1.075 | 0.4119 | No | ns | -10.93 to 13.08 |
| 26 | KP buffer vs FH 10-15 | 3.175 | 1.217 | No | ns | -8.830 to 15.18 |
| 27 | KP buffer vs FH 12-13 | 2.575 | 0.9867 | No | ns | -9.430 to 14.58 |
| 28 | CR1 mAb J3B11 vs Factor H full | -17.25 | 6.610 | Yes | ** | -29.25 to -5.245 |
| 29 | CR1 mAb J3B11 vs FH 8-15 | -25.50 | 9.772 | Yes | **** | -37.50 to -13.50 |
| 30 | CR1 mAb J3B11 vs FH 10-15 | -23.40 | 8.967 | Yes | **** | -35.40 to -11.40 |
| 31 | CR1 mAb J3B11 vs FH 12-13 | -24.00 | 9.197 | Yes | **** | -36.00 to -12.00 |
| 32 | Factor H full vs FH 8-15 | -8.250 | 3.161 | No | ns | -20.25 to 3.755 |
| 33 | Factor H full vs FH 10-15 | -6.150 | 2.357 | No | ns | -18.15 to 5.855 |
| 34 | Factor H full vs FH 12-13 | -6.750 | 2.587 | No | ns | -18.75 to 5.255 |
| 35 | FH 8-15 vs FH 10-15 | 2.100 | 0.8047 | No | ns | -9.905 to 14.10 |
| 36 | FH 8-15 vs FH 12-13 | 1.500 | 0.5748 | No | ns | -10.50 to 13.50 |
| 37 | FH 10-15 vs FH 12-13 | -0.6000 | 0.2299 | No | ns | -12.60 to 11.40 |

# APPENDICES

| | 1way ANOVA<br>Tabular results | | | | | |
|---|---|---|---|---|---|---|
| 1 | Table Analyzed | Data 1 | | | | |
| 2 | | | | | | |
| 3 | One-way analysis of variance | | | | | |
| 4 | P value | < 0.0001 | | | | |
| 5 | P value summary | *** | | | | |
| 6 | Are means signif. different? (P < 0.05) | Yes | | | | |
| 7 | Number of groups | 7 | | | | |
| 8 | F | 16.25 | | | | |
| 9 | R square | 0.8228 | | | | |
| 10 | | | | | | |
| 11 | ANOVA Table | SS | df | MS | | |
| 12 | Treatment (between columns) | 2656 | 6 | 442.7 | | |
| 13 | Residual (within columns) | 572.0 | 21 | 27.24 | | |
| 14 | Total | 3228 | 27 | | | |
| 15 | | | | | | |
| 16 | Tukey's Multiple Comparison Test | Mean Diff. | q | Significant? P < 0.05? | Summary | 95% CI of diff |
| 17 | Binding medium vs KP buffer | 6.425 | 2.462 | No | ns | -5.580 to 18.43 |
| 18 | Binding medium vs CR1 mAb J3B11 | 33.00 | 12.65 | Yes | *** | 21.00 to 45.00 |
| 19 | Binding medium vs Factor H full | 15.75 | 6.035 | Yes | ** | 3.745 to 27.75 |
| 20 | Binding medium vs FH 8-15 | 7.500 | 2.874 | No | ns | -4.505 to 19.50 |
| 21 | Binding medium vs FH 10-15 | 9.600 | 3.679 | No | ns | -2.405 to 21.60 |
| 22 | Binding medium vs FH 12-13 | 9.000 | 3.449 | No | ns | -3.005 to 21.00 |
| 23 | KP buffer vs CR1 mAb J3B11 | 26.58 | 10.18 | Yes | *** | 14.57 to 38.58 |
| 24 | KP buffer vs Factor H full | 9.325 | 3.573 | No | ns | -2.680 to 21.33 |
| 25 | KP buffer vs FH 8-15 | 1.075 | 0.4119 | No | ns | -10.93 to 13.08 |
| 26 | KP buffer vs FH 10-15 | 3.175 | 1.217 | No | ns | -8.830 to 15.18 |
| 27 | KP buffer vs FH 12-13 | 2.575 | 0.9867 | No | ns | -9.430 to 14.58 |
| 28 | CR1 mAb J3B11 vs Factor H full | -17.25 | 6.610 | Yes | ** | -29.25 to -5.245 |
| 29 | CR1 mAb J3B11 vs FH 8-15 | -25.50 | 9.772 | Yes | *** | -37.50 to -13.50 |
| 30 | CR1 mAb J3B11 vs FH 10-15 | -23.40 | 8.967 | Yes | *** | -35.40 to -11.40 |
| 31 | CR1 mAb J3B11 vs FH 12-13 | -24.00 | 9.197 | Yes | *** | -36.00 to -12.00 |
| 32 | Factor H full vs FH 8-15 | -8.250 | 3.161 | No | ns | -20.25 to 3.755 |
| 33 | Factor H full vs FH 10-15 | -6.150 | 2.357 | No | ns | -18.15 to 5.855 |
| 34 | Factor H full vs FH 12-13 | -6.750 | 2.587 | No | ns | -18.75 to 5.255 |
| 35 | FH 8-15 vs FH 10-15 | 2.100 | 0.8047 | No | ns | -9.905 to 14.10 |
| 36 | FH 8-15 vs FH 12-13 | 1.500 | 0.5748 | No | ns | -10.50 to 13.50 |
| 37 | FH 10-15 vs FH 12-13 | -0.6000 | 0.2299 | No | ns | -12.60 to 11.40 |

