Isolation, purification and characterisation of the pentraxins from *Limulus polyphemus* and *Mustelus canis*

By Sarah Bailey

Thesis submitted for the degree of Doctor of Philosophy

September 2008

Keele University

ORIGINAL COPY TIGHTLY

BOUND

THE FOLLOWING ITEMS HAVE BEEN EXCLUDED UNDER INSTRUCTION FROM THE UNIVERSITY

FIGURES

Fig 1.1, Fig 1.2 pg. 19, Fig 1.3 Fig 1.4 pg. 32,

Fig 1.5 pg. 41, Fig 1.7 pg. 54, Fig 1.8 pg. 55, Fig 6.1 pg. 274,

Fig 6.2 pg. 275.

Abstract

CRP and SAP form part of the innate immune system in humans, and homologues have been found in mammals, amphibians, fish, elasmobranchs, and invertebrates. The aim of this research was to broaden knowledge of CRP and SAP in horseshoe crabs, and to attempt to determine the structure and function of the CRP and SAP proteins from the elasmobranch smooth dogfish.

In the invertebrate horseshoe crabs, CRP comprises a heterogeneous family of proteins, whilst SAP is present as a single protein. The phylogenetic relationships between the horseshoe crab CRP and SAP sequences were determined, which showed clustering of homologous sequences. The conservation of horseshoe crab CRP and SAP sequences at sites corresponding to functional areas on human CRP was determined. The calcium ion binding site was the most conserved area, and variation at the area corresponding to the ligand binding site suggests a possibility of different binding properties between horseshoe crab pentraxins. Studies on the sequences of horseshoe crab CRP and SAP and the known human CRP and *Limulus* SAP structures, offers insight into new areas of conservation that may be important in horseshoe crab CRP and SAP function.

The Atlantic horseshoe crab, *Limulus polyphemus*, CRP and SAP were found to be composed of different subunit and molecular aggregate forms, and also bind to the novel ligands ribose-5-phosphate and AMP. *Limulus* CRP appeared to have a higher affinity than SAP for ribose-5-phosphate and AMP. The subunit composition and molecular aggregate surface hydrophobicity were analysed for whole *Limulus* CRP and SAP, and for that isolated using ribose-5-phosphate and AMP.

Attempts were made to isolate the genes in smooth dogfish (*Mustelus canis*) for CRP and SAP, but these proved unsuccessful primarily due to the availability of only limited amino acid sequence data resulting in primers for PCR that were too short and degenerate.

Contents

List of Figures.

Chapter 5. Isolation of the genes for pentraxins CRP and SAP from *Mustelus canis*

Chapter 6. Phylogenetic and structural analysis of *Limulus* CRP and SAP

Chapter 7. Conclusions and future work.

List of Tables

Amino acid table

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

Abbreviations.

AMP – Adenosine-5-monophosphate

CRP – C-reactive protein

cCRP – *Carcinoscorpius rotundicauda* C-reactive protein

EDTA – Ethylenediamine tetraacetic acid

FPLC – Fast protein liquid chromatography

HIC – Hydrophobicity interaction column

MPD – Methane pentane diol

PC – Phosphocholine

PCR - Polymerase Chain Reaction

PE – Phosphoethanolamine

PEG – Polyethylene glycol

SAP – Serum amyloid P-component

SDS – Sodium dodecyl sulphate

SDS-PAGE – Sodium dodecyl sulphate polyacrylamide gel electrophoresis

tCRP – *Tachypleus tridentatus* C-reactive protein

Acknowledgements

I would like to thank my supervisors Dr Annette Shrive and Dr John Mills, and my advisor Prof Trevor Greenhough for providing me with the opportunity to conduct this research degree. I am grateful for the help and support that they have all given me, and feel privileged to have been a part of the structural biology research group at Keele University. The assistance from Dr Sheila Hope is very much appreciated in my transition between structural and molecular biology. I would also like to thank the School of Life Sciences for providing me with a bursary to pay my fees. Also, the collaboration with Prof Peter Armstrong and his research group at the Marine Biological Laboratory, Woods Hole, Massachusetts, was a great learning experience and was highly appreciated.

The technical assistance from Ian Burns is much valued, I will miss the light-hearted jokes in the lab and learning much more about astronomy than I did before. Thanks also go out to Dr Paul Nield and Jenny Paterson for introducing me to chromatography and crystallography techniques.

I would like to thank my mum and dad for supporting me, and always being there for me during my research. I would especially like to thank my sister Jenna who has listened and helped me when things were not going to plan. She has always been honest and re-assuring when I needed it, and helped me to recover from my broken arm. Together we have learnt so much, especially not to ice skate during writing up your thesis.

The work described in this thesis was carried out building on the knowledge and research, and within the infrastructure, of the structural biology research group at Keele University. It was performed entirely by myself except where mentioned in "Chapter 5. Isolation of the genes for pentraxins CRP and SAP from *Mustelus canis*": isolation of dogfish CRP and SAP, cyanogen bromide digest of dogfish CRP, and trypsin digest of dogfish CRP were performed by Ian Burns and Alisha Roberts. Technical assistance with chromatography was provided by the research group technical support staff Ian Burns and Dr Paul Nield.

Chapter 1. Introduction to CRP and SAP proteins

1.1 Overview

The innate immune system is the first line of defence against infection, and plays an important role in the activation of the adaptive immune system in vertebrates. It is composed of many proteins and cells that work synergistically to rid the body of pathogens. During the acute phase response to tissue injury or infection, there is a rise in concentration of some plasma proteins synthesised in the liver which have been termed acute phase proteins (Kushner 1982) and form part of the innate immune system in humans. C-reactive protein (CRP) is one of these acute phase proteins which can increase in concentration up to 1000 fold and more above normal serum levels. CRP belongs to the family of proteins known as the pentraxins, so called because of their five-fold symmetry. Serum amyloid P component (SAP) also belongs to the pentraxin family but is not classed as an acute phase protein. Homologues of CRP and SAP are found in mammals, birds, fish, amphibians, and elasmobranchs, and in the invertebrate horseshoe crab family.

In mammals, there have only been identified single genes and proteins of CRP (Lei *et al.*, 1985; Woo *et al.*, 1985) and SAP (Mantzouranis *et al.*, 1985), whilst in the horseshoe crabs there have been identified multiple genes and mature proteins with different primary structures and binding properties (Nguyen *et al.*, 1986a; Shrive *et al.*, 1999; Iwaki *et al.*, 1999; Tharia *et al.*, 2002; Ng *et al.*, 2004). In the absence of adaptive immunity in the horseshoe crabs, the pentraxins may contribute a major part of the crab immune system and have a variety of different binding properties allowing possible recognition of different pathogens.

1.2 Human C-reactive protein (CRP)

1.2.1 The discovery of human CRP

Studies with patients acutely ill with Pneumonia, *Streptococcus*, *Staphylococcus*, and rheumatic fever, showed that their sera precipitated C-polysaccharide fraction of *Pneumococci*, but that this precipitation did not occur with sera from patients that were not acutely ill (Tillet and Francis 1930; Abernathy and Avery 1940; Macleod and Avery 1940a). Precipitation occurred only in the presence of calcium ions and was identified as being protein related (Abernathy and Avery 1940) hence the name "C-reactive protein" (Macleod and Avery 1940b).

1.2.2 Human CRP interaction with phosphocholine

C-polysaccharide (CPS) contains phosphocholine (PC) residues (Tomasz 1967), and phosphocholine inhibits the CRP-CPS binding reaction by binding CRP with high affinity (Volanakis and Kaplan 1971). Therefore, it was thought that CRP binds C-polysaccharide through PC residues. CRP binds C-polysaccharide (and hence PC) only in the presence of calcium ions (Abernethy and Avery 1940), but binds to calcium in the absence of phosphate. Therefore, calcium ion binding is not phosphate dependent, but phosphate binding to CRP is calcium dependent (Gotschlich and Edelman 1967). CRP binds to PC and also phosphoethanolamine (PE) through phosphate group interactions and position and charge of the choline/amine group (Oliveira *et al.*, 1980; Gotschlich *et al.*, 1982; Barnum *et al.*, 1982).

1.2.3 Human CRP interaction with polycations

CRP also binds to polycations such as poly-L-lysine, poly-L-arginine, poly-L-ornithine and protamine sulphate, but binds them calcium independently (DiCamelli *et al.*, 1980; Potempa *et al.*, 1981; Lee *et al.*, 2002). Precipitation of CRP with polycations was

dependent on the polycation size and ratio, whilst addition of calcium suppressed precipitation (Potempa *et al.*, 1981). Inhibition studies suggest that the PC and polycation sites on CRP do not overlap (Lee and Lee 2003).

1.2.4 Human CRP interaction with fibronectin and laminin

CRP binds to fibronectin that is immobilized (and *vice versa*) calcium dependently, which can be enhanced by a reduction in the pH, but inhibited by a calcium ion concentration greater than 1mM (Salonen *et al.*, 1984; Tseng and Mortensen 1988). The affect of pH on CRP-fibronectin binding suggests that CRP-fibronectin binding is inhibited under normal pH, but occurs at sites of inflammation where the pH is lower and may play a role in tissue repair. CRP-fibronectin binding occurs via the calcium binding site/PC binding site on CRP (Tseng and Mortensen 1988; Suresh *et al.*, 2004) at the mid-molecule region of fibronectin (Tseng and Mortensen 1986). Mutational studies on CRP residues 40 and 42 indicated that they were in the vicinity of the fibronectin binding area, as they altered CRP-fibronectin binding (Agrawal *et al.*, 1992). CRP was also found to bind to laminin which is a basement membrane protein, calcium dependently through the PC binding site on CRP, more specifically through amino acids 47-63 (Tseng and Mortensen 1989). This may provide an explanation as to why CRP deposits at sites of inflammation and suggests a role of CRP as a leukocyte ligand (Swanson *et al.*, 1989; Swanson and Mortensen 1990).

1.2.5 Human CRP lectin properties

Human CRP can act as a lectin, binding preferentially to carbohydrates with terminal galactosyl residues, and those which have the galactose conformation and phosphorylated at carbon position 6 (Kottgen *et al.*, 1992; Culley *et al.*, 2000; Lee and Lee 2003). Modified CRP (i.e. CRP monomers that show new (neo) epitopes) interaction with BSA

conjugates was inhibited far less by non-phosphorylated carbohydrates such as glucuronic acid, than phosphorylated ones (Lee and Lee 2003) indicating the importance of phosphorylation on CRP binding. This was also indicated by CRP-PC binding studies that showed binding was inhibited most greatly with PC then PE, 2-deoxyribose 5-phosphate, galactose-6-phosphate, ribose-5-phosphate, glucose-6-phosphate, AMP, and mannose-6-phosphate (Lee *et al.*, 2002).

Native CRP can express new or "neo" CRP antigens after treatment with acid, urea, or heating in the absence of calcium which corresponds to the production of free CRP subunits called modified CRP (Potempa *et al.*, 1987). Modified CRP also binds PC and galactose derivatives similar to native CRP (Volanakis and Nakartes 1981), but can bind the latter in the absence of calcium (Lee and Lee 2003). Modified CRP binding to phosphorylated carbohydrates was of greater affinity than native CRP binding, and occurred optimally in the absence of calcium (Lee and Lee 2003) suggesting a change in conformation of subunit. Phosphate containing carbohydrates mannose-1-phosphate, galactose-6-phosphate, glycerophosphate, mannose-6-phosphate, galactose-1-phosphate, glucose-6-phosphate, ribose-5-phosphate and AMP all inhibited modified CRP binding to galactose more efficiently than PC in both the presence and absence of calcium (Lee and Lee 2003). This suggested a change in structure of modified CRP from native CRP, and therefore perhaps a change in function. Antibodies against native CRP identified some identical areas on modified CRP (Ying *et al.*, 1989) which has been detected in the walls of blood vessels (Diehl *et al.*, 2000) indicating not only that it is a naturally occurring form of CRP but that it shows some similarity in structure to native CRP.

1.2.6 Human CRP interaction with membranes

CRP was shown to interact with liposome membranes resulting in activation of the classical complement pathway, and complement damage to liposomes (Richards *et al.*,

1977). The most susceptible of these membranes were positively charged and contained PC and a glycolipid (Richards *et al.*, 1977; 1979). Despite the presence of PC residues, it seemed that the interaction between CRP and the liposome membranes was more like the CRP interaction with polycations rather than with PC (Mold *et al.*, 1981b). CRP binding to liposomes was also dependent on the presence of stearylamine in the membrane, and binding increased with an increase in the membrane fluidity (Mold *et al.*, 1981b). CRP also binds to erythrocytes with PC residues on their surface in a calcium dependent reaction, which was inhibited by PC (Nakarates and Volanakis 1981; Volanakis and Nakartes 1981; and Volanakis and Kearney 1981). The binding of CRP to model membranes required incorporation of lysophosphatidylcholine into the bilayer (Volanakis and Wirtz 1979), and CRP did not show binding to normal unaltered vesicle lipid membranes (Narkates and Volanakis 1982). Therefore, alteration of the bilayer to expose PC residues must be required for CRP binding. This may be a mechanism by which CRP helps to clear up necrotic cells therefore reducing autoimmunity. It was also proposed that CRP binds to oxidised PC bearing phospholipids on low density lipoprotein and plasma membranes of apoptotic cells resulting in their removal and preventing autoimmunity (Chang *et al.*, 2002). The CRP binding affinity for PC increases when there are multiple PC residues on the membrane, suggesting the presence of multiple PC binding sites within the CRP pentamer (Lee and Lee 2003) which may anchor CRP to a PC coated surface more strongly. This was in agreement with earlier studies showing a single binding site for PC on each CRP protomer (Roux *et al.*, 1983) and therefore five PC binding sites on each pentamer. Discovery of the structures of CRP with and without PC bound showed that this was indeed the case (Shrive *et al.*, 1996; Thompson *et al.*, 1999). It was suggested that CRP exists as a pentamer whilst bound to a membrane in a calcium dependent manner, but in the absence of calcium CRP monomers bind to negatively charged membranes (Wang *et al.*, 2002).

1.2.7 Human CRP interaction with membranes and complement activation

Similar to CRP-liposome membrane binding, binding of CRP to erythrocytes with PC present on the surface caused cell lysis (Narkates and Volanakis 1982) by activation of the classical complement cascade through binding of C1q (Volanakis and Nakartes 1981). CRP binding to *Streptococcus pneumoniae* in the presence of calcium activated the classical complement cascade, and had an effect on opsonization of some strains (Mold *et al.*, 1982). CRP also binds apoptotic cells and enhances their opsonisation and phagocytosis, but also protected them from the terminal complement components therefore preventing lysis (Gershov *et al.*, 2000). CRP mediated activation of complement was influenced by the charge of liposome membranes, with positively charged membranes inducing more CRP mediated complement activation than negatively charged ones (Richards *et al.*, 1979). This may be due to a greater binding affinity of CRP to positively charged membranes than negatively charged ones, and a higher concentration of CRP bound to membranes may affect C1q activation.

1.2.8 Human CRP and infection

There has been observed an association between increases in the CRP concentration in serum and severe bacterial infections (Morley and Kushner 1982).The ability of CRP to bind PC and PE residues, polycations, and phosphorylated sugars may contribute to their protective capacity during infection. For example in mice, human CRP protected against infection with *Streptococcus pneumoniae* and *Salmonella enterica* (Mold *et al.* 1981a; Yother *et al.*, 1982; Szalai *et al.*, 1995; Szalai *et al.*, 2000a), but optimal activity required a functional complement system (Horowitz *et al.*, 1987). However, it has been shown that mice depleted of complement protein C4 were protected against *S. pneumoniae* infection by pre-treatment with human CRP (Nakayama *et al.*, 1983) which suggests that one is not

dependent on the other. In serum, human CRP protected against *S. pneumoniae* serotype 27 by activating the classical complement pathway, and enhancing the opsonisation of the bacteria (Edwards *et al.*, 1982). Mouse CRP agglutinated several species of bacteria and could be an opsonin showing similar functions with human CRP (Patterson and Higginbotham 1965). CRP also causes an increase in splenic removal of erythrocytes coated with pneumococcal C-polysaccharide, and a decrease in hepatic removal in a similar way to antibodies (Nakayama *et al.*, 1982).

Infection of human CRP transgenic mice with non-virulent *Salmonella* showed an increase in IgG concentration compared to non-transgenic mice, suggesting that CRP enhances the humoral response in mice, which may also occur in humans (Szalai *et al.*, 2000b). CRP also enhanced the phagocytic response of human neutrophils and monocytes to *Streptococcus pneumoniae* serotype 27 (Holzer *et al.*, 1984) through opsonisation of cell wall and capsule of *S. pneumoniae*, and subsequently activated complement.

CRP from humans and rodents has also been shown to bind to parasites such as *Leishmania* (Pritchard *et al.*, 1985; Culley *et al.*, 1996; Culley *et al.*, 2000), *Plasmodium* (Pied *et al.*, 1989), and fungi such as *Aspergillus fumigatus* (Baldo *et al.*, 1977; Richardson *et al.*, 1991) and *Candida albicans* (Richardson *et al.*, 1991). When binding to *Leishmania donovani* promastigotes, CRP interacted with the lipophosphoglycan component of the promastigote cell surface (Culley *et al.*, 1996).

1.2.9 Human CRP interaction with complement

Isolation of C1 of complement from sera identified a protein which binds C1q in a calcium dependent manner and also binds IgG, and therefore was proposed to be the link between IgG and C1q binding (Assimeh and Painter 1975a, b). This protein was CRP, although was identified as "C1t" at the time. Incubation of sera with acute phase proteins

such as CRP, showed depletion of C1, C4 and C2 but minimal depletion of C3-9, and required calcium ions and substrate (Siegel *et al.*, 1974). CRP enhanced the activity of C1 (Assimeh and Painter 1975a) and activated the classical complement pathway by its interaction with C1q. Modified CRP (i.e. CRP protomers) were shown not to activate the complement system (Potempa *et al.*, 1988) which suggests that aggregation of CRP induces conformational change that allows C1q binding. Studies showed that CRP trimers and dimers activated complement through the collagen-like region of C1q, but that CRP monomers did not (Jiang *et al.*, 1991b). It was also shown that PC did not inhibit CRP trimer-C1q binding which was calcium independent, suggesting the presence of two different binding sites: one which binds ligands such as PC through calcium interaction, and another which binds C1q without calcium interaction.

Later studies using modelling of the C1q structure and CRP structure, indicated that it was the globular head structure of C1q not the collagen-like region, that interacted with CRP as it could be accommodated by the negatively charged central pore of CRP (Gaboriaud *et al.*, 2003). On this model residues 112 and 175 from CRP subunits A and E which had previously been shown to be important in C1q binding (Agrawal *et al.*, 2001), come into contact with residues 175 and 200 from the C1q globular head subunits B and A. A more recent study has shown that antibodies against the globular head region of C1q prevented CRP binding to C1q, but that antibodies against the collagen-like region did not (McGrath *et al.*, 2006). This would leave the collagen-like region of C1q available for interaction with C1r, C1s and cell receptors (McGrath *et al.*, 2006). Therefore, it appears that CRP interacts with the globular head region of C1q not the collagen-like region as previously thought. It was also shown that interaction between C1q and CRP or IgG in solution inhibits the binding of C1q to immobilized IgG or CRP, which indicates not only that CRP can bind C1q in solution, but that the binding sites on C1q for CRP and IgG are overlapping or close to each other (McGrath *et al.*, 2006).

CRP was shown to induce the consumption of human complement and C3 conversion through C1q on reaction with pneumococcal C-polysaccharide (Kaplan and Volanakis 1974). This reaction was inhibited by PC and also occurred with choline phosphatides, lecithin and sphingomyelin and therefore complement consumption was thought to be attributable to activation by CRP complexes. However, later studies indicated that CRP does not require PC interaction to activate complement (McGrath *et al.*, 2006). CRP had also been found to activate complement in the absence of PC in mildly acidic conditions which gave a conformational change to the CRP molecule, which may indicate CRP function at sites of inflammation where the pH is more acidic (Miyazawa and Inoue 1990).

It appeared that CRP activation of complement leads only to C3 convertase (Kaplan and Volanakis 1974; Volanakis 1982) and it is possible that CRP and ligand that have C3 bound to them are targeted for opsonophagocytosis, so that both CRP and C1q work together as opsonic and anti-inflammatory proteins (Kishore *et al.*, 2004). Studies using infection of human CRP-transgenic mice with *S. pneumoniae* suggested that complement and CRP work synergistically with each other and amplify protection against infection (Szalai *et al.*, 1996).

1.2.10 Human CRP interaction with receptors of phagocytes

Human CRP binds to, and affects the function of phagocytic cells such as monocytes, macrophages and neutrophils. CRP binding to these cells is fast and saturable (Ballou *et al.*, 1989; Tebo and Mortensen 1990; Dobrinich and Spagnuolo 1991) and occurs through a site associated with the IgG receptor FcγR (Muller and Fehr 1986; Buchta *et al.*, 1987; Zeller *et al.*, 1989; Marnell *et al.*, 1995). Rabbit CRP has been shown to compete with human CRP to bind neutrophils, suggesting a conservation of function through mammals (Heurtz *et al.*, 2005). There is a greater binding of IgG complexes to neutrophils than CRP complexes (Romero *et al.*, 1998) indicating lower affinity of IgG receptors to CRP than

IgG. Only native and non-aggregated CRP binds via the IgG receptor (Zeller and Sullivan 1993), although earlier studies indicated otherwise (Zeller et al., 1986a). Native CRP was not shown to bind IgG (Ballou and MacIntyre 1990; Dobrinich and Spagnuolo 1991) but modified CRP binds to both aggregated and monomeric IgG, optimally at pH 5.5 which indicates a function of modified CRP with immune complexes at sites of inflammation (Motie et al., 1996).

It was suggested that CRP binds mouse macrophages through a receptor distinct from IgG FcγR, but that the receptor may have been influenced by FcγR (Zahedi et al., 1989). This was similar to what early studies showed for human CRP and monocyte binding, where binding was not through FcγRI or FcγRII, but instead through a distinct receptor (Tebo and Mortensen 1990). This also appeared to be the case for monocyte ingestion of erythrocytes coated with pneumococcal C-polysaccharide-CRP, as cell attachment was not through FcγR (Kilpatrick and Volanakis 1985). Also, CRP appeared to induce cell activation through FcγRI but without binding it (Zeller and Sullivan 1993).

Other studies, however, indicated that FcγR was the receptor for CRP. It was initially shown that human CRP binds to FcγRI on monocytes and neutrophils (Crowell et al., 1991), but other studies with transfected COS cells showed that FcγRIIa was the main receptor (Bharadwaj et al., 1999). The binding of CRP to monocytes via FcγRI is calcium dependent (Ballou et al., 1989) and is not inhibited by PC suggesting different binding sites on CRP for FcγRI and PC (Bodman-Smith et al., 2002). The affinity of CRP binding to FcγRII was allele specific with stronger binding to R131 than H131 (Stein et al., 2000). In experiments with mouse leukocytes, the receptors for CRP binding were FcγRI and FcγRIIb (Stein et al., 2000; Mold et al., 2001). The production of TNF-α and IL-1β by monocytes in response to S. pneumoniae was enhanced by human CRP (Mold and Du Clos 2006). This enhancement was blocked by PC, FcγRI inhibitors, and antibodies towards FcγRI and FcγRII, suggesting that CRP interaction with monocytes is through both FcγRI

and FcγRII. Later studies indicated that blocking FcγRI binding using antibodies abolished

CRP binding to monocytes, but blocking of FcγRII binding using the same method did not

affect CRP binding (Tron *et al.*, 2008). It was also shown that CRP binds to COS-7 cells

through FcγRI, and binding was enhanced by the γ-chain, which suggested its important

role in CRP binding (Röcker *et al.*, 2007). It could be that the different results for CRP-

FcγR interaction are due to differential expression and/or aggregation patterns of FcγR on

monocyte, macrophage and neutrophil membranes (Volanakis 2001).

1.2.11 Human CRP and induction of phagocytosis

CRP bound to PC-BSA does not activate neutrophils suggesting that when complexed,

CRP does not activate immune cells (Romero *et al.*, 1998). However, CRP stimulated

phagocytosis of bacteria such as *S. aureaus, E. coli,* and *Klebsiella aerogenes* (Kindmark

1971). CRP appears to enhance phagocytosis through FcγRI binding (Hokama *et al.*, 1962;

Mortensen and Duszkiewicz 1977; Mold *et al.*, 2001; Mold *et al.*, 2002; Thomas-Rudolph

et al., 2007), FcγRIIa binding (Bodman-Smith *et al.*, 2004) or FcγRIII binding (Mold *et

al.*, 2002) probably by acting as a cross linker of neutrophil to bacteria (Bodman-Smith *et

al.*, 2002). Experiments with CRP coated erythrocytes showed that ingestion by monocytes

was proportional to the amount of CRP on the erythrocyte, and was reduced by blocking

the monocyte receptors with aggregated IgG (Mortensen *et al.*, 1976). Ingestion of

erythrocytes coated with C-polysaccharide is dependent on both CRP on erythrocyte

surface and treatment with complement (Mortensen and Duszkiewicz 1977). CRP

complexed with C-polysaccharide was cleared by neutrophils, a function which was

enhanced by activation of complement (Shephard *et al.*, 1986).

1.2.12 Human CRP and phagocytic cell functions

CRP was shown to activate macrophages without cytokine being present (Zahedi and Mortensen 1986). Native CRP, CRP peptides from neutrophil degradation and modified CRP, have all been shown to affect phagocytic cell functions. Phagocyte superoxide production, secretion of vitamin B12-binding protein, chemotaxis, degradation, phagocytosis, IL-1 production, TNF production, extracellular signalling, metalloproteinase production, increase in cAMP levels, and intracellular calcium movement were all affected (Hokama *et al.*, 1962; Zeller *et al.*, 1986b; Buchta *et al.*, 1987; Robey *et al.*, 1987; Shepherd *et al.*, 1988; Potempa *et al.*, 1988; Miyagawa *et al.*, 1988; Kew *et al.*, 1990; Dobrinich and Spagnuolo 1991; Mortensen and Zhong 2000; Zouki *et al.*, 2001; Woollard *et al.*, 2002; Abe *et al.*, 2005; Mold and Du Clos 2006). CRP peptides 51-58 and 181-187 enhanced macrophage phagocytosis, and CRP peptides 37-58, 51-58, 173-187, and 181-187 inhibited some superoxide production, lysozyme release and vitamin B12 binding protein release from neutrophils (Buchta *et al.*, 1986).

CRP has been shown to affect monocyte interaction with endothelial cells (Woollard *et al.*, 2002). Native CRP and CRP peptides 174-185 and 201-206 inhibited adhesion of neutrophils to endothelial cells by binding FcγRIIIb and downregulating expression of L-selectin on the neutrophil surface (Zouki *et al.*, 1997). Therefore, it is possible that CRP accumulates at a site of inflammation, and acts as an anti-inflammatory regulator of phagocytic cell function by providing a negative feedback system in which CRP peptides produced by phagocyte digestion affect cell function. Inhibition of macrophage chemotaxis by CRP could also be a way in which CRP plays a role in inflammation by causing accumulation of macrophages (Miyagawa *et al.*, 1988).

Studies indicated that CRP is expressed by macrophages and it appears on the membrane where it may act as a selectin-like adhesion molecule (Kolb-Bachofen *et al.*,

1995). Rat macrophages have also been shown to synthesize their own membrane-bound C-reactive protein protomer during the acute phase response, and have a binding protein in their plasma membrane that is specific for CRP (Egenhofer *et al.*, 1993).

1.2.13 Human CRP and prevention of autoimmunity

CRP was found to bind to chromatin, but only in the presence of histone H1 (Du Clos *et al.*, 1988; Du Clos *et al.*, 1991). CRP-chromatin binding was calcium dependent and inhibited by PC suggesting a site of chromatin binding associated with the PC binding site (Robey *et al.*, 1984). Interactions were also found to occur between histones H2A, H2B and the H2A-H2B complex with CRP probably through the PC binding site (Du Clos *et al.*, 1988; Du Clos *et al.*, 1991 (a, b)). The CRP-histone interaction was eliminated when the histones were re-constituted with DNA (Du Clos *et al.*, 1991 (a)). Therefore, the interaction of CRP with histones did not appear to be responsible for CRP-chromatin binding (Du Clos *et al.*, 1991 (a)). CRP binding with chromatin and histones also seemed to be different to that of SAP (Hicks *et al.*, 1992).

CRP was also shown to bind the U1 small ribonuclear protein particle possibly through the PC binding site and in a calcium dependent manner (Du Clos 1989; Du Clos *et al.*, 1991 (a)). Upon injection into VERO cells, CRP was shown to locate to the nucleus, as was SAP (Du Clos *et al.*, 1990) which may be due to binding properties for histones, chromatin and nucleosome core particles.

CRP was shown to prolong survival of mice injected with chromatin-coated beads and reduced the antibody response to histones and DNA, although did not affect nucleosome core particle clearance (Du Clos *et al.*, 1994). The interactions of CRP with nuclear antigens did not appear to occur at physiological ionic strength, and did not activate complement or cleave DNA (Butler *et al.*, 1990). To determine whether CRP binding to

nuclear antigens was likely to occur *in vivo,* experiments were performed using human cells under near physiological conditions (Pepys *et al.,* 1994). It was found that CRP bound to small ribonucleoproteins, and SAP bound to chromatin and the nucleolus. It was suggested that the controversy about CRP binding was probably due to experimental procedures that involved denaturation and non-physiological conditions (Pepys *et al.,* 1994).

1.2.14 Human CRP implications in heart disease

CRP is a useful non-specific marker for inflammation, and has been used to show relationships between an increase in CRP production and future atherothrombotic events (Pepys and Hirschfield 2003). Studies on stable and unstable angina showed that an increase in baseline CRP values is associated with a significantly increased risk of future coronary events (Pepys 2008).

Raised CRP concentrations in circulation can indicate a higher risk of coronary events in patients with angina (Haverkate *et al.,* 1997) and may be an important factor in inflammation during angina (Liuzzo *et al.,* 1994). Elevated CRP levels were also shown to predict the future risk of ischaemic stroke, peripheral arterial disease and coronary heart disease (Danesh *et al.,* 1998). CRP induced the generation of reactive oxygen species by vascular smooth muscle cells, which controls pro-inflammatory activities of the cells and may promote atherogenesis (Ryu *et al.,* 2007). Atherosclerosis and plaque instability in atherothrombotic events are inflammatory and produce TNF-α, IL-1 and IL-6, and association with an increase in CRP levels could arise this way as they regulate CRP gene expression. CRP is co-deposited with activated complement in myocardial infarcts as it is a site of inflammation and an acute phase stimulus (Pepys and Hirschfield 2003). The CRP levels rise according to muscle tissue damage after myocardial infarction, and elevated CRP levels 14 days after the event suggest persistent inflammation and predict recurrent

events (Morrow *et al.*, 1998). Studies showed that injection of human CRP into rats enhanced myocardial infarct size by ~40%, and this was aided by complement as CRP activates complement it can lead to inflammation and more damage but did not show any effect in healthy rats (Griselli *et al.*, 1999).

Therefore, as CRP levels increase due to inflammation and tissue damage, CRP may not be directly involved in atherothrombosis. Recent studies using mice with and without the human CRP transgene have shown that human CRP is not pro-atherogenic, pro-inflammatory or pro-atherothrombotic (Tennent *et al.*, 2008).

However, other studies suggest human CRP directly affects athero-thrombotic events. Human CRP was shown to affect the growth of coronary artery endothelial cells and muscle cells (Yang *et al.*, 2005; Cirillo *et al.*, 2005) suggesting that it plays a part in thrombosis and/or myocardial infarction. Human CRP also promoted the adhesion of platelets to endothelial cells via up-regulated P-selectin expression, linking it with inflammation and thrombosis (Yaron *et al.*, 2006). Inflammatory levels of CRP induce monocytes to express tissue factor, which is a potent coagulant suggesting that CRP may also affect coagulation as well as thrombosis (Cermak *et al.*, 1993). However, platelet aggregation was inhibited by CRP (Fiedel *et al.*, 1982), and the onset of coagulation delayed (Fiedel and Ku 1986) as was activation of platelet factor 3 and release of β-glucuronidase (Fiedel and Gewurz 1976 (a, b)). CRP and C1q inhibited platelet adhesion to IgG and human serum albumin (Skoglund *et al.*, 2008). In contrast, modified CRP induced aggregation of platelets and potentiated their activation which may mean that conversion of CRP to modified CRP at sites of inflammation may be important to platelet function (Potempa *et al.*, 1988). CRP complexes with poly-L-lysine and protamine, and heat aggregated CRP also induced platelet aggregation, but CRP-C-polysaccharide complex did not (Fiedel *et al.*, 1982; Potempa *et al.*, 1988). Therefore, native CRP inhibits platelet

aggregation and clot formation, but modified CRP and CRP-complexes (except with C-polysaccharide) activated platelet aggregation.