# APPENDICES

| 1way ANOVA Column statistics | | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|---|
| | | Binding mediu | KP buffer | CR1 mAb J3 | Factor H fu | FH 8-15 | FH 10-15 | FH 12-13 |
| 1 | Number of values | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 2 | | | | | | | | |
| 3 | Minimum | 43.00 | 36.50 | 6.000 | 24.00 | 36.00 | 38.00 | 33.00 |
| 4 | 25% Percentile | 44.25 | 36.88 | 8.000 | 25.75 | 37.00 | 38.00 | 34.25 |
| 5 | Median | 49.00 | 38.90 | 17.00 | 32.50 | 40.00 | 38.00 | 39.50 |
| 6 | 75% Percentile | 50.75 | 48.95 | 20.00 | 38.50 | 44.50 | 39.20 | 43.25 |
| 7 | Maximum | 51.00 | 52.00 | 20.00 | 40.00 | 46.00 | 39.60 | 44.00 |
| 8 | | | | | | | | |
| 9 | Mean | 48.00 | 41.58 | 15.00 | 32.25 | 40.50 | 38.40 | 39.00 |
| 10 | Std. Deviation | 3.559 | 7.080 | 6.633 | 6.652 | 4.123 | 0.8000 | 4.690 |
| 11 | Std. Error | 1.780 | 3.540 | 3.317 | 3.326 | 2.062 | 0.4000 | 2.345 |
| 12 | | | | | | | | |
| 13 | Lower 95% CI | 42.34 | 30.31 | 4.445 | 21.67 | 33.94 | 37.13 | 31.54 |
| 14 | Upper 95% CI | 53.66 | 52.84 | 25.56 | 42.83 | 47.06 | 39.67 | 46.46 |

| Table format: Column | | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|
| | | Binding medium | KP buffer | CR1 mAb J3B11 | CR1 15-17 | CR1 10-11 | CR1 17 |
| | | Y | Y | Y | Y | Y | Y |
| 1 | Expt 1 | 40.00 | 43 | 2.00 | 7.0 | 8 | 14.0 |
| 2 | Expt 2 | 45.00 | 41 | 18.00 | 28.0 | 12 | 22.0 |
| 3 | Expt 3 | 49.00 | 41 | 18.00 | 20.0 | 27 | 10.0 |
| 4 | Expt 4 | 41.00 | 38 | 10.00 | 18.0 | 15 | 20.0 |

| Table format: Column | | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|
| | | Binding medium | KP buffer | CR1 mAb J3B11 | CR1 20-23 | CR1 21-22 | CR1 21 |
| | | Y | Y | Y | Y | Y | Y |
| 1 | Expt 1 | 44.00 | 34 | 3.00 | 38.0 | 20 | 28.0 |
| 2 | Expt 2 | 51.00 | 47 | 13.00 | 30.0 | 47 | 33.0 |
| 3 | Expt 3 | 40.00 | 41 | 12.00 | 32.0 | 41 | 31.0 |
| 4 | Expt 4 | 39.00 | 40 | 16.00 | 34.0 | 40 | 36.0 |

# APPENDICES

PatRosdata09_3.pzf: 1way ANOVA of Data 1 - Tabular results - Thursday, 24 June 2010

| | 1way ANOVA<br>Tabular results | | | | | |
|---|---|---|---|---|---|---|
| 1 | Table Analyzed | Data 1 | | | | |
| 2 | | | | | | |
| 3 | One-way analysis of variance | | | | | |
| 4 | P value | < 0.0001 | | | | |
| 5 | P value summary | **** | | | | |
| 6 | Are means signif. different? (P < 0.05) | Yes | | | | |
| 7 | Number of groups | 7 | | | | |
| 8 | F | 16.25 | | | | |
| 9 | R square | 0.8228 | | | | |
| 10 | | | | | | |
| 11 | ANOVA Table | SS | df | MS | | |
| 12 | Treatment (between columns) | 2656 | 6 | 442.7 | | |
| 13 | Residual (within columns) | 572.0 | 21 | 27.24 | | |
| 14 | Total | 3228 | 27 | | | |
| 15 | | | | | | |
| 16 | Tukey's Multiple Comparison Test | Mean Diff. | q | Significant? P < 0.05? | Summary | 95% CI of diff |
| 17 | Binding medium vs KP buffer | 6.425 | 2.462 | No | ns | -5.580 to 18.43 |
| 18 | Binding medium vs CR1 mAb J3B11 | 33.00 | 12.65 | Yes | **** | 21.00 to 45.00 |
| 19 | Binding medium vs Factor H full | 15.75 | 6.035 | Yes | ** | 3.745 to 27.75 |
| 20 | Binding medium vs FH 8-15 | 7.500 | 2.874 | No | ns | -4.505 to 19.50 |
| 21 | Binding medium vs FH 10-15 | 9.600 | 3.679 | No | ns | -2.405 to 21.60 |
| 22 | Binding medium vs FH 12-13 | 9.000 | 3.449 | No | ns | -3.005 to 21.00 |
| 23 | KP buffer vs CR1 mAb J3B11 | 26.58 | 10.18 | Yes | **** | 14.57 to 38.58 |
| 24 | KP buffer vs Factor H full | 9.325 | 3.573 | No | ns | -2.680 to 21.33 |
| 25 | KP buffer vs FH 8-15 | 1.075 | 0.4119 | No | ns | -10.93 to 13.08 |
| 26 | KP buffer vs FH 10-15 | 3.175 | 1.217 | No | ns | -8.830 to 15.18 |
| 27 | KP buffer vs FH 12-13 | 2.575 | 0.9867 | No | ns | -9.430 to 14.58 |
| 28 | CR1 mAb J3B11 vs Factor H full | -17.25 | 6.610 | Yes | ** | -29.25 to -5.245 |
| 29 | CR1 mAb J3B11 vs FH 8-15 | -25.50 | 9.772 | Yes | **** | -37.50 to -13.50 |
| 30 | CR1 mAb J3B11 vs FH 10-15 | -23.40 | 8.967 | Yes | **** | -35.40 to -11.40 |
| 31 | CR1 mAb J3B11 vs FH 12-13 | -24.00 | 9.197 | Yes | **** | -36.00 to -12.00 |
| 32 | Factor H full vs FH 8-15 | -8.250 | 3.161 | No | ns | -20.25 to 3.755 |
| 33 | Factor H full vs FH 10-15 | -6.150 | 2.357 | No | ns | -18.15 to 5.855 |
| 34 | Factor H full vs FH 12-13 | -6.750 | 2.587 | No | ns | -18.75 to 5.255 |
| 35 | FH 8-15 vs FH 10-15 | 2.100 | 0.8047 | No | ns | -9.905 to 14.10 |
| 36 | FH 8-15 vs FH 12-13 | 1.500 | 0.5748 | No | ns | -10.50 to 13.50 |
| 37 | FH 10-15 vs FH 12-13 | -0.6000 | 0.2299 | No | ns | -12.60 to 11.40 |