Platelet cytotoxicity against *Schistosoma mansoni* in rats was correlated with the CRP concentration (Bout *et al.*, 1986) indicating a function of CRP on platelet activity in rats as well as humans. More recent studies have shown that injection of natural human CRP into humans does not induce the acute phase response, or activate coagulation and is therefore not pro-inflammatory (Pepys *et al.*, 2005).

1.2.15. Human CRP and drug design

A drug, 1,6-bis(phosphocholine)-hexane, was developed that inhibits CRP binding calcium dependently to ligands including PC, PE, low density lipoprotein, and apoptotic or necrotic cells, and also blocked complement activation (Pepys *et al.*, 2006). This compound was based on a drug designed for inhibition of human SAP binding (Pepys *et al.*, 2002). 1,6-bis(phosphocholine)-hexane binds CRP in a calcium dependent manner and cross-links two CRP pentamers face to face via their PC binding sites. The drug was tolerated in rats and mice without adverse effects, and prevented death in rats with coronary artery ligation compared to rats without the drug.

1.2.16 The gene and peptide sequence of human CRP

An early study used Edman degradation and peptide sequencing to determine the amino acid sequence of human CRP (Oliviera *et al.*, 1979). This study showed that each human CRP protomer was 187 amino acids long with a single disulphide bond. Later studies that isolated the human CRP gene using a genomic library, showed that the derived amino acid sequence was actually 206 amino acids long as there were an additional 19 amino acids between positions 61 and 62 of the original determined amino acid sequence (Lei *et al.*,

1985; Woo *et al.*, 1985). The gene from human CRP is found on chromosome 1 (Whitehead *et al.*, 1983) and the coding area of the gene contained a single intron of 278 base pairs that was positioned after the codon for amino acid position three (Asp) (Lei *et al.*, 1985; Woo *et al.*, 1985). The gene also showed a sequence of 54 nucleotides encoding an 18 amino acid signal peptide as previously found by Tucci *et al.*, (1983) which was Met-Glu-Lys-Leu-Leu-Cys-Phe-Leu-Val-Leu-Thr-Ser-Leu-Ser-His-Ala-Phe-Gly.   The signal peptide is cleaved after translocation of the protein into membrane vesicles (Tucci *et al.*, 1983).

In the CRP gene the polyadenylation signal was 1.2kbp from the termination codon, and the first potential promoter box of the gene was 133 nucleotides upstream of the ATG start codon (Lei *et al.*, 1985). In the 5' portion of the gene there are three regions similar to the *Drosophila* heat shock sequences which may lead to enhancement of CRP synthesis during the acute phase response (Woo *et al.*, 1985). There are also splice junctions in the 5' region and the 3' region which may mean poly mRNA forms for human CRP, or different mRNA processing occurring in tissues other than the liver and/or in non-acute phase liver (Woo *et al.*, 1985). Human CRP mRNA synthesis produced mRNA 2.2kbp long (Tucci *et al.*, 1983; Lei *et al.*, 1985) which gave a mature protein of ~24kDa subunit size after processing (Tucci *et al.*, 1983). Although only one gene has been found for human CRP, there have been found isoforms of CRP monomers using SDS-PAGE and isoelectric focusing which vary in size and expression in serum from patients with different diseases (Das *et al.*, 2003).

1.2.17 The site of formation of human CRP

In adults the average CRP baseline concentration is 0.8mg/l but after an acute phase stimulus it can increase from <50µg/l to >500mg/l (i.e. a potential increase of 10,000 fold) (Pepys and Hirschfield 2003). CRP is produced by the liver 16-24 hours after infection

(Hurlimann *et al.*, 1965). In transgenic mice, human CRP is transcribed exclusively in the liver, and expression is dependent on inflammation indicating that regulation of CRP is primarily at the transcriptional level (Ciliberto *et al.*, 1987). It was found that this regulation of CRP gene transcription is controlled by two inducible elements of *cis* acting DNA sequences at the 5' flanking area of the CRP gene which interleukin-6 (IL-6) and interleukin-1 (IL-1) activate (Li *et al.*, 1990; Ganapathi *et al.*, 1991). IL-1 upregulation of CRP gene expression (Pied *et al.*, 1989) has no effect when only IL-1 is used, but it did act synergistically with IL-6 (Castell *et al.*, 1990; Zhang *et al.*, 1995). Monocytes have been shown to induce CRP synthesis in hepatocytes (Goldman and Liu 1987) probably by release of cytokines IL-1 and IL-6. Hepatocyte nuclear factors have been shown to interact with the IL-6 responsive elements in the CRP gene and have both positive and negative regulation of CRP gene expression (Li and Goldman 1996). However, nitric oxide prevented the induction of CRP expression in the presence of both IL-6 and IL-1 (Voleti and Agrawal 2005) as did tumour necrosis factor (TNF) (Yap *et al.*, 1991), although TNF had no effect on CRP synthesis alone (Ganapathi *et al.*, 1991).

1.2.18 The structure of human CRP

Early studies showed that human CRP is an aggregate of subunits/protomers each held together by non-covalent interactions (Gotschlich and Edelman 1965). There is a single intrachain disulphide bond in each protomer (Gotschlich and Edelman 1965).

The 3.0Å crystal structure of human CRP was found to be a pentamer, each protomer having 206 amino acids (Shrive *et al.*, 1996) (see figure 1.1). The pentamer was part of a decamer resulting from crystal packing that also caused a loss of calcium in a single protomer per pentamer. Interprotomer interactions were: salt bridges 155-118 and 101-201 which are conserved in human SAP; and 123-197 and 118-202 which were unique to human CRP. A majority of the interprotomer interactions were between 115-123 loop, and 40-42 and 197-202 of the adjacent protomer.

Figure 1.1. Human CRP pentamer (Shrive *et al.*, 1996). α-helices are shown in red, β-sheets are shown in blue, loop regions are shown in yellow, and calcium ions are shown in pink.

The overall structure of CRP is a β-jellyroll topology consisting of two β-sheets of anti-parallel β-strands and a helix folded against it (see figure 1.2). This is similar to human SAP (see figure 1.4), however, there are structural differences between human CRP and human SAP: CRP has a helical turn in a loop connecting β-sheet strands C and D which displaces loops between strands E and F, and G and H from the position they are in human SAP.

Figure 1.2 Human CRP protomer. α-helices are shown in red, β-sheets are shown in blue, loop regions are shown in yellow, orange shows disulphide bridges and calcium ions are shown in pink (Source: Shrive *et al.*, 1999).

On one face of the CRP protomer is the calcium binding site, and on the other face is a cleft is not seen in the human SAP structure. This agrees with previous studies that suggest the calcium binding site is on the surface of CRP (Swanson *et al.*, 1991). The cleft on the

opposite face to the calcium binding site is formed by parts of the N-terminus, C-terminus and 177-182 loop and C-terminal end of the α-helix. The conserved disulphide bridge is located beneath the floor of this cleft.

The first calcium ion in human CRP is coordinated by Asp60, Asn61, Glu138, Asp140 and the main chain carbonyl of Gln139, whilst the second calcium ion is coordinated by Glu138, Asp140 Gln150 and Glu147. Each calcium ion is coordinated by five ligands each as there is one co-ordination from each residue except Glu147 which has two co-ordinations with the second calcium ion, and assuming that Asp60 shows single co-ordination. This suggests equal calcium binding affinity by human CRP. Trp67 and Arg47 also contribute to topology of the calcium binding region and calcium binding loop. There are hydrogen bonds between Arg47 and the main chain carbonyls of Ser151 and Ser149 which surround Gln150 that binds to calcium. Trp67 affects the orientation of Phe66 which is involved in PC binding (Shrive *et al.*, 1996).

The structure of human CRP indicates that the phosphate group of PC binds to the calcium ions, and that the three methyl groups of the choline group of PC interact with Phe66 in a hydrophobic pocket made by Phe66, Leu64 and Thr76. This agrees with previous work using antibodies against CRP peptide 134-148 that determines part of the calcium binding region which inhibited PC binding. This suggested that the calcium binding region on CRP is responsible for a minor structural effect of calcium on CRP (Swanson *et al.*, 1991). At one end of the hydrophobic pocket is the calcium binding site and at the other end are the side chains of Ser68, Ser74, and Glu81. Phe66 is conserved in human, mouse, rabbit, rat, and frog CRP, but is tyrosine in human, mouse, rat and hamster SAP (Shrive *et al.*, 1996) suggesting that the hydrophobic pocket (Phe66) methyl group interaction is for CRP only, which explains why human SAP does not bind to PC.

There were differences between the calcium binding area in protomers with two calcium ions bound and those with no calcium bound caused by crystal packing (Shrive *et al.*, 1996). In the calcium bound protomers, loop 140-150 was packed against the body of the protein with Glu147 coordinating with the second calcium ion meaning that the proteolytic cleavage site 145-147 was not exposed. However, in the calcium free protomer the 140-150 loop had shifted and had no interaction with the main body of protein, meaning that the proteolytic cleavage site was exposed. Early reports had shown that the presence of calcium had helped CRP resist urea or heat structural denaturation (Potempa *et al.*, 1983) which may be due to this structural change. Re-arrangement of the calcium binding loop also disrupts main chain Ser149 and Ser151 hydrogen bonds with Arg47. These differences may have been due to crystal packing which also makes identification of structure movement difficult.

The 3.15Å structure of human CRP without calcium bound revealed more detail on the difference between the calcium bound and non-calcium bound protomers (Ramadan *et al.*, 2002). The crystal structure was of two pentamers forming a face to face (α-helix to α-helix) decamer with the calcium binding regions on the outside faces of the decamer. A study has shown that pure CRP can form fibril like structures where pentamers are stacked face to face (Wang *et al.*, 2002), but the CRP decamer was thought to be formed due to crystal packing. It was stabilized by five surface contacts between the two pentamers through Thr173 to Pro179 for each protomer. The calcium binding loop 140-150 is the site of major conformational change between calcium bound and non-calcium bound CRP protomers. When calcium is bound to the protomer, the calcium binding loop folds in to bind Glu147 to calcium ion two and therefore protects the protease-sensitive site from cleavage. This agrees with findings that native CRP is resistant to enzymatic cleavage in the presence of calcium (Ying *et al.*, 1992a). When calcium is not bound, Glu147 does not interact with the calcium ions and the loop moves away from the body of the CRP protomer and is available for external interaction and proteolytic cleavage. Also available

for external interaction after loss of calcium is Glu81, Asp60 and Asn61 which are candidates for polycation binding due to their negative charges. Therefore, it can be said that binding of calcium stabilizes the CRP protomer and makes PC binding possible, whilst loss of calcium exposes proteolytic cleavage sites but also exposes possible areas of polycation binding which is calcium independent (DiCamelli *et al.*, 1980; Potempa *et al.*, 1981; Lee *et al.*, 2002).

1.2.19 The structure of human CRP with PC bound

The structure of human CRP with PC bound was solved in 1999 by Thompson, Pepys and Wood. The overall topology of CRP is the same as that previously published by Shrive *et al.*, (1996) with differences in some side chains causing change of 116-42 inter-subunit bond to 116-85 intra-subunit hydrogen bond. The β-sheets of the core of CRP protomer are at 22° rotation towards the five-fold axis and each pentamer shows five α-helices on one face (A face) and 10 calcium ions on the other face (B face). The electron density map shows one PC per protomer at the calcium binding site. Two oxygens from the phosphate group of PC interact with the two calcium ions, and the third oxygen of the phosphate group points away from the calcium binding site into the solvent. This would allow CRP to bind PC when the phosphate is in ester linkage with other molecules. The hydrophobic pocket formed by Thr76, Glu81, Gly79, Asn61 and Thr76 would encourage binding of PC analogues at position 2 of PC. Previous mutational analysis showed that Thr76 was a determinant of the PC binding site, and that Trp67, Lys57 and Arg58 do not directly contact PC but are required for the proper conformation of the binding site (Agrawal *et al.*, 1992, 1997). The choline group of PC interacts with Phe66, and the quaternary nitrogen (positively charged) interacts with the negatively charged side chain of Glu81 which agreed with later mutational studies that show that Phe66 and Glu81 are required for PC binding (Agrawal *et al.*, 2002). This same study also showed that mutations which

inhibited CRP binding at 66 and 81 do not affect fibronectin binding, and therefore the binding sites for PC and fibronectin are distinct.

On the A face there is a cleft seen previously in the structure by Shrive *et al.,* (1996) whose side walls are constructed from Ser5, Arg6, Gln203, Pro206, Trp187, Arg188, Asn160, Gly177, Leu176, Tyr175, His95 and Asp112; and the bottom of the cleft is lined by Asn158, His38, Leu37, Val94, and Asp112. The cleft runs halfway into the pore of the pentamer ending at Asp112 which is negatively charged, therefore there is a ring of negative charges around the pore. Mutagenesis studies of CRP complement activation indicated that Asp112 plays a direct role in C1q binding possibly by its position within the C1q binding site of CRP, whilst Lys114 and Arg116 affect C1q binding and complement activation indirectly (Agrawal and Volanakis 1994). Mutation of Lys114 to an amino acid with a negatively charged, neutral or non-polar side chain enhanced C1q binding and complement activation (Agrawal *et al.,* 2001).

It therefore appears that the pocket at the open end of the cleft is the C1q binding site CRP. It was proposed that this cleft and the negatively charged pore play a part in targeting C1q to the surface of cells with disturbed phospholipid membranes or bacterial ligands as the CRP pentamers bind to the cell membrane, stabilized by the multivalent PC binding sites. The close proximity of CRP pentamers bound to the cell membrane or bacterial ligand should allow each of the six C1q globular heads to bind to a CRP pentamer via the cleft and/or negatively charged pore, which then activates the complement system and provides cell lysis of compromised or bacterial cells. Mutational analysis also indicates that Glu88 influences the conformational change in C1q necessary for complement activation and that Asn158 and His38 probably contribute to correct geometry of the binding site (Agrawal *et al.,* 2001). Single mutations at Lys114, Asp169, Thr173, Tyr175 and Leu176 which are all situated around the cleft, also affect CRP-C1q binding (Bang *et al.,* 2005).

There are indications that the binding sites of human CRP for FcγRI, FcγRIIa and C1q overlap as mutational studies also showed that Thr173, Leu176 and Asn186 are important for binding of CRP to FcγRI and FcγRIIa, whilst Lys114 was implicated in binding to FcγRI but not FcγRIIa (Bang *et al.*, 2005). It was proposed that the hydrophobic area formed by the α-helix and loop 177-186 is the area of contact on CRP for FcγR (Bang *et al.*, 2005).

1.2.20 Human CRP and SAP interaction

Human SAP has been shown to bind human CRP in a 1:~1.2-1.5 ratio but only when it is immobilized and not in solution (Swanson *et al.*, 1992 (a, b), Christener and Mortensen 1994). The binding is calcium dependent, and is most likely to occur through the calcium binding region (Swanson *et al.*, 1992 (a, b)), residues 134-148 and/or the carboxy terminal region 191-206 on CRP (Christner and Mortensen 1994) but not through the PC binding site (Swanson *et al.*, 1992 (b)). Deglycosylated SAP still bound CRP and so SAP–CRP binding is not due to CRP lectin binding properties (Christner and Mortensen 1994).

1.3 Human serum amyloid P component (SAP)

1.3.1 Human SAP location in the body

Human SAP is a component of human serum and also amyloid tissue where it is found as amyloid P component (AP) (Bladen *et al.*, 1966; Shirahama and Cohen 1967). In humans, the normal levels of SAP in serum are ~30μg/ml which remains constant (Pepys and Baltz 1982; Steel and Whitehead 1994). SAP and AP are the same protein according to morphology, size, and antigenicity (Painter *et al.*, 1982), and both are pentamers of subunit size 25kDa (Pepys *et al.*, 1994). SAP is also a normal part of the glomerular basement membrane (Dyck *et al.*, 1980; Hutchcraft *et al.*, 1982), and elastic fibres in the connective

tissue of the skin, lung, gut, heart, and blood vessel walls (Breathnach *et al.*, 1981).