# APPENDICES

| 1way ANOVA Column statistics | | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|
| | | Binding mediu | KP buffer | CR1 mAb J3E | CR1 15-17 | CR1 10-11 | CR1 17 |
| 1 | Number of values | 4 | 4 | 4 | 4 | 4 | 4 |
| 2 | | | | | | | |
| 3 | Minimum | 40.00 | 38.00 | 2.000 | 7.000 | 8.000 | 10.00 |
| 4 | 25% Percentile | 40.25 | 38.75 | 4.000 | 9.750 | 9.000 | 11.00 |
| 5 | Median | 43.00 | 41.00 | 14.00 | 19.00 | 13.50 | 17.00 |
| 6 | 75% Percentile | 48.00 | 42.50 | 18.00 | 26.00 | 24.00 | 21.50 |
| 7 | Maximum | 49.00 | 43.00 | 18.00 | 28.00 | 27.00 | 22.00 |
| 8 | | | | | | | |
| 9 | Mean | 43.75 | 40.75 | 12.00 | 18.25 | 15.50 | 16.50 |
| 10 | Std. Deviation | 4.113 | 2.062 | 7.659 | 8.655 | 8.185 | 5.508 |
| 11 | Std. Error | 2.056 | 1.031 | 3.830 | 4.328 | 4.093 | 2.754 |
| 12 | | | | | | | |
| 13 | Lower 95% CI | 37.21 | 37.47 | -0.1879 | 4.477 | 2.475 | 7.736 |
| 14 | Upper 95% CI | 50.29 | 44.03 | 24.19 | 32.02 | 28.52 | 25.26 |

| 1way ANOVA Column statistics | | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|
| | | Binding mediu | KP buffer | CR1 mAb J3E | CR1 20-23 | CR1 21-22 | CR1 21 |
| 1 | Number of values | 4 | 4 | 4 | 4 | 4 | 4 |
| 2 | | | | | | | |
| 3 | Minimum | 39.00 | 34.00 | 3.000 | 30.00 | 20.00 | 28.00 |
| 4 | 25% Percentile | 39.25 | 35.50 | 5.250 | 30.50 | 25.00 | 28.75 |
| 5 | Median | 42.00 | 40.50 | 12.50 | 33.00 | 40.50 | 32.00 |
| 6 | 75% Percentile | 49.25 | 45.50 | 15.25 | 37.00 | 45.50 | 35.25 |
| 7 | Maximum | 51.00 | 47.00 | 16.00 | 38.00 | 47.00 | 36.00 |
| 8 | | | | | | | |
| 9 | Mean | 43.50 | 40.50 | 11.00 | 33.50 | 37.00 | 32.00 |
| 10 | Std. Deviation | 5.447 | 5.323 | 5.598 | 3.416 | 11.75 | 3.367 |
| 11 | Std. Error | 2.723 | 2.661 | 2.799 | 1.708 | 5.874 | 1.683 |
| 12 | | | | | | | |
| 13 | Lower 95% CI | 34.83 | 32.03 | 2.093 | 28.06 | 18.31 | 26.64 |
| 14 | Upper 95% CI | 52.17 | 48.97 | 19.91 | 38.94 | 55.69 | 37.36 |

# APPENDICES

| | 1way ANOVA<br>Tabular results | | | | | |
|---|---|---|---|---|---|---|
| 1 | Table Analyzed | Data 1 | | | | |
| 2 | | | | | | |
| 3 | One-way analysis of variance | | | | | |
| 4 | P value | < 0.0001 | | | | |
| 5 | P value summary | *** | | | | |
| 6 | Are means signif. different? (P < 0.05) | Yes | | | | |
| 7 | Number of groups | 6 | | | | |
| 8 | F | 18.56 | | | | |
| 9 | R square | 0.8376 | | | | |
| 10 | | | | | | |
| 11 | ANOVA Table | SS | df | MS | | |
| 12 | Treatment (between columns) | 3900 | 5 | 779.9 | | |
| 13 | Residual (within columns) | 756.2 | 18 | 42.01 | | |
| 14 | Total | 4656 | 23 | | | |
| 15 | | | | | | |
| 16 | Tukey's Multiple Comparison Test | Mean Diff. | q | Significant? P < 0.05? | Summary | 95% CI of diff |
| 17 | Binding medium vs KP buffer | 3.000 | 0.9257 | No | ns | -11.57 to 17.57 |
| 18 | Binding medium vs CR1 mAb J3B11 | 31.75 | 9.797 | Yes | *** | 17.18 to 46.32 |
| 19 | Binding medium vs CR1 15-17 | 25.50 | 7.868 | Yes | *** | 10.93 to 40.07 |
| 20 | Binding medium vs CR1 10-11 | 28.25 | 8.717 | Yes | *** | 13.68 to 42.82 |
| 21 | Binding medium vs CR1 17 | 27.25 | 8.408 | Yes | *** | 12.68 to 41.82 |
| 22 | KP buffer vs CR1 mAb J3B11 | 28.75 | 8.871 | Yes | *** | 14.18 to 43.32 |
| 23 | KP buffer vs CR1 15-17 | 22.50 | 6.943 | Yes | ** | 7.932 to 37.07 |
| 24 | KP buffer vs CR1 10-11 | 25.25 | 7.791 | Yes | *** | 10.68 to 39.82 |
| 25 | KP buffer vs CR1 17 | 24.25 | 7.482 | Yes | *** | 9.682 to 38.82 |
| 26 | CR1 mAb J3B11 vs CR1 15-17 | -6.250 | 1.928 | No | ns | -20.82 to 8.318 |
| 27 | CR1 mAb J3B11 vs CR1 10-11 | -3.500 | 1.080 | No | ns | -18.07 to 11.07 |
| 28 | CR1 mAb J3B11 vs CR1 17 | -4.500 | 1.389 | No | ns | -19.07 to 10.07 |
| 29 | CR1 15-17 vs CR1 10-11 | 2.750 | 0.8485 | No | ns | -11.82 to 17.32 |
| 30 | CR1 15-17 vs CR1 17 | 1.750 | 0.5400 | No | ns | -12.82 to 16.32 |
| 31 | CR1 10-11 vs CR1 17 | -1.000 | 0.3086 | No | ns | -15.57 to 13.57 |

| Table format:<br>Column | | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|---|
| | | Binding medium | KP buffer | CR1 mAb J3B11 | Factor H full | FH 8-15 | FH 10-15 | FH 12-13 |
| | | Y | Y | Y | Y | Y | Y | Y |
| 1 | Expt 1 | 48.00 | 38.0 | 6.00 | 24.0 | 36 | 38.0 | 33.0 |
| 2 | Expt 2 | 50.00 | 52.0 | 20.00 | 31.0 | 40 | 38.0 | 44.0 |
| 3 | Expt 3 | 51.00 | 39.8 | 20.00 | 40.0 | 46 | 39.6 | 38.0 |
| 4 | Expt 4 | 43.00 | 36.5 | 14.00 | 34.0 | 40 | 38.0 | 41.0 |

# APPENDICES

| | 1way ANOVA<br>Tabular results | | | | | |
|---|---|---|---|---|---|---|
| 1 | Table Analyzed | Data 1 | | | | |
| 2 | | | | | | |
| 3 | One-way analysis of variance | | | | | |
| 4 | P value | < 0.0001 | | | | |
| 5 | P value summary | *** | | | | |
| 6 | Are means signif. different? (P < 0.05) | Yes | | | | |
| 7 | Number of groups | 6 | | | | |
| 8 | F | 12.80 | | | | |
| 9 | R square | 0.7805 | | | | |
| 10 | | | | | | |
| 11 | ANOVA Table | SS | df | MS | | |
| 12 | Treatment (between columns) | 2671 | 5 | 534.2 | | |
| 13 | Residual (within columns) | 751.0 | 18 | 41.72 | | |
| 14 | Total | 3422 | 23 | | | |
| 15 | | | | | | |
| 16 | Tukey's Multiple Comparison Test | Mean Diff. | q | Significant? P < 0.05? | Summary | 95% CI of diff |
| 17 | Binding medium vs KP buffer | 3.000 | 0.9289 | No | ns | -11.52 to 17.52 |
| 18 | Binding medium vs CR1 mAb J3B11 | 32.50 | 10.06 | Yes | *** | 17.98 to 47.02 |
| 19 | Binding medium vs CR1 20-23 | 10.00 | 3.096 | No | ns | -4.517 to 24.52 |
| 20 | Binding medium vs CR1 21-22 | 6.500 | 2.013 | No | ns | -8.017 to 21.02 |
| 21 | Binding medium vs CR1 21 | 11.50 | 3.561 | No | ns | -3.017 to 26.02 |
| 22 | KP buffer vs CR1 mAb J3B11 | 29.50 | 9.134 | Yes | *** | 14.98 to 44.02 |
| 23 | KP buffer vs CR1 20-23 | 7.000 | 2.167 | No | ns | -7.517 to 21.52 |
| 24 | KP buffer vs CR1 21-22 | 3.500 | 1.084 | No | ns | -11.02 to 18.02 |
| 25 | KP buffer vs CR1 21 | 8.500 | 2.632 | No | ns | -6.017 to 23.02 |
| 26 | CR1 mAb J3B11 vs CR1 20-23 | -22.50 | 6.967 | Yes | ** | -37.02 to -7.983 |
| 27 | CR1 mAb J3B11 vs CR1 21-22 | -26.00 | 8.050 | Yes | *** | -40.52 to -11.48 |
| 28 | CR1 mAb J3B11 vs CR1 21 | -21.00 | 6.502 | Yes | ** | -35.52 to -6.483 |
| 29 | CR1 20-23 vs CR1 21-22 | -3.500 | 1.084 | No | ns | -18.02 to 11.02 |
| 30 | CR1 20-23 vs CR1 21 | 1.500 | 0.4644 | No | ns | -13.02 to 16.02 |
| 31 | CR1 21-22 vs CR1 21 | 5.000 | 1.548 | No | ns | -9.517 to 19.52 |