Amyloidosis is a disorder where soluble proteins are deposited extracellularly as abnormal insoluble fibrils, accumulation of which disrupts tissue structure and function (Pepys and Hirschfield 2003). In amyloid fibrils, AP is not an integral part of the amyloid fibril but is distinct and only 10-15% of the fibril (Skinner *et al.*, 1982). It was found that when the calcium concentration is low, SAP pentamers form a lattice or another regular array which was suggested to be the framework on which amyloid deposits form (Pinteric and Painter 1979). However, SAP binding to amyloid fibrils has been observed when the amyloid fibrils are mature (Holm-Nielsen *et al.*, 2000) through its calcium dependent affinity for glycosaminoglycans on the surface of the amyloid fibril (Pepys *et al.*, 1977, Pepys *et al.*, 1979b; Stenstad *et al.*, 1993; Urbanyi *et al.*, 2003). Several glycosaminoglycans and proteoglycans are associated calcium dependently such as dermatan/chondroitin sulphate and heparin sulphate through amyloid-P component (Stenstad *et al.*, 1993).

Aggregation of human SAP into polymers dependent on calcium and SAP concentration (Hamazaki 1989; Landsmann *et al.*, 1994) is inhibited by sulphated saccharides especially heparin but not unsulphated ones (Hamazaki 1989). This may be relevant to the deposition of SAP on amyloid fibrils of Alzheimers disease which have sulphated proteoglycans on their surface (Stenstad *et al.*, 1993; Holm-Nielsen *et al.*, 2000). Aggregation does not occur with mouse SAP (Baltz *et al.*, 1982) suggesting it is a property specific to human SAP, although mouse SAP is an acute phase protein like human CRP, which may explain the difference in aggregation properties. However, SAP enhances the induction of murine amyloidosis and may play an important role in the pathogenesis of human amyloidosis including Alzheimers disease in a similar way (Togashi *et al.*, 1997). Binding of SAP to amyloid is enhanced under slightly acidic pH such as in sites of inflammation, suggesting a role of SAP during inflammation (Holm-Nielsen *et al.*, 2000).

SAP is thought to play a pathogenic role in Alzheimers disease not only by showing

toxicity at normal serum levels against neuronal cells (Urbanyi *et al.,* 1994; Duong *et al.,*

1998) and possibly inducing apoptosis of neuronal cells after translocating to the nuclei

(Urbanyi *et al.,* 2003), but also by binding to the β-amyloid fibrils of Alzheimers disease

(Hamazaki 1995), neurofibrillary tangles and senile plaques (Perlmutter *et al.,* 1995). SAP

binding to mature amyloid fibrils of Alzheimers disease may contribute to their persistence

by preventing their degradation, as SAP prevents proteolysis when bound (Tennent *et al.,*

1995). Reduction in amyloid fibril formation when the SAP gene is deleted shows a direct

participation of SAP to amyloidosis *in vivo* (Botto *et al.,* 1997). SAP levels have been

found to be high in cerebrospinal fluid of humans with Alzheimers disease which may be

affected by the presence of cerebral amyloidosis (Hawkins *et al.,* 1994). Complement C4 –

binding protein has been found on cerebral cortical amyloid deposits associated with SAP

which suggests involvement of complement in the pathogenesis of Alzheimers disease

(Kalaria and Kroon 1992).

Drugs have been designed to inhibit SAP binding *in vivo,* and therefore hinder new

amyloid fibril formation and possibly reduce the stability of existing amyloid deposits to

promote their regression. The drug R-1-[6-[R-2-carboxy-pyrrolidin-1-yl]-6-oxo-

hexanonyl]pyrrolidine-2-carboxylic acid (CPHPC) was developed which clinical trials

showed is tolerated in the body, and rapidly removes SAP from the blood by the liver

(Pepys *et al.,* 2002). This alters the SAP:AP ratio and therefore SAP flows from tissues

into the blood where it is targeted by the drug. This therefore leads to removal of SAP from

the amyloid deposits.

1.3.2 Human SAP binding properties

As well as binding to heparin and sulphated sugars such as heparin sulphate, dermatan

sulphate and chondroitin-6-sulphate that are present in amyloid fibril deposits including

those present in Alzheimers disease (Hamazaki *et al.*, 1987; Loveless *et al.*, 1992; Heegard *et al.*, 1996; Danielson *et al.*, 1997), SAP also binds to sugars such as galactose, N-acetylglucosamine and glucuronic acid with a sulphate at carbon position 3 (Loveless *et al.*, 1992). SAP calcium dependent binding to heparin is through residues in the sequence areas 27-38, 88-120, and 192-203 (Cardin and Weintraub 1989; Heegard *et al.*, 1996).

Like CRP, SAP shows binding to sugars phosphorylated through carbon 6, but also to those through carbon 1 (Loveless *et al.*, 1992). It has also been proposed that SAP functions as a mannose binding protein as it binds to ligands such as ovalbumin, thyroglobulin, and β-glucuronidase which contain the mannose oligosaccharide sequence αDMan1-3DMan, and αDMan1-6DMan (Kubak *et al.*, 1988). SAP binds to methyl 4,6-O-(1-carboxyethylidene)-β-D-galactopyranoside) (MOβDG) which is a 4,6 cyclic pyruvate acetal of galactose, but binds less to 4,6 pyruvate acetal of mannose, and does not bind to 4,6 pyruvate acetal of glucose suggesting a preference of SAP binding to the galactose structure and not the mannose structure or glucose structure (Hind *et al.*, 1985). SAP-MOβDG binding inhibits calcium dependent binding reactions of SAP (Hind *et al.*, 1984) and removal of sialic acids from SAP reduced the affinity for agarose (which contains MOβDG) by 7% suggesting that sialic acids are partly responsible for SAP-agarose binding (Hamazaki 1990).

CRP and SAP can be distinguished from each other by their calcium binding affinity for PC and PE respectively. Although CRP can bind to both PC and PE, SAP does not bind PC at all or if so very weakly (Schwalbe *et al.*, 1992, Christner and Mortensen 1994). Human SAP binds to PE-agarose and agarose, but only slightly to PC-sepharose, whilst human CRP binds to PC-sepharose and PE-agarose indicating that the choline side chain affects SAP binding to PC (Christner and Mortensen 1994). This suggested that there were differences between the PC/PE binding sites of CRP and SAP which was supported by

different reactions with site specific antibodies (Christner and Mortensen 1994). There is

also evidence of SAP forms with various agarose binding affinity that show molecular

weights of 25.5kDa and 24.5kDa which suggests microheterogeneties or allotypic forms

(Kubak *et al.*, 1988).

Human SAP has also been shown to bind to fibronectin (like human CRP does) and C4

binding protein calcium dependently and in serum (De Beer *et al.*, 1981; Sørensen *et al.*,

1995). This binding only occurred when SAP was dimerized, and native unbound and

aggregated SAP did not bind to fibronectin or C4 binding protein (De Beer *et al.*, 1981).

As SAP in amyloid tissue can form a dimer (Prelli *et al.*, 1985) this indicates that

fibronectin and C4 binding protein may be bound when SAP is bound to amyloid fibrils,

perhaps increasing their pathology. There are indications that the C4 binding protein

binding site is the same as, or associated with, that for phosphate containing compounds

because human SAP-C4 binding protein complexes were disrupted using PE and also PC,

phosphate and glycerophosphate (Schwalbe *et al.*, 1992).

1.3.3 Human SAP interaction with complement

SAP was shown to agglutinate complement-coated erythrocytes by calcium dependent

binding to C3b and C4b (Hutchcraft *et al.*, 1981, 1982). Later studies showed that SAP

interacts with the classical complement system both through direct interaction with C1q

and by regulation through C4b-binding protein. SAP monomers bound C1q calcium

dependently but they did not induce complement activation, whilst SAP trimers and dimers

both bound to C1q at the collagen like region calcium independently and induced

complement activation (Ying *et al.*, 1993). SAP trimers showed native immunoreactivity

so any changes in structure when aggregated were small (Ying *et al.*, 1993), but the

difference in C1q activation suggests that aggregated SAP has different binding properties

to monomeric SAP. It is possible that at sites of aggregation of SAP in the human body, the classical complement cascade is activated. SAP regulates the classical complement pathway by forming a calcium dependent complex with C4b-binding protein, which inhibits its binding to C4b (Garcia de Frutos and Dahlback 1994). The SAP-C4b binding protein interaction is inhibited by PE and heparin suggesting that they bind at an overlapping or identical site on SAP.

1.3.4 Human SAP and the prevention of autoimmunity

One possible role of SAP in the human body is to prevent antinuclear autoimmunity by binding to nuclear antigens such as chromatin. SAP binds to single and double-stranded DNA (Pepys and Butler 1987), chromatin (Pepys and Butler 1987; Breathnach *et al.*, 1989, Hicks *et al.*, 1992, Bharadwaj *et al.*, 2001), nuclei (Breathnach *et al.*, 1989), histones (Hicks *et al.*, 1992), and nucleosome core particles (Pepys and Butler 1987; Butler *et al.*, 1990). It has been shown to slow down degradation of long chromatin by solubilizing it, and displacing H1-type histones that are used for chromatin condensation and folding (Butler *et al.*, 1990; Bickerstaff *et al.*, 1999). Binding of SAP to histones and chromatin was calcium dependent (Hicks *et al.*, 1992) similar to PE binding, indicating it is through a similar or identical binding site. SAP-chromatin binding did not activate the classical complement pathway suggesting another method of clearance of nuclear antigens (Bickerstaff *et al.*, 1999). Mice lacking SAP develop antinuclear autoimmunity suggesting that SAP does indeed play a role in managing nuclear debris from dead cells (Bickerstaff *et al.*, 1999). SAP has also been shown to have calcium dependent binding to apoptotic cells via exposed PE residues suggesting that SAP is associated with dealing with apoptosis *in vivo* (Familian *et al.*, 2001).

1.3.5 Human SAP and cellular immunity

SAP binds to human polymorphonuclear leukocytes in a calcium dependent manner, and this binding is inhibited by SAP, CRP, and IgG (Landsmann *et al.*, 1994). SAP was not degraded by neutrophils normally, but peptides from SAP degradation by neutrophil enzymes also inhibited SAP binding to human polymorphonuclear leukocytes (Landsmann *et al.*, 1994). SAP binding to leukocytes occurs via FcγRI and FcγRII and enhances the phagocytosis of particles (Bharadwaj *et al.*, 2001; Mold *et al.*, 2001), which may provide clearance of SAP ligands such as chromatin, and hence play apart in prevention of antinuclear autoimmunity and also host defence (Bharadwaj *et al.*, 2001). Mouse SAP also binds macrophages but it is thought binding is through a mannose-6-phosphate receptor on the macrophage surface not FcγR (Siripont *et al.*, 1988).

1.3.6 The human SAP gene and amino acid sequence

Initial studies using N-terminal sequencing of cyanogen bromide, trypsin and V-8 protease cleaved peptides of SAP gave a primary structure which was 50% identical to CRP and showed a conserved disulphide bond (Anderson and Mole 1982). The human SAP amino acid sequence was later determined to be 204 amino acids long with an additional 19 amino acid signal peptide (Mantzouranis *et al.*, 1985). The gene for human SAP was localized to chromosome 1, and it was suggested that it was close to the gene for human CRP (Mantzouranis *et al.*, 1985). Additional sequence studies between human CRP and human SAP show a 51-54% amino acid identity between the two (Woo *et al.*, 1985; Mantzouranis *et al.*, 1985).

1.3.7 The human SAP structure

SAP is a macromolecule of non-covalently bound subunits each ~25kDa (Thompson and Enfield 1978) composing a decamer ~255kDa (Painter *et al.*, 1982; Kubak *et al.*, 1988). Early studies indicated that SAP was a pentamer in solution, but decamerisation occurred in physiological conditions (Wood *et al.*, 1988). Electron microscopy showed that SAP was pentameric in low calcium concentrations and formed aggregates only in the absence of serum (Painter *et al.*, 1982). This was confirmed by later studies which indicated that SAP was a pentamer in serum, but changed to a decamer when purified by affinity chromatography and ion-exchange chromatography, and *vice versa* when incorporated into serum (Sørensen *et al.*, 1995). Density gradient ultracentrifugation also indicated that SAP was a pentamer in serum and not a decamer (Hutchinson *et al.*, 2000). Therefore, it is likely that SAP in serum is prevented from aggregation by calcium and ligands, but when serum is removed, calcium promotes SAP aggregation (Hutchinson *et al.*, 2000).

The crystal structure of the SAP pentamer (see figure 1.3) showed five subunits each 204 amino acids long with a structure of anti-parallel β-strands arranged in two sheets with β-jelly roll topology (Emsley *et al.*, 1994). The two β sheets form a hydrophobic core one end of which is formed by the amino terminus and carboxy terminus in the solvent, and the other end of the core is involved in protomer-protomer interactions. The β strands are in planes normal to the five fold axis and there is an α-helix folded on top of the β-sheet (see figure 1.4).

Figure 1.3 Human SAP pentamer. α-helices are shown in red, β-sheets are shown in blue, loop regions are shown in yellow, and calcium ions are shown in pink. Source: T. J. Greenhough.

Figure 1.4 Human SAP protomer. α-helices are shown in red, β-sheets are shown in blue, loop regions are shown in yellow, orange shows the disulphide bridges and calcium ions are shown in pink. (Source: Shrive *et al.*, 1999).

Previous studies suggested the presence of a single intrachain disulphide bond (Thompson and Enfield 1978; Anderson and Mole 1982; Painter *et al.*, 1982) which is found between Cys36 and Cys95 and links strands B and C in one sheet. There are also strand interactions with strands from other protomers such as strands G, I, and J interacting with strand N and loops between strands A and B, C and D, G and H, and K and L of the

adjacent protomer. The interactions between protomers compose of hydrogen bonds between the main chain of amino acids, three salt bridges, and some hydrophobic contacts.

There is an N-linked oligosaccharide at Asn32 which only shows electron density for one saccharide residue (Emsley *et al.*, 1994). This agrees with work that showed each SAP subunit/protomer contains a single N-linked oligosaccharide at Asn32 (Anderson and Mole 1982; Prelli *et al.*, 1985; Hamazaki 1990: Pepys *et al.*, 1994). The SAP oligosaccharide is a biantennary complex of which there was only a single component of Galβ4GlcNAcβ2Manα6(Galβ4GlcNAcβ2Manα3)-Manβ4GlcNAcβ4GlcNAc-ol (Hamazaki 1990; Pepys *et al.*, 1994). Multiple isoforms of SAP were shown in mice and humans that differed in size and isoelectric point, which was not due to sialic acid content (Nybo *et al.*, 1998). During acute phase infection in mice, the concentration and number of isoforms of SAP increased, indicating an *in vivo* modification of SAP in mice which may also occur in humans (Nybo *et al.*, 1998).

The calcium ion binding sites identified in the SAP structure (Emsley *et al.*, 1994) are as follows: calcium ion one is coordinated by side chains of Asp58, Asn59, Glu136, Asp138 and the main chain carbonyl of Gln137; calcium ion two is coordinated by side chains of Glu136, Asp138 and Gln148. This means calcium ion two has fewer ligands than calcium ion one, and may be preferentially unloaded. The major site for protease cleavage of SAP (144-145) may be susceptible to cleavage when SAP is not bound to calcium as the calcium ions hold this area of sequence close to the protomer body. Binding of 4,6-O-(1-carboxyethylidene)-β-D-galactopyranoside (MOβDG) was shown to be through the two calcium ions via the carboxyl group, and two hydrogen bonds between the oxygen atoms at positions 4 and 6 and the amide nitrogen atoms of Gln148 and Asp59. Therefore, only the methyl 4,6-O-(1-carboxyethylidene) ring forms interactions with the calcium ions and not the galactopyranoside. The methyl group sits in a hydrophobic pocket formed by Leu62,

Tyr64 and Tyr74 which is not present in human CRP and may explain poor binding of

MOβDG by CRP (Emsley *et al*, 1994). Analysis of electron density maps indicated that PE

binding by SAP is at the same site of MOβDG suggesting a direct interaction between the

phosphate group and calcium ions. As binding occurs through the phosphate group on PE

it is also likely that SAP binds DNA through the phosphate backbone of DNA.