# APPENDICES

## Appendix F

# A

**SCSLDHKFHTNINTEYTE**GRKPCYERNEKRFSNEGEAKCGSDKIRDYGIKSAGGACAPF
RRQNLCDRNLEYLINKNTNTTHDLLGNVLVTAKYEGDSIVNNHPDKNSSGNKSSICTAL
ARSFADIGDIVRGRDMFKPNDADKVEKGLQVVFGKIYNSLPSPAQKHYAHDDGSGNYYK
LREDWWAINRKEVWKAITCRAPNEANFFRNISGNMKAFTSQGYCGHSETNVPTNLDYVP
QFLRWFDEWAEEFCRIRKIKLENVKKECRDEPNNKYCSGDGHDCKRTYLKDNTIFIDLN
CPRCENACSNYTKWIEIQRKQFDKQKRKYMNEIKIKTNISNNENDKEFYENLDKKGYST
INTFLESLNHGKQCQDNIDKKNKTNFKNNLETFGPSGY**CEACP**

# B

SCS LDHKFHTNIN TEYTEGRKPC YERNEKRFSN EGEAKCGSDK IRDYGIKSAG
GACAPFRRQN LCDRNLEYLI NKNT**Q**TTHDL LGNVLVTAKY EGDSIVNNHP DK**Q**SSG**Q**KSS
ICTALARSFA DIGDIVRGRD MFKPNDADKV EKGLQVVFGK IYNSLPSPAQ KHYAHDDGSG
NYYKLREDWW AINRKEVWKA ITCRAPNEAN FFR**Q**ISGNMK AFTSQGYCGH SETNVPTNLD
YVPQFLRWFD EWAEEFCRIR KIKLENVKKE CRDEPNNKYC SGDGHDCKRT YLKDNTIFID
LNCPRCENAC S**Q**YTKWIEIQ RKQFDKQKRK YMNEIKIKT**Q** ISNNENDKEF YENLDKKGYS
TINTFLESLN HGKQCQDNID KK**Q**KTNFKNN LETFGPSGYC EACPIYGVKC SNEKCTPVTE
NEWNSNNRLP TDTSTKNL**Q**A TNIDMLVNDG IGNAIDNELE K**Q**CTKYGILK GIKKQKWQCQ
YLNNIDQCKI NNVMNSGYFD NKIAFNVLFQ RWLRYFVRDH NRLKEKIDVC IKKENINENI
CIKRCKTNCE CVGKWLEKKE AEWDKINQHY NQKNHIMFIL IPYWITGFYE KITFPNDFFK
ALEDVDTINV LDTLKECQDT HCKIEKIRSI DVDLIKEIIS WLQNKIEVCK SHHDEDKHEY
CCDILPKSVD DDEEDDEEVD EEKEESSQTT KR**Q**ISQKGGT KSASCVKGAC AIVKGVLQQK
S**Q**GSIDNCNA KNRKKNEW**Q**C DKNTFVDGNE GVCMPPRRKS ICIH**Q**LTLEE QTKNKYQLRE
AFIKCAAKET NLLWDKYKND KNEAEELLKK GKIPEDFMRI MFYTFGDFRD FCLENDMGKD
VDKVKKNINK VF**QQ**SSKRGF KKIDPENWWN ENGPQIWNGM LCALIHADTK DSIKNKDNYK
YEKVTILAKR DGSNGMTLSE FAKKPKFLRW FVEWYDDYCK ERQKYLTEVA STCKSIDGGQ
LKCDRGCNNK CDEYKKYMRK KKEEWNLQDK YYKDKRENKG IDKGPIGIIV KDYVLANAKE
YLKKKFTASC VTSSGKAQNS ATEEVKKNIE LLSEEQYYDA DQYCGCT LVPRGS HHHHHH

# C

**DBLα-CIDR-DBLγ DNA sequence from Gene Art**

GGTACCTGCAGGATCTTGTTCTTTGGACCACAAGTTCCACACTAACATCAACACTGAGTA
CACTGAGGGTAGAAAGCCATGTTACGAGAGAAACGAGAAGAGATTCTCCAACGAGGGTGA
AGCTAAGTGTGGTTCCGACAAGATTAGAGACTACGGTATCAAGTCTGCCGGTGGTGCTTG
TGCTCCATTCAGAAGACAGAACCTTTGCGACAGAAACTTGGAGTACTTGATCAACAAGAA
CACTCAGACTACTCACGACTTGTTGGGTAACGTTTTGGTTACTGCTAAGTACGAGGGAGA
CTCCATTGTTAACAACCACCCAGACAAGCAATCTTCCGGTCAAAAGTCCTCCATCTGTAC
TGCTTTGGCTAGATCCTTCGCTGACATCGGTGACATTGTTAGAGGTAGAGACATGTTCAA
GCCAAACGACGCTGACAAGGTTGAAAAGGGATTGCAAGTTGTTTTCGGAAAGATCTACAA
CTCTTTGCCATCCCCAGCTCAAAAGCATTACGCTCACGATGATGGTTCTGGTAACTACTA
CAAGTTGAGAGAAGATTGGTGGGCTATCAACAGAAAAGAAGTTTGGAAGGCTATCACTTG
TAGAGCCCCAAACGAGGCTAATTTCTTCAGACAGATCTCCGGTAACATGAAGGCTTTCAC
TTCCCAAGGATACTGTGGTCACTCCGAGACTAACGTTCCAACTAACTTGGACTACGTTCC
ACAGTTCTTGAGATGGTTCGACGAATGGGCTGAAGAGTTCTGTAGAATCAGAAAGATCAA
GTTGGAGAACGTTAAGAAAGAATGTAGAGATGAGCCAAACAACAAGTACTGTTCCGGTGA