1.3.8 The human SAP structure with MOβDG bound

A detailed structure of SAP bound with MOβDG showed that not only did it bind

directly to two calcium ions through the carboxyl group, and hydrogen bonds made

between 4' oxygen and 6' oxygen of the sugar ring and the amide nitrogens of Gln148 and

Asn59, but that it also bound using the 3' oxygen of the sugar ring and Gln148, and 1'

oxygen of the sugar ring and Lys79 (Thompson *et al.*, 2002). It also showed that the

calcium ions also oriented the Gln148 and Asn59 side chains to interact with the 4' and 6'

oxygen of the sugar ring. The binding observed by 1' oxygen of sugar ring and Lys79 may

be due to crystal contact differences. There were two crystal forms in the same study, one

of which showed only a single calcium and water replacing the second calcium ion in one

of the subunits, and the other showed two calciums and one MOβDG in all subunits. It is

possible that the concentration of calcium at physiological strength is not high enough to

saturate all sites. The water in place of a single calcium does not appear to affect the

protomer structure of SAP but this may not be the case if both calcium ions are absent. The

crystal forms of SAP were decamers of two pentamers stacked face A to face B, not face A

to face A as found earlier for dAMP (Hohenester *et al.*, 1997). The bound MOβDG form a

layer between the pentamers in the decamer, but only interacted with the pentamer which

they were bound to and not the adjacent pentamer. The arrangement of pentamers in the

decamer was the same as those in a single pentamer, and there were few contacts between

pentamers: Arg146-Asp3, Lys143-Glu188, and Lys79-Glu167 per protomer.

1.3.9 The human SAP structure with dAMP ligand bound

Ligands were screened for their ability to inhibit calcium induced aggregation of SAP. dAMP and AMP were found to be as effective as PE and MOβDG at inhibiting aggregation of SAP so crystallization trials were performed using dAMP (Hohenester *et al.*, 1997). The structure of the SAP decamer with dAMP bound was found at 2.8Å resolution (Hohenester *et al.*, 1997). The pentamers in the decamer interacted through their calcium binding sites and there were a few weak contacts across the interface of the two pentamers through the dAMP molecules. The interaction of pentamer to pentamer was through base stacking of the adenine rings in dAMP and there were no other hydrophobic contacts stabilizing the decamer. Decamerisation of SAP by same face interactions by dAMP prevents larger aggregation of SAP as the faces only interact with each other. dAMP binds the calcium ions via the phosphate group and one oxygen has an extra hydrogen bond from the side chain of Asn59. Other interactions of dAMP with SAP are hydrogen bonds from Gln148 and Tyr64 to the 2'oxygen of the deoxyribose ring, and from Asp145 to the amino group of purine ring at position 6. The specific affinity of SAP for the adenine ring of dAMP explains the poor affinity of binding of other nucleotides that contain other bases such as guanine or the pyrimidines (Hohenester *et al.*, 1997).

1.4 Pentraxins from other organisms

1.4.1 Pentraxins in amphibians

CRP and SAP like proteins have also been found in a range of other organisms from invertebrates, like the horseshoe crab, to fish, amphibians and other mammals. Amphibians have both CRP-like proteins and antibodies, and can be seen as a transition between organisms which require CRP to act as a major host defence protein as they have no or little adaptive immunity such as invertebrates, and those which have more complex immune systems and require CRP to a lesser degree such as mammals. The African clawed

toad (*Xenopus laevis*) has a CRP like protein 222 amino acids long, which is longer than that of human CRP. It does not have a heat shock consensus sequence or IL-6 responsive element within its gene unlike human CRP (Lin and Liu 1993). It is most probably formed in the liver like human CRP, but is only expressed at levels of <1μg/ml and expression is not induced by inflammation so it is therefore not an acute phase protein (Lin and Liu 1993).

1.4.2. Pentraxins in elasmobranchs

CRP and SAP like proteins have been identified in the elasmobranch smooth dogfish, *Mustelus canis*. Dogfish CRP binds to phosphocholine, and the dogfish SAP binds to sepharose calcium dependently (Robey *et al.*, 1983). The dogfish CRP protomer was ~26kDa and the SAP protomer ~27.5kDa. Although some heterogeneity in dogfish CRP has been indicated (Robey and Liu 1983; Samudzi *et al.*, 1993), the whole sequences of CRP and SAP have yet to be found and so firm conclusions cannot be drawn.

1.4.3 Pentraxins in fish

Pentraxin like proteins have been found in fish, although the properties are much less consistent than those of mammals. CRP has been identified in rainbow trout (*Salmo gairdneri* and *Oncorhynchus mykiss*) which shows calcium dependent binding to pneumococcal C-polysaccharide (Murai *et al.*, 1990; Murata *et al.*, 1995) and an acute phase response to bacteria *Vibrio anguilarum* (Kodama *et al.*, 1989) and environmental pollutants (Winkelhake *et al.*, 1983). In rainbow trout, CRP activated the complement system in the presence of C-polysaccharide (Nakanishi *et al.*, 1991). SAP has also been found which binds to agarose calcium dependently and has N-terminal amino acid homology to SAP from plaice, humans, hamsters, rat and mouse (Murata *et al.*, 1994).

Rainbow trout CRP was 66kDa in size with subunits of approximately 26.6kDa, whilst SAP has a molecular weight of 100kDa composed of subunits ~32kDa.

CRP has also been identified in common carp (*Cyprinus carpio*) as a PC binding protein with N-linked glycosylation (Cartwright *et al.*, 2004). In carp (*Catla catla*) CRP is also an acute phase protein like in rainbow trout but varies in subunit size (22 and 29kDa) with varying glycosylation (Paul *et al.*, 1998) meaning that CRP in carp represents a family of proteins rather than a single protein. CRP basal levels in serum from European carp lines ranged from 2.9 ± 0.15 to $12.57\pm1.19\mu g/ml$, and increased several fold during infection with *Aeromonas hydrophila* indicating that it is an acute phase protein (MacCarthy *et al.*, 2008).

CRP and SAP have also been identified in salmon (*Salmo salar*) where CRP reacts with phosphocholine calcium independently and exists in monomeric and dimeric forms of molecular weights 80 and 160kDa respectively (Lund and Olafsen 1998b). The salmon SAP was found to be an acute phase protein which had calcium dependent agarose binding properties, and was a pentamer formed from glycosylated subunits 37kDa in size (Lund and Olafsen 1999). The same study also identified pentraxin-like proteins from the common wolffish (*Anarhichas lupus*) (25kDa subunit), cod (*Gadus morhua*) (30kDa subunit), and halibut (*Hippoglossus hippoglossus*) (28kDa subunit). Both wolffish and halibut pentraxins were agarose binding proteins, but the cod pentraxin was a PC binding protein. Despite all the pentraxins mentioned coming from fish, they had approximately equal N-terminus sequence identity with other fish, as they did with human CRP and SAP N-termini.

A pentraxin like protein from the Snapper (*Pagrus auratus*) also bound calcium dependently to agarose, but increased in concentration only two fold after LPS injection making it a minor acute phase protein, although it did have the ability to activate

37

complement (Cook *et al.*, 2003). It was ~200kDa in size composed of non covalently

bound subunits 23-26kDa in size (Cook *et al.*, 2003). CRP found in the spotted murrel

(*Channa punctatus*) by calcium dependent binding to PC, and was found to be a

pentameric acute phase protein (Ghosh and Bhattacharya 1992) of molecular weight

141kDa (Mitra and Bhattacharya 1992).

1.4.4 Pentraxins in mammals

Rat CRP is a pentamer with two subunit types 23.1kDa and 26.5kDa both glycosylated

and is 211 amino acids long with a 19 amino acid signal sequence (Rassouli *et al.*, 1992).

In rats, CRP is the major acute phase protein whilst SAP is not an acute phase protein (De

Beer *et al.*, 1982). Rat SAP is a single pentamer of subunits 24.5kDa in size (De Beer *et

al.*, 1982) and 208 amino acids long (Dowton and McGrew 1990). Mice CRP is 206 amino

acids long with a 19 amino acid signal sequence (Whitehead *et al.*, 1990). There is

conflicting evidence surrounding mice CRP whether it is an acute phase protein (Bodmer

and Siboo 1977; Le *et al.*, 1982 (a)) or present in small amounts showing some acute phase

properties (Pepys *et al.*, 1979; Whitehead *et al.*, 1990). Mouse SAP on the other hand, is a

major acute phase protein (Pepys *et al.*, 1979) and can increase up to 10 times in

concentration during the acute phase of a disease (Sipe *et al.*, 1982). This corresponds to an

increase of SAP synthesis by the liver (Le *et al.*, 1982 (a)) which may be caused by IL-1

and macrophages (Sipe *et al.*, 1982; Le *et al.*, 1982 (b)). Mouse SAP is 204 amino acids

long (Nishiguchi *et al.*, 1988) and forms a glycosylated pentamer (Taylor *et al.*, 1982) with

glycosylation occurring at amino acids 32-34 of each subunit (Nishiguchi *et al.*, 1988).

Each subunit is of size ~25.3kDa and the pentamer can be isolated using calcium

dependent affinity chromatography to sepharose (Pepys *et al.*, 1979 (a)).

Rat CRP was shown to bind C-polysaccharide, but did not agglutinate C-polysaccharide

coated erythrocytes or initiate complement (De Beer *et al.*, 1982). However, more recent studies have shown that rat CRP does activate complement (Padilla *et al.*, 2005). Mouse CRP interacts with C-polysaccharide calcium dependently (Bodmer and Siboo 1977; Pepys 1979). In mice it was also suggested that SAP binds to rough bacteria and reduces their clearance from the blood, but had no effect on smooth bacteria (Nourasadeghi *et al.*, 2000).

In Syrian hamsters a protein called female protein was found with N-terminus homology to human CRP and SAP, and similar overall pentameric structure of subunits ~30kDa in size that were non-covalently attached (Coe *et al.*, 1981). Similar to rat CRP, Syrian hamster female protein had a primary peptide structure of 211 amino acids, but had a 22 amino acid signal peptide (Rudnick and Dowton 1993a). Female protein is 69% identical in amino acid sequence to human SAP, and 50% identical to human CRP (Dowton *et al.*, 1985). Female protein was found to be a constituent of Syrian hamster amyloid tissue similar to human SAP and human amyloid, but with gender-linked increases in amyloidosis in females (Coe and Ross 1985). Control of protein expression is performed at the transcriptional level (Dowton *et al.*, 1985), and female protein is found in low levels in males and increases in concentration five fold during the acute phase response, whilst in females it is in much higher basal levels but decreases in concentration during the acute phase response (Coe and Ross 1983). Female protein expression decreased with the presence of Il-1, IL-6 and tumour necrosis factor in female hamsters but not male hamsters, suggesting the presence of IL-1 and IL-6 responsive elements in the female protein gene in females but not males (Rudnick and Dowton 1993a). The increase in female protein concentration in males that have been castrated or treated with oestrogen, and the suppression of female protein in females treated with testosterone, suggested that control of female protein expression is hormone linked (Coe 1977, Rudnick and Dowton 1993). However, in Armenian hamsters, SAP (or female protein in Syrian hamsters) is not a major acute phase protein and does not differ in expression by males or females (Rudnick and

Dowton 1993b). Syrian hamsters also had CRP which is expressed regardless of gender of the hamster (Dowton and Holden 1991). The hamster CRP gene is 206 amino acids long with 19 amino acids signal peptide and gene transcription is affected by IL-1, IL-6 and TNF (Dowton and Holden 1991).

In rabbits, CRP is an acute phase reactant which localises at sites of inflammation (Kushner and Kaplan 1961) and has been shown to partially protect rabbit CRP-transgenic mice from lethal doses of LPS compared to litter mates with suppressed CRP expression (Xia and Samols 1997). Rabbits have only one CRP gene copy like the other mammalian CRP's, which produces a protein 205 amino acids long with a 20 amino acid signal sequence (Syin *et al.*, 1986; Hu *et al.*, 1986). The N-terminus of rabbit CRP has significant homology with the N-terminus of human CRP (Osmand *et al.*, 1977). Mutational studies with rabbit CRP showed that mutation of Phe66 and Glu81 reduced PC binding, and that mutation of Tyr175 affected complement activation (Black *et al.*, 2003). This showed a high similarity of rabbit CRP PC binding and complement interaction to human CRP.

CRP-like proteins have also been found in larger mammals. In horses, CRP was isolated from serum using its calcium dependent affinity to C-polysaccharide and was shown to be a pentamer of non-glycosylated subunits 23kDa each (Takiguchi *et al.*, 1990). In the harbour porpoise (*Phoca vitulina*), CRP was purified by calcium dependent PC chromatography and had a subunit of size ~25kDa (Funke *et al.*, 1997). An internal peptide from harbour seal CRP showed it to be similar in sequence to other mammalian CRPs.

Overall, CRP and SAP appear to have kept the PC/C-polysaccharide and agarose or sepharose binding respectively. In some organisms it is clear that CRP acts as the major acute phase protein whether in mammals and fish; in others SAP is the major acute phase protein (e.g. mice); and in some it is unclear whether CRP or SAP is the major acute

protein. It is also interesting that CRP in humans has been shown to activate complement, as it does in trout and snapper, but there are conflicting views on rat CRP activation of complement.

1.5 Physiology and immunity in horseshoe crabs

1.5.1 Horseshoe crab physiology

Horseshoe crabs fall into three genera: *Limulus, Tachypleus*, and *Carcinoscorpius* all of which are under the order Xiphosurida and suborder Limulina (Størmer 1952). At present there are four distinguishable species of horseshoe crab; three that inhabit the East coast of Asia, and one that inhabits the East coast of North America. The three Pacific horseshoe crabs are *Tachypleus tridentatus, Tachypleus gigas,* and *Carcinoscorpius rotundicauda* and the single Atlantic horseshoe crab is *Limulus polyphemus* (see figure 1.5). It is assumed that the Atlantic and Asian horseshoe crabs diverged approximately 135 million years ago (Sekiguchi 1988).

Figure 1.5 *Limulus polyphemus.* Source: http://www.mbl.edu/marine_org/marine_org.

Horseshoe crabs live between the intertidal zone of a bay to the continental shelf, only coming to land to breed and lay their eggs except *Carcinoscorpius rotundicauda* which can migrate into fresh water (Størmer 1952). Egg development to hatching takes approximately 14 days, after which juvenile horseshoe crabs fend for themselves in intertidal and shallow

water areas. Maturity of juveniles into adult horseshoe crabs takes approximately 18 growth stages (instars) in 10 years and depends on diet, feeding rate and environmental conditions. Horseshoe crabs typically feed on molluscs and worms by digging into sand and probing with their legs. Although they are Arthropods, horseshoe crabs are not actually crabs but are chelicerates as they have feeding appendages called chelicerae and they have an exoskeleton made of chitin. Underneath the hard carapace that has two compound eyes, there are 13 pairs of appendages which have roles in locomotion, feeding, burrowing and water flow across the book gills. Horseshoe crabs also have a telson (tail) which is used to correct itself when it is turned over on its underside. Also underneath the carapace are the book gills composed of lamella that adsorb oxygen from the water into the blood. By moving the gills, the horseshoe crab keeps a fresh flow of water over them, thereby providing more oxygen and allowing de-sorption of carbon dioxide from the blood.

The blood of a horseshoe crab is blue due to the oxygen carrying protein being copper based haemocyanin, rather than the iron-based haemoglobin in humans. This haemocyanin is the most abundant protein in the haemolymph, and exists extracellularly unlike haemoglobin in humans which is within erythrocytes. The haemolymph contains only one cell; the motile granular amoebocyte also called granulocyte or haemocyte, which is the main defence against invading pathogens within the haemolymph. On recognizing invading bacteria, the cell de-granulates releasing components that provide clotting of the haemolymph thereby trapping bacteria so that they cannot move. Also present in the haemolymph of horseshoe crabs are a variety of proteins including those that have similar properties to proteins found in humans such as α_2-macroglobulin, lectins and the pentraxins.

1.5.2 Immunity in horseshoe crabs

The first defence of the horseshoe crab against invading pathogens is the hard

exoskeleton on which there have been found cuticular proteins similar to known antimicrobial peptides (Ijjima *et al.*, 2005). A chitin binding transglutaminase called Caraxin in the horseshoe crab cuticle may be involved in host defence (Matsuda *et al.*, 2007a) as could Stabling, a haemocyte derived protein which binds chitin and agglutinates bacteria by promoting clotting mesh formation at sites of injury (Matsuda *et al.*, 2007b). If the exoskeleton is breached, microbes can enter the tissue and blood of the horseshoe crab where there are arrays of antimicrobial defences. The amoebocytes contain both large and small secretory granules which contain specific proteins: small granules contain the antimicrobial or bacterial agglutinating proteins such as big defensin (Saito *et al.*, 1995; Kawabata *et al.*, 1997) and tachyplesins, whilst large granules contain clotting factors, protease inhibitors and some proteins with bacterial agglutinating activities (Iwanaga *et al.*, 1998). The secretory granules of the amoebocyte are released spontaneously in the absence of endotoxin, but released at a faster rate in the presence of endotoxin (Armstrong *et al.*, 1980; Armstrong and Rickles 1982). The granular amoebocyte becomes adhesive and motile when the secretory granules are released suggesting the possibility of migration into tissue (Armstrong 1977, 1979, 1980).

The amoebocytes contain proteins which form a clot in the presence of an endotoxin. The clotting cascade can be activated by either gram negative bacteria or LPS (Bang 1956, Levin and Bang 1964), or (1-3)-β-D-glucan from fungi cell walls (Iwanaga *et al.*, 1992) (see figure 1.6). Factor C in the cascade has a high affinity LPS binding site which activates it and starts the cascade (Tan *et al.*, 2000). Two forms of factor C have been found in *Carcinoscorpius rotundicauda* both of which have endotoxin receptors to activate their catalytic sites (Ding *et al.*, 1993). The clotting cascade is inhibited by certain proteins released from the amoebocyte granules which may indicate some regulation (Armstrong and Quigley 1985).

```
Endotoxin or LPS                              1,3-β-D-Glucan

Factor C  →  Factor C        Factor G  ←      Factor G

Factor B  →  Factor B

Pro-clotting enzyme  →  →  Clotting enzyme

Coagulin and peptide C  ←  Coagulogen protein

Cross linking and gel clot
formation
```
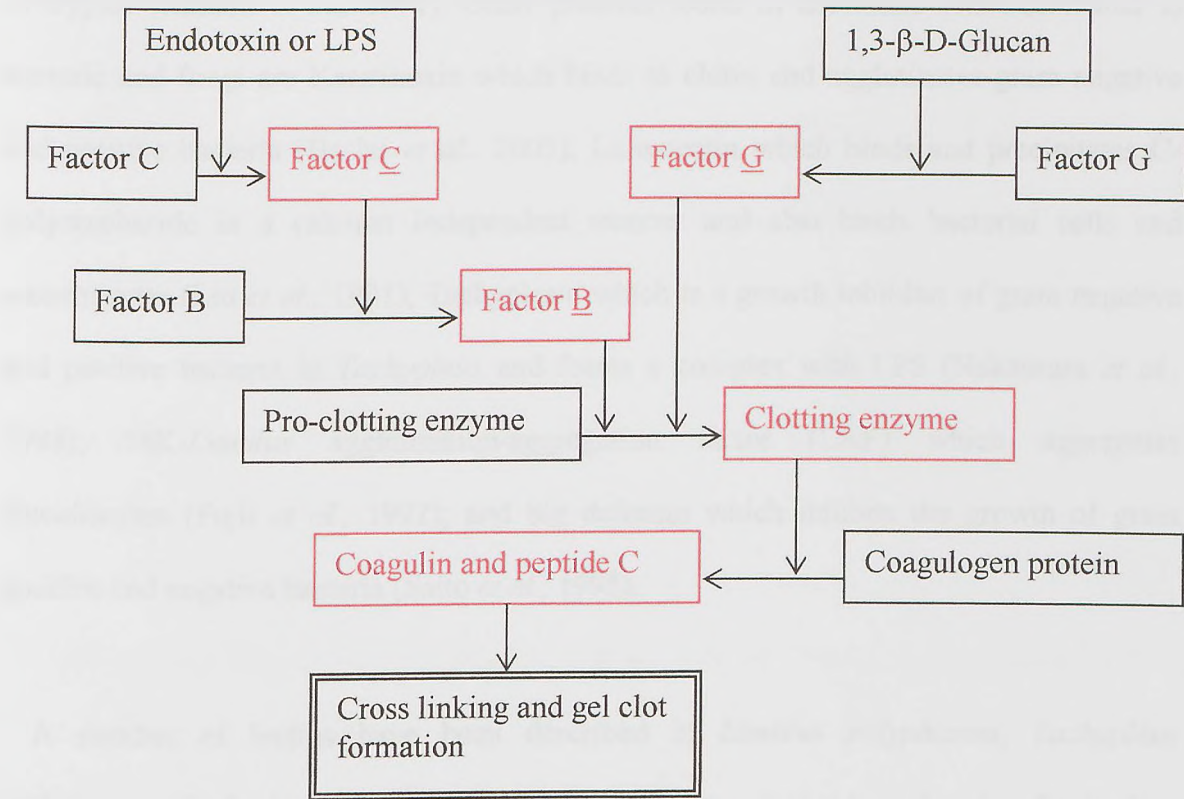
Figure 1.6 Clotting cascade of *Limulus polyphemus* (adapted from Shuster 2003). Red boxes and text highlight activated enzymes.

The coagulin forms the fibrils of the clot trapping any invading bacteria or fungal spores. The clotting cascade is used in a commercial way as the *Limulus* amoebocyte lysate (LAL) assay to test for the presence of endotoxin in pharmaceuticals, food, water supplies, and blood (Levin 1985). It is likely that the bactericidal activity exhibited by *Limulus* serum varies according to bacteria present, environmental conditions, and individual horseshoe crabs (Furman and Pistole 1976).

The amoebocytes also contain many other proteins and peptides that are released during de-granulation in addition to the clotting cascade. There is also present in amoebocytes an anti-coagulent, anti-LPS factor (LALF) which inhibits activation of the clotting cascade and shows antibacterial activity against gram negative bacteria (Aketagawa *et al.*, 1986). LALF peptide 31-52 was shown to have anti-inflammatory properties on human cells (Vallespi *et al.*, 2000). *Limulus* endotoxin binding protein-protease inhibitor also found in the amoebocytes, binds specifically to LPS and *E. coli*, and inhibits the proteolytic activity

of trypsin (Minetti *et al.*, 1991). Other proteins found in horseshoe crabs that bind to bacteria and fungi are Keratinoxin which binds to chitin and agglutinates gram negative and positive bacteria (Hashii *et al.*, 2005); Limunectin which binds and precipitates C-polysaccharide in a calcium independent manner and also binds bacterial cells and amoebocytes (Liu *et al.*, 1991); Tachyplesin which is a growth inhibitor of gram negative and positive bacteria in *Tachypleus* and forms a complex with LPS (Nakamura *et al.*, 1988); 18K-*Limulus* agglutination-aggregation factor (LAF) which aggregates amoebocytes (Fujii *et al.*, 1992); and big defensin which inhibits the growth of gram positive and negative bacteria (Saito *et al.*, 1995).

A number of lectins have been described in *Limulus polyphemus*, *Tachypleus tridentatus*, *Tachypleus gigas*, and *Carcinoscorpius rotundicauda*. In *Tachypleus tridentatus*, Tachylectins 5a and 5b agglutinate gram positive and negative bacteria (Chen *et al.*, 2001), and enhance the antimicrobial activity of big defensin (Gokudan *et al.*, 1999). Homologues of Tachylectins have been found in *Carcinoscorpius rotundicauda* called Carcinolectins 5a and 5b which also recognize pathogens (Zhu *et al.*, 2005) probably through LPS (Ng *et al.*, 2007). Galactose binding proteins have also been found in both *Tachypleus tridentatus* (Chiou *et al.*, 2000) and *Carcinoscorpius rotundicauda* (Ng *et al.*, 2007). *Tachypleus tridentatus* plasma also contains proteins that bind to *S. aureus* and *E. coli* (Chiou *et al.*, 2000), and Tachyplesins which cause disruption of the plasma membranes of secretory granules causing release of cytolytic peptides (reviewed in Shuster 2003). *Tachypleus gigas* has lectins with haemagglutitinating activity which are specific for sialic acids (Tsuboi *et al.*, 1993a), as does *Tachypleus tridentatus* (Shimizu *et al.*, 1977; Tsuboi *et al.*, 1993c), *Limulus polyphemus* (Tsuboi *et al.*, 1993b), and *Carcinoscorpius rotundicauda* (Carcinoscorpin) (Bishayee and Dorai 1980; Dorai *et al.*, 1981; Mohan *et al.*, 1982; Srimal *et al.*, 1985). In *Limulus polyphemus*, Limulin is the most characterized sialic acid binding protein.

Limulin was first discovered by Marchalonis and Edelman (1968) and is a haemagglutinating *Limulus* CRP-like protein with sialic acid binding properties (Roche and Monsigny 1974; Robey and Liu 1981; Quigley *et al.*, 1984). Limulin has a native molecular weight of ~300kDa with subunit size ~33kDa (Armstrong *et al.*, 1996), and shows binding properties to PE-agarose and fetuin-sepharose through which it is isolated from the other *Limulus* pentraxins (Armstrong *et al.*, 1996). Limulin is less than 1% of the total pentraxin present in the haemolymph of *Limulus poly*phemus (Armstrong *et al.*, 1996) and shows calcium dependent haemolytic activity at concentrations as low as 3-5nM (Armstrong *et al.*, 1996) and agglutinating activity at ~2ng/ml (Roche *et al.*, 1975). Limulin haemolytic activity is inhibited by the presence of *Limulus* α_2-macroglobulin that has been reacted with proteases suggesting that α_2-macroglobulin modulates Limulin haemolytic function (Swarnakar *et al.*, 2000).

Limulus α_2-macroglobulin was discovered as a proteinase inhibitor that was released from the granules of the amoebocytes in *Limulus* haemolymph (Armstrong and Quigley 1985). It also contains the common features of mammalian α_2-macroglobulin: a bait region internal thiol-ester site, and a receptor-binding domain (Iwaki *et al.*, 1996) suggesting a similarity in binding and function between species. Peptides of *Limulus* α_2-macroglobulin align with sequences of human α_2-macroglobulin showing that the thiol-ester site had 67% identity (Sottrup-Jensen *et al.*, 1990). *Limulus* α_2-macroglobulin has the same function of mammalian α_2-macroglobulin as it binds proteases that could have been released from lysed cells or invading pathogens, and is cleared from the plasma apparently by *Limulus* amoebocytes (Melchior *et al.*, 1995). *Limulus* α_2-macroglobulin reacts with proteinases without forming a covalent link between itself and the proteinase but undergoes major conformational change (Enghild *et al.*, 1990).

Limulus α_2-macroglobulin was not the only mammalian protein homologue to be found in horseshoe crabs. Homologues of complement were found in *Tachypleus tridentatus*

(Ozaki *et al.*, 2005) and *Carcinoscorpius rotundicauda* (Zhu *et al.*, 2005). In *Carcinoscorpius*, CrC3 resembled human C3, and CrC2/bf resembled human C2 of complement. CrC3 binds a wide range of microbes and initiated phagocytosis by amoebocytes (Zhu *et al.*, 2005). Homologues of CrC3 and CrC2/bf with over 95% sequence similarity were found in *Tachypleus tridentatus* as TtC3 and TtC2/bf (Ozaki *et al.*, 2005). C2/bf and factor C (clotting cascade) from *Carcinoscorpius rotundicauda* are both serine proteases which interact with three core members of the pathogen-recognition receptors: carcinolectin 5, galactose-binding protein and CRP (Le Saux *et al.*, 2008).

As well as α_2-macroglobulin and complement, proteins were found in the horseshoe crabs *Limulus polyphemus, Tachypleus tridentatus* and *Carcinoscorpius rotundicauda* that are similar to the pentraxin family of proteins found in mammals, amphibians and fish.

1.5.3 Horseshoe crab pentraxins

Pentraxin-like proteins have been found in *Limulus polyphemus, Tachypleus tridentatus* and *Carcinoscorpius rotundicauda*. In *Limulus polyphemus* there have been found three sequences for CRP-like proteins, Limulin (a sialic acid binding CRP form at <1% of total pentraxin), and an SAP-like protein. In *Tachypleus tridentatus* there have been twenty-two CRP-like proteins found which could be separated into three groups based on binding properties and phylogenetic analysis (Iwaki *et al.*, 1999). In *Carcinoscorpius rotundicauda* there have been found eight CRP-like proteins (Ng *et al.*, 2004), and CrOctin which is similar to SAP in binding to PE, but interacts with CRP during infection (Li *et al.*, 2007).

Attempts at isolating haemolytic proteins from the plasma of *Tachypleus tridentatus* gave three CRP-like proteins: tCRP1 which was the most abundant CRP, tCRP2 and tCRP3 (Iwaki *et al.*, 1999). Amino acid sequencing of peptides formed from cyanogen bromide cleavage of tCRP1-3 allowed not only the primary structure of tCRP1-3 to be determined,

47

but also allowed generation of oligonucleotide primers which were used to identify clones of the tCRP1-3 genes. It was found that none of the tCRP genes contained introns and were mainly expressed in the haemocytes and hepatopancreas. The amino acid sequences deduced from the tCRP genes all had a N-linked glycosylation site at Asn123, and tCRP3 had another N-linked glycosylation site at Asn58. They also possessed an identical N-terminus sequence to the *Limulus* CRPs. The clone sequences were placed into three groups which corresponded to their homology to the tCRP1-3 protein sequences already determined. Each tCRP group showed individual properties as detailed below:

- tCRP1 was of native mass 300kDa and had subunits of sizes 29 and 31kDa of which the different sizes were not due to glycosylation. tCRP1 showed no haemolytic or sialic acid binding properties but did bind to PE calcium dependently, all of which are similar properties of *Limulus* CRPs. tCRP1 also showed lack of agglutinating and bacterial agglutinating activity.

- tCRP2 was of native mass 330kDa and had three subunits of sizes 29, 31, and 33kDa of which size difference was not due to glycosylation. tCRP2 showed calcium dependent haemolytic properties but much less than tCRP3. This was also the case for the calcium dependent agglutination as tCRP2 needed to be in higher concentration than tCRP3 to show the same agglutination, although agglutination by tCRP2 was inhibited by colominic acid from *E. coli* whilst tCRP3 was not. tCRP2 was also the only CRP to show bacterial agglutination, but only agglutinated *E. coli* with polysialic acids. Similar to tCRP1, tCRP2 precipitated PE-BSA but much less so, and was able to precipitate fetuin unlike tCRP1 which could not.

- tCRP3 was of native mass 340kDa of subunit sizes 31 and 33kDa of which size difference was not due to glycosylation, although there were differences in N-terminal sequences from the 33kDa subunit suggesting there are at least two forms of this subunit. tCRP3 showed the greatest haemolytic ability which was calcium dependent and abolished by desialylation similar to Limulin properties. It was suggested that haemolysis by tCRP3 was performed by it forming pores 2.4 to

3.5nm wide in the cell membrane. tCRP3 also showed the greatest agglutination activity and the greatest fetuin precipitation activity, but showed no precipitation with PE-BSA nor did it bind PE-sepharose. Fetuin precipitation by tCRP2 and tCRP3 was calcium dependent and inhibited by 2-keto-3-deoxyoctonic acid, N-acetylneuraminic acid and PE.

As tCRP1 and tCRP3 showed quite different properties, their structure was observed under electron microscopy to determine whether they were similar in shape. Both were hexagonal ring structures which possibly formed dimers.