# APPENDICES

**A**

TGGTCACGACTGTAAGAGAACTTACTTGAAGGACAACACTATCTTCATCGACTTGAACTG
TCCAAGATGTGAGAACGCTTGTTCCCAGTACACTAAGTGGATCGAGATCCAGAGAAAGCA
ATTCGACAAGCAGAAAAGAAAGTACATGAACGAGATCAAGATCAAGACTCAGATCTCCAA
CAACGAGAACGACAAAGAGTTCTACGAGAACTTGGACAAGAAGGGATACTCCACTATCAA
CACTTTCTTGGAGTCCTTGAACCACGGAAAGCAATGTCAAGACAACATCGACAAGAAGCA
AAAGACTAACTTCAAGAACAACTTGGAGACTTTCGGACCATCTGGTTACTGTGAGGCTTG
TCCAATCTACGGTGTTAAGTGTTCCAACGAGAAGTGTACTCCAGTTACTGAGAACGAGTG
GAACTCCAACAACAGATTGCCAACTGACACTTCCACTAAGAACTTGCAGGCTACAAACAT
CGACATGTTGGTTAACGACGGTATCGGTAACGCTATTGACAACGAGTTGGAGAAGCAGTG
TACTAAGTACGGTATCTTGAAGGGTATCAAGAAACAAAAGTGGCAGTGTCAGTACTTGAA
CAACATCGACCAGTGTAAGATCAACAACGTTATGAACTCCGGTTACTTCGACAACAAGAT
CGCTTTCAACGTTTTGTTCCAGAGATGGTTGAGATACTTTGTTAGAGATCACAACAGATT
GAAAGAGAAGATCGATGTCTGCATCAAGAAAGAGAACATCAACGAGAACATCTGTATCAA
GAGATGTAAGACTAACTGTGAGTGTGTTGGAAAGTGGCTTGAGAAGAAAGAAGCTGAGTG
GGACAAGATTAACCAGCACTACAACCAGAAAAACCACATCATGTTCATCTTGATCCCATA
CTGGATCACTGGTTTCTACGAGAAGATCACTTTCCCAAACGATTTCTTCAAGGCTTTGGA
GGACGTTGACACTATCAACGTTTTGGACACTTTGAAAGAGTGTCAGGACACTCACTGTAA
GATCGAGAAGATCAGATCCATCGACGTTGACTTGATCAAAGAGATCATCTCCTGGTTGCA
GAACAAGATTGAAGTTTGTAAATCCCACCACGATGAGGATAAGCACGAGTACTGTTGTGA
CATCTTGCCAAAGTCTGTTGATGATGACGAAGAGGACGACGAAGAAGTTGACGAAGAGAA
AGAAGAGTCCTCCCAGACTACTAAGAGACAGATTTCCCAGAAGGGTGGTACTAAGTCTGC
TTCCTGTGTTAAGGGAGCTTGTGCTATCGTTAAGGGAGTTTTGCAACAGAAGTCCCAGGG
TTCCATTGATAACTGTAACGCTAAGAACAGAAAGAAAAACGAGTGGCAGTGTGACAAGAA
CACTTTCGTTGACGGTAACGAGGGAGTTTGTATGCCACCAAGAAGAAAGTCCATCTGTAT
CCACCAGTTGACTTTGGAGGAACAGACTAAGAACAAGTACCAGTTGAGAGAGGCTTTCAT
CAAGTGCGCTGCTAAAGAGACTAACTTGCTTTGGGACAAGTACAAGAACGATAAGAACGA
GGCTGAGGAATTGTTGAAGAAGGGAAAGATCCCAGAGGACTTCATGAGAATCATGTTCTA
CACTTTCGGTGACTTCAGAGACTTCTGTTTGGAGAACGACATGGGAAAGGATGTTGACAA

**B**

GGTTAAGAAGAACATCAACAAGGTTTTCCAGCAGTCCTCTAAGAGAGGTTTCAAGAAGAT
CGACCCAGAAAACTGGTGGAACGAGAACGGTCCACAAATTTGGAACGGAATGTTGTGTGC
TTTGATCCACGCTGACACTAAGGACTCCATCAAGAACAAGGACAACTACAAGTACGAGAA
GGTTACAATCTTGGCTAAGAGAGATGGTTCCAACGGAATGACTTTGTCCGAGTTCGCTAA
GAAGCCAAAGTTTTTGAGATGGTTTGTTGAGTGGTACGACGACTACTGTAAAGAGAGACA
GAAGTACTTGACTGAAGTTGCTTCCACTTGTAAGTCCATTGACGGTGGTCAATTGAAGTG
TGACAGAGGTTGTAACAACAAGTGTGACGAGTACAAGAAATACATGAGAAAGAAAAAAGA
AGAGTGGAACTTGCAGGACAAGTACTACAAGGACAAGAGAGAGAACAAGGGTATTGACAA
GGGTCCAATCGGTATCATCGTTAAGGACTACGTTTTGGCTAACGCTAAAGAGTACTTGAA
GAAGAAGTTCACTGCTTCTTGCGTTACTTCCTCCGGAAAGGCTCAAAACTCTGCTACTGA
AGAGGTTAAGAAAAATATCGAGTTGTTGTCCGAGGAACAATACTACGACGCTGACCAGTA
CTGTGGTTGTACTTTGGTTCCAAGAGGTTCTCATCATCACCATCACCACTAGTAGTCTAG
AGCTC
TCGAG

***Sequences of part of PfEMP1 used for the study (DBLs)***

(A)Template from DBLα(M)- Same sequence without the black highlighted area was the sequence used for preparing DBLα(P)(B) DBLα-CIDR-DBLγ sequence with N-glycosylation sites knocked out plus cleavable His-Tag: (C) DBLα-CIDR-DBLγ DNA sequence from Gene

# APPENDICES

***Additional Optimization sensogram on C1 chip***