CRP-like proteins were found in *Carcinoscorpius rotundicauda* by using LPS-sepharose in affinity chromatography of *Carcinoscorpius* haemolymph (Ng *et al.*, 2004). There were a few proteins in the haemolymph that bound LPS but the most dominant was a protein of subunit size ~25kDa which was shown to be similar to *Limulus* and *Tachypleus* CRPs. The CRP-LPS binding was calcium independent and was the first direct evidence that horseshoe crab CRP played a role in recognizing endotoxins. Based on peptides generated from mass spectrometry, oligonucleotide primers were designed and used to clone *Carcinoscorpius* CRP from a genomic library. There were eight different isoproteins of *Carcinoscorpius* CRP (cCRP) which fell into two groups: three cCRPs were homologous to tCRP1 therefore are cCRP1; and five cCRPs were homologous to tCRP2 therefore are cCRP2. Both groups had LPS binding activity but it was unknown whether all forms did. The hepatopancreas was shown to be the main site of CRP synthesis, and it appeared that the CRP gene expression was upregulated by 60 fold when the horseshoe crab was infected with *Pseudomonas aeruginosa*, correlating with a reduction in the number of bacteria in the haemolymph. This shows an effective antimicrobial defense system in *Carcinoscorpius rotundicauda* of which CRP may play a very important role.

Another part of the antimicrobial defense system in *Carcinoscorpius rotundicauda* may involve CrOctin: a protein from *Carcinoscorpius* haemolymph which bound to *S. enterica* and was eluted from it using PE (Li *et al.*, 2007). CrOctin has subunits of size 28kda and 29kDa and showed differences analysis to cCRP and other *Carcinoscorpius* proteins by mass spectrometry. The CrOctin name came from similarities to *Limulus* SAP because of its PE binding and subunit size. Therefore, the name CrOctin was after *Carcinoscorpius rotundicauda* (<u>Cr</u>Octin) and the octameric shape of *Limulus* SAP (Cr<u>O</u>ctin). CrOctin binds avidly to PE with much higher affinity than cCRP. However, cCRP binds to CrOctin immobilized by PE, and binds much more avidly from infected *Carcinoscorpius rotundicauda* than non-infected. This suggests that infection causes a conformational change in cCRP increasing its affinity for CrOctin immobilized on a PE covered surface. This may be a model for binding of CrOctin to PE on an LPS molecule. CrOctin peptides from mass spectrometry were used to probe for cDNA clones which showed there were two major groups of CrOctin containing eight possible sub-isoforms. CrOctin amino acid sequences were shown to have a sequence similar to the classical pentraxin motif: HxCxS(T)WxS, but phylogenetic analysis showed that it did not cluster with horseshoe crab CRP or SAP indicating it is a new family. The CrOctin gene was mainly expressed in the hepatopancreas and was consistently expressed in females during infection but not males. CrOctin was more highly expressed than CRP in females than in males suggesting that CrOctin levels were affected by hormone levels.

CRP-like proteins and SAP-like protein have been discovered in *Limulus polyphemus*. *Limulus* CRP has affinity for PC and PE but was first isolated using its affinity for PC (Robey and Liu 1981). The PC binding was calcium dependent and *Limulus* CRP was also shown to bind phosphate, PC-BSA and C-polysaccharide by precipitation reactions (Robey and Liu 1981). Two subunits of *Limulus* CRP of molecular weight 18kDa and 24kDa were identified and it was suggested that these two subunits come together to form a doubly stacked hexamer of molecular weight 500kDa (Robey and Liu 1981).

To more fully characterize *Limulus* CRP, two studies were carried out in tandem to identify the nucleotide and amino acid sequence. Unlike human CRP which has one gene, there are at least three genes for *Limulus* CRP (Nguyen *et al.*, 1986a). These genes were identified by screening a cDNA library with oligonucleotide probes determined from the amino acid sequence of peptide fragments of *Limulus* CRP (Nguyen *et al.*, 1986a). Only three clones (named 1.1, 1.4 and 3.3) gave strong hybridization signals to all the probes, and sequencing of the *Limulus* CRP gene in these three clones showed three genes encoding a signal peptide of 24 amino acids followed by 218 amino acids for the primary CRP structure. All the three genes shared the same protomer and had the promoter and 5' flanking region within 200 base pairs upstream of the initiation codon. The heat shock consensus sequence in human CRP is missing from the *Limulus* CRP genes, as is the intron separating the coding regions of the gene. From the gene sequencing, it was uncertain whether all three genes were expressed in the same cell, and there was also the possibility of more CRP sequences present. The gene sequencing also confirmed the presence of microheterogeneities in amino acid composition observed in peptides resulting from Edman degradation of *Limulus* CRP (Nguyen *et al.*, 1986b). The amino acid sequences for *Limulus* CRP 1.1, 1.4, and 3.3 were deduced from Edman degradation peptides which matched the sequences derived from the cDNA clones (Nguyen *et al.*, 1986b). Therefore, *Limulus* CRP was not a single protein, but a family of proteins of which there were at least three different subunits that were expressed.

The three subunits (CRP 1.1, 1.4 and 3.3) shared an identical amino acid sequence from 1-44 and 206-218 with microheterogeneity between residue 44 and 206 at 10-11% (Nguyen *et al.*, 1986b). All three sequences show six cysteines which forms three disulphide bonds: Cys38-Cys101, Cys88-Cys120, Cys183-Cys217, of which Cys38-Cys101 is conserved amongst pentraxins. *Limulus* CRP 1.1 has a seventh cysteine at residue 45 which CRP 1.4 and 3.3 do not possess (Nguyen *et al.*, 1986a). All three CRPs also have a glycosylation site at Asn123-Ala-Thr (Nguyen *et al.*, 1986b). It was suggested

51

that all three subunits existed in equimolar amounts and aggregated to form a hexamer of

two of each subunit (Nguyen *et al.*, 1986b).

Further analysis showed that *Limulus* CRP was indeed composed of at least three

subunits as indicated by SDS-PAGE (Tennent *et al.*, 1993). Mass spectrometry indicated

the presence of many subunit sizes all approximately 25kDa, some of which varied in

glycosylation (Tennent *et al.*, 1993). The mature protein weight of *Limulus* CRP in the

study was determined by native-PAGE to be approximately 300kDa which may be an

aggregate of 12 subunits of CRP (each at 25kDa in size) (Tennent *et al.*, 1993). Unlike the

previous study which used pooled haemolymph (Nguyen *et al.*, 1986b) it was suggested

that there may be differential expression of CRP genes and/or glycosylation within and

between individual horseshoe crabs as there seems to be a predominance of one CRP

species in individual crab haemolymph. Later studies identified Limulin as a sialic acid

binding form of *Limulus* CRP which also binds PE the same as *Limulus* CRP but is <1% of

total pentraxin (Quigley *et al.*, 1994; Armstrong *et al.*, 1996).

Recent studies on *Limulus* CRP showed that it had a permeabilising effect on liposomes

in the presence of calcium by intercalating into the lipid bilayer to form pores 1.9±0.2nm

wide which allow movement of ions across the lipid bilayer (Harrington *et al.*, 2008).

Limulin and *Limulus* SAP also showed this effect on liposome permeability. Liposomes of

lipids of some natural origin were sensitive to permeabilisation, whilst liposomes from

PC:PE, phosphate, polysaccharide, sphingomyelin, and cholesterol were insensitive to

permeabilisation. *Limulus* CRP was also shown to agglutinate liposomes containing *E.

coli* lipids both calcium independently and dependently, and forms fibres that encapsulate

the liposomes within sheet-like structures. The agglutination is optimal at acidic pH, but is

observed at pH 7.8 unlike permeabilisation which was optimal at pH 5.8, but unobservable

at pH 7.8.

As *Limulus* CRP bound to PE (Quigley *et al.*, 1994), PE affinity chromatography was used to elute it from *Limulus* plasma (Shrive *et al.*, 1999). *Limulus* CRP was confirmed to show three bands on an SDS-PAGE gel and could be eluted from a PE column using 10mM PC (Shrive *et al.*, 1999). It was during isolation of *Limulus* CRP from a PE column that *Limulus* SAP was discovered. *Limulus* plasma had been applied to a PE column, with Limulin eluted using sialic acid and CRP eluted using PC (Shrive *et al.*, 1999). The remaining protein on the column was eluted using EDTA which showed the presence of three bands at 25-30kDa in size, and three bands at 70kDa in size on SDS-PAGE (Shrive *et al.*, 1999). This new protein was termed *Limulus* SAP due to its PE affinity and lack of PC affinity. The multiple bands present on SDS-PAGE may have indicated different glycosylation and/or sequences.

Limulus SAP has approximately 0.12-0.17mg/ml concentration in plasma so is only 8-19% of total *Limulus* pentraxin. Oligonucleotide probes were designed from an electron density map of the *Limulus* SAP structure and peptides generated from cyanogen bromide cleavage (Tharia *et al.*, 2002). The nucleotide sequence of *Limulus* SAP was determined from screening of a cDNA library using these probes. The *Limulus* SAP gene encodes a 234 amino acid sequence of which the first 17 amino acids are part of a signal peptide that has extensive hydrophobicity and a net positive charge. The gene had no introns similar to the *Limulus* CRP genes, and there only appeared to be one single sequence and not multiple sequences present as for *Limulus* CRP. However, peptides generated by cyanogen bromide cleavage did not entirely match the sequence derived from the nucleotide sequence which may suggest the presence of SAP variants. The amino acid sequence determined from the *Limulus* SAP gene, gave a subunit mass of 23.8kDa and showed two potential glycosylation sites at Asn81 and Asn118, but the electron density map showed only glycosylation at Asn81. The amino acid sequence showed 32% identity with *Limulus* CRPs, 31-35% identity with *Tachypleus* CRPs and 22% identity to human CRP and SAP.

The structure of *Limulus* SAP had been determined before the nucleotide and amino acid sequence were found using diffraction data (Myles *et al.,* 1990) and a pentraxin polyalanine protomer model from human CRP (Shrive *et al.,* 1999). The 3Å structure of *Limulus* SAP was the first invertebrate pentraxin structure to be determined and showed that *Limulus* SAP was a doubly stacked cyclic octameric ring (see figure 1.7) not a pentamer like human CRP and SAP. The protomers of this hexadecamer were identical, and each showed the core β-fold and pentraxin helix (see figure 1.8) as shown in human SAP and human CRP. Each protomer also showed the conserved pentraxin disulphide bridge (Cys38-Cys101) and two additional disulphide bridges as seen in *Limulus* CRPs.

Figure 1.7 *Limulus* SAP view down the 8 fold axis (Source: A. K Shrive).

The calcium binding site on *Limulus* SAP was similar in topology to that of the mammalian pentraxins, but was situated on the external edge of the octamer rather than on one face. The pentraxin helix was also in a different position in *Limulus* SAP than in mammalian pentraxins as it was on the internal edge of the octamer rather than on one face. Interactions between protomers were non-covalent and a pair of protomers formed a

continuous twisted β-sheet structure across the hexadecamer. There was also no conservation of salt bridges between *Limulus* SAP and other pentraxins suggesting different relationships between protomers. The putative C1q cleft in human CRP is not present in either human SAP or *Limulus* SAP, suggesting similarity in function between human SAP and *Limulus* SAP, rather than with *Limulus* SAP and human CRP.

Figure 1.8 *Limulus* SAP protomer showing labelling of the secondary structures (Source: Shrive *et al.,* 1999). α-helices are shown in red, β-sheets are shown in blue, loop regions are shown in yellow, and disulphide bridges are shown in orange .

The presence of CRP and SAP in horseshoe crabs and humans suggests that they were present in an ancestor of chordates and arthropods, but there is a possibility that they evolved along the separate lines by gene duplication. The presence of both CRP and SAP in humans and horseshoe crabs also indicates that they are performing separate but essential functions (Shrive *et al.,* 1999).

1.6 Research aims

The overall aim of this research was to gain more insight into the pentraxins CRP and SAP which appear throughout evolution from invertebrates to mammals.

Three CRP proteins have been identified in *Limulus polyphemus* (Nguyen *et al.,* 1986a), and there is a possibility that these proteins have different binding properties such as those in *Tachypleus tridentatus* (Iwaki *et al.,* 1999), and that there may even be more than three

forms. There is also the possibility that there is more than one form of *Limulus* SAP, because of multiple bands identified on SDS-PAGE (Shrive *et al.*, 1999). One aim of this research was to isolate CRP and SAP proteins from *Limulus polyphemus* plasma using human CRP ligands, in an attempt to identify CRP and SAP proteins with different binding properties. It was also attempted to obtain the crystal structure of *Limulus* CRP with and without ligand bound, and also *Limulus* SAP ligand bound. Determining the structure of *Limulus* CRP, and discovery of *Limulus* CRP and SAP binding properties, would allow evolutionary and functional comparisons with human CRP and SAP.

A second aim of this research was to phylogenetically compare CRP and SAP sequences from horseshoe crabs with those from other invertebrates, fish, amphibians, and mammals, and also CRP and SAP sequences within the horseshoe crabs themselves. The key residues corresponding to calcium ion binding, ligand binding, and putative C1q binding sites on human CRP were compared to corresponding residues in the horseshoe crab CRPs and SAP to identify areas of possible structure/function conservation. These studies would give more insight into possible differences in binding properties of CRP and SAP in horseshoe crabs, both within and between species.

Attempts were made to isolate CRP and SAP genes from the smooth dogfish, *Mustelus canis*, so that they could be sequenced providing additional CRP and SAP amino acid sequences for phylogenetic analysis, and to fit onto the crystal structure of dogfish CRP and SAP (unpublished data). Obtaining the sequence and structure of dogfish CRP and SAP would also allow comparisons between CRP and SAP from invertebrates to elasmobranchs to mammals, and indicate evolutionary change or conservation in areas of proposed function.

Chapter 2. Materials and Methods

2.1 Introduction to isolation of horseshoe crab and dogfish pentraxins

The CRP and SAP like proteins were isolated from the plasma of *Limulus polyphemus* and *Mustelus canis* by their affinity to the ligands phosphoethanolamine (PE) and phosphocholine (PC) (Robey and Liu 1981; Robey *et al.*, 1983; Shrive *et al.*, 1999). Both the CRP and SAP bind to PE, but only CRP binds to PC. *Limulus* and *M. canis* plasma was added directly to a PE column which bound both CRP and SAP, which were then eluted separately using their different ligand binding properties. Once separation was achieved, the CRP and SAP were dialysed to remove the elution ligands, and then concentrated to be used in gel electrophoresis showing the sizes of the eluted proteins. Purified *Limulus* protein solutions were then either used in crystal trials, reapplied to a PE column and eluted again with a variety of ligands, or applied to a hydrophobicity interaction column (HIC) (see figure 2.1). *M. canis* CRP was used in cyanogen bromide digestion and trypsin digestion as detailed in section 5.2.
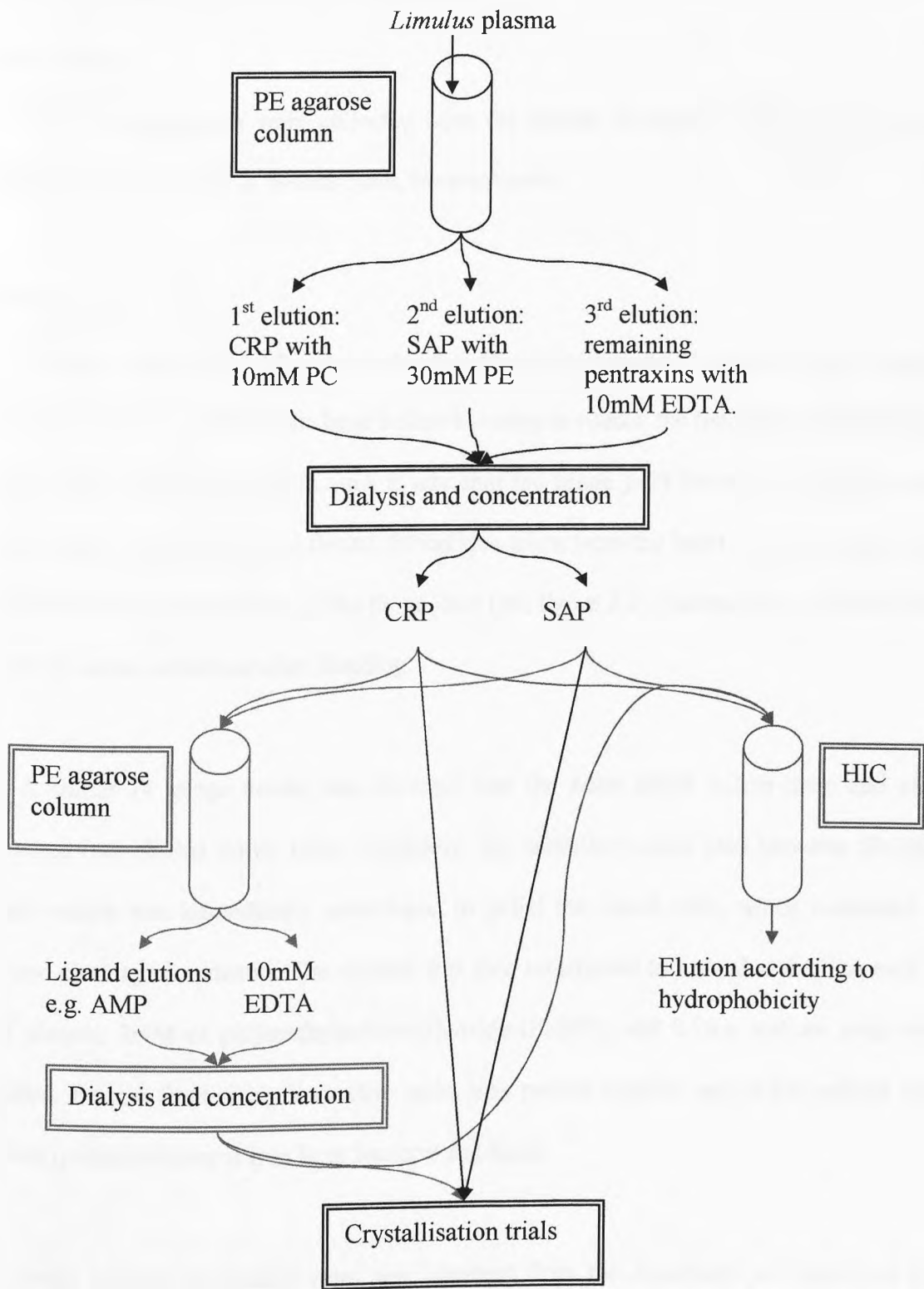
Figure 2.1. Diagram showing the procedure of isolation of *Limulus* pentraxins from plasma and subsequent processing. First the pentraxins CRP and SAP were separated from each other using affinity chromatography from a PE-agarose column using their affinity for PC and PE. Pentraxins were then dialysed to remove elution ligands and then re-applied to a PE-agarose column, applied to a HIC column, or used in crystallisation trials. CRP and SAP re-applied to a PE-agarose column were eluted using the desired ligand and EDTA, and then dialysed and used in either crystallisation trials or applied to a HIC. CRP and SAP that were applied to a HIC were eluted according to their surface hydrophobicity using an elution buffer.

2.1.1 Obtaining horseshoe crab plasma

Materials:

Limulus polyphemus were collected from the Marine Resources Centre at the Marine Biological Laboratory in Woods Hole, Massachusetts.

Procedure:

Animals were checked for previous signs of cardiac puncture to prevent over bleeding, and then placed at 4°C for one hour before bleeding to reduce the risk of the blood clotting. Horseshoe crabs were held in such a way that the hinge joint between the prosoma and opisthosoma was visible and flexed. Blood was taken from the heart of the horseshoe crab which lies along the midline of the hinge joint (see figure 2.2). Animals were released back into the ocean unharmed after bleeding.

A sterile 14 gauge needle was inserted into the heart about 1-2cm deep and blood drained into chilled sterile tubes. Typically, the horseshoe crabs bled between 50-100ml each which was immediately centrifuged to pellet the blood cells, which contained the blood clotting components. The plasma was then transferred to a sterile tube. For each ml of plasma, 1mM of polymethylsulfonylfluoride (PMSF), and 0.2mg sodium azide were added. Plasma from many horseshoe crabs was pooled together and 0.2% sodium azide (NaN_3) added to keep it free from bacteria and fungi.

Large volumes of plasma were also obtained from the Associates of Cape Cod Inc., Falmouth, Massachusetts. Due to the large volumes, plasma was concentrated using polyethylene glycol (PEG) cut procedure. This involved adding PMSF to concentration of 1mM, and PEG 8000 molecular weight to concentration of 3%, to approximately 1.5litre of plasma. The preparation was incubated at 4°C for 3 to 6 hours, and then centrifuged for 30 minutes at 20000rpm to pellet the PEG and haemocyanin. PMSF was added to the supernatant to a final concentration of 1mM, and PEG 8000 added to bring the final

concentration to 10% PEG. The supernatant was incubated for approximately 1 hour at 4°C and centrifuged for 10 minutes at 20000rpm. The supernatant from the second centrifugation was discarded and the pellet containing the pentraxins was dissolved in calcium free buffer (0.15M NaCl, 50mM Tris, 1mM EDTA, 0.2mg/ml NaN_3 pH 7.5) with 1mM PMSF and left overnight at 4°C. Calcium chloride ($CaCl_2$) was then added to make the final concentration 5mM.
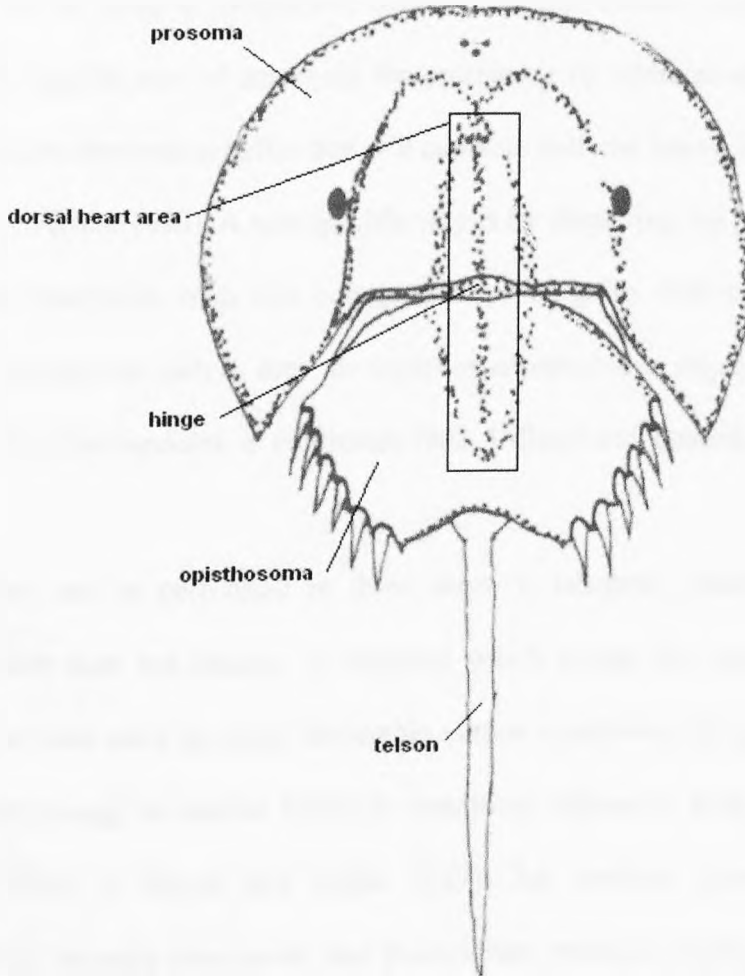
Figure 2.2. Diagram to show where blood was taken from *Limulus polyphemus*. The animal was flexed to expose the hinge joint and the heart punctured through this. (Adapted from www.dnr.state.mo.us.com).

2.1.2 Obtaining dogfish plasma

Dogfish plasma was obtained as detailed in section 5.2.

2.1.3 Introduction to affinity chromatography

Affinity chromatography was first introduced in 1968 (Cuatrecases *et al.*, 1968) and is based upon recognition of a macromolecule such as protein, for a natural or artificial molecule that can be immobilized on a matrix (Wilchek and Chaiken, in Bailon *et al.*, 2000). Once the protein is selectively bound to the column and contaminants are washed through, the protein can be desorbed by either disturbing the protein-ligand interaction on the column or using a competitive ligand in elution buffer (Catsimpoolas, in Heftmann 1983). A specific way of desorbing the proteins is by addition of analogues of ligand or new ligand to the elution buffer that will compete with the ligand binding site (Villems and Toomik, in Kline 1993). A non-specific way is by disturbing the protein-ligand interaction by using detergents, high salt concentration, change in elution buffer pH, addition of protein denaturants such as urea, or depletion of metal ions responsible for the nativity of the protein (Catsimpoolas, in Heftmann 1983; Villems and Toomik in Kline 1993).

Elutions can be performed in three ways: i) isocratic which means elution buffer composition does not change, ii) stepwise which means the elution buffer composition changes at least once for more favourable elution conditions, iii) gradient elution which is a constant change in elution buffer to conditions favouring protein desorbtion from the column (Roe, in Harris and Angal 1989). An isocratic elution allows a dynamic equilibrium between association and dissociation which is dependent on the equilibrium constant for the protein-ligand interaction (Wilchek and Chaiken, in Bailon *et al.*, 2000). Therefore, affinity is reflected in the elution volume: the larger the elution volume, the lower the affinity.

Isolation of CRP and SAP pentraxins from *Limulus* plasma was achieved by affinity chromatography which utilises their specific binding to PC and PE. Like human CRP that binds to both PC (Volanakis and Kaplan 1971) and PE (Schwalbe *et al.*, 1992), *Limulus*

CRP has affinity for PC (Robey and Liu 1981) and PE (Quigley *et al.*, 1994). *Limulus* SAP was also shown to bind PE (Shrive *et al.*, 1999; Tharia *et al.*, 2002). Therefore, CRP and SAP could be isolated from *Limulus* plasma using their affinity to PE, then eluted sequentially using PC to elute CRP, and PE to elute SAP. Affinity chromatography of the pentraxins requires the addition of a co-factor: calcium, that allows binding of CRP and SAP to PC/PE and PE respectively (see figure 2.3).Attempts at isolation of subgroups of *Limulus* CRP and SAP from fractions of pure protein was performed using affinity chromatography with a variety of ligands.

Figure 2.3. Diagram showing basis of affinity chromatography for the pentraxins of *Limulus polyphemus* and *Mustelus canis*. Part 1: i-PE linked to agarose beads (solid phase). ii- calcium ions iii- pentraxin. Part 2 and 3: Pentraxin requires calcium ions to attach to PE ligand. Part 4: iv- ligand for pentraxin e.g. PC or PE. Part 5: ligand binds pentraxin through calcium ions therefore removing it from the solid phase.

Isolation of CRP and SAP pentraxins from *Mustelus canis* plasma was also achieved by affinity chromatography using PC and PE, as *M. canis* CRP and SAP have specific binding to PC and PE respectively (Robey *et al.*, 1983).

For *Limulus* and *Mustelus* pentraxins, the ligand used on the column was PE which shows group selectivity of binding both to CRP and SAP. To isolate *Limulus* and *Mustelus* pentraxins from their plasma, low pressure column chromatography is used with PE linked beaded agarose. This is packed into a 25ml column over which buffer is pumped at a low flow rate of 0.5ml/min that prevents the agarose compacting which could reduce the flow rate and also reduce the available area for the pentraxins to bind with PE. Agarose has high

stability and an inert surface (Roe, in Harris and Angal 1989) so that the pentraxins do not interact with it in preference to PE. Specific elution of CRP and SAP with ligand was chosen rather than non-specific elution that may have denatured the proteins and rendered them useless for structure analysis. The possibility of other PE binding proteins being present cannot be ruled out.

The *Limulus* pentraxins were eluted from the PE-agarose column using either an isocratic elution of a single ligand buffer at a given molarity (for example 10mM PC) or a gradient elution of ligand, for example 0-20mM PC. *Mustelus* pentraxins were eluted using isocratic elutions of PC and PE. Once the *Limulus* CRP and SAP were isolated from the horseshoe crab plasma, they were either reapplied to an affinity chromatography column, used in crystal trials, or used in fast protein liquid chromatography (FPLC) a form of high pressure liquid chromatography (HPLC) specific for protein separation using a hydrophobicity interaction column (HIC).

Materials:

Equipment:

- 0.2µm pore acetate filter (Sartorius).
- Biologic Low Pressure System and HP controller (BIORAD)
- *O*-Phosphoethanolamine agarose beads (Sigma).

Chemicals:

- Calcium chloride dihydrate (Sigma and VWR).
- Sodium chloride (VWR)
- Tris(hydroxymethyl)methylamine (VWR)
- Phosphocholine chloride calcium salt tetrahydrate (Sigma)
- *O*-Phosphoethanolamine (Sigma)
- Ethylenediaminetetraacetic acid (EDTA) (Sigma)

  - Sodium azide (Sigma)

Procedure:

Isolation of pentraxins CRP and SAP from *Limulus* plasma was performed using a BIORAD Biologic LP system run by a BIORAD Biologic HR controller, and an affinity chromatography column composed of *O*-Phosphoethanolamine attached via the amine group to epoxy activated beaded agarose. The column was equilibrated with calcium wash buffer (see table 2.1) at a flow rate of 0.5ml/min to remove any previous buffer. Calcium wash buffer was used because the pentraxins require calcium to bind to the PE column and also the elution ligands.

Table 2.1 Table showing the buffers used for affinity chromatography of *Limulus* and *Mustelus* pentraxins. All buffers were at pH 7.4, and degassed to remove air bubbles which may block and dry out the column, and filtered with 0.2µm pore cellulose acetate filter to remove bacterial and fungal spores.

Buffer	Components	pH
Calcium wash buffer	50mM Tris, 10mM $CaCl_2$, 150mM NaCl	7.4
Phosphocholine (PC) buffer	10mM Phosphocholine chloride, 50mM Tris, 10mM $CaCl_2$, 150mM NaCl	7.4
Phosphoethanolamine (PE) buffer	30mM O-Phosphoethanolamine, 50mM Tris, 10mM $CaCl_2$, 150mM NaCl	7.4
EDTA buffer	50mM Tris, 150mM NaCl, 10mM EDTA	7.4
Regeneration buffer A	500mM NaCl, 200mM Tris, 10mM EDTA	7.4
Regeneration buffer B	500mM NaCl, 50mM Tris, 10mM $CaCl_2$	7.4
Regeneration buffer C	50mM Tris, 10mM $CaCl_2$, 150mM NaCl, 0.02% NaN_3	7.4

Limulus plasma was filtered through a 0.2µm cellulose acetate filter to remove any contaminants such as bacterial and fungal spores. 5-10mls of plasma was diluted two fold with calcium wash buffer to reduce the salt concentration, which would otherwise have affected binding of the pentraxins to the column. The diluted and filtered plasma was then added to the column and washed through with calcium wash buffer to remove any unbound proteins. *Limulus* CRP was then eluted from the column using 10mM PC in calcium wash buffer (see table 2.1), and *Limulus* SAP was eluted from the column using 30mM PE in calcium wash buffer (see table 2.1). The PC and PE act as competitive ligands with the PE attached to the agarose on the column, and eluted most of their corresponding proteins from the column. However, some proteins do remain and they were removed from the column with ethylene-diamine-tetra-acetic acid (EDTA) buffer (see table 2.1) which acts

by chelating the calcium ions that allow the pentraxins to attach to the PE on the column. A similar procedure was used with *M. canis* plasma to isolate CRP and SAP (see section 5.2). After EDTA buffer, the column was subjected to regeneration with regeneration buffers A, B, and C (see table 2.1). These contained high salt concentrations and EDTA to elute the remaining proteins, and sodium azide to protect the column from fungal and bacterial growth during storage.

The above basic procedure is that used for elution of the CRP and SAP pentraxins from the plasma of the horseshoe crab and dogfish. To analyse binding of either *Limulus* CRP or SAP, they were individually reapplied to the PE column and eluted using buffers that contained the desired ligand to test as described in sections 3.2 and 4.2. Such an example would be 10mM ribose-5-phosphate in calcium wash buffer which was: 10mM ribose-5-phosphate, 50mM Tris, 10mM $CaCl_2$, and 150mM NaCl.

2.1.4 Introduction to Fast Protein Liquid Chromatography (FPLC)

To identify whether *Limulus* CRP or SAP isolated using affinity chromatography contained multiple molecular aggregate forms, FPLC was performed using surface hydrophobicity as the variant between molecular aggregate forms. Proteins can have patches of hydrophobicity on their surface due to the side-chains of non-polar amino acids, and also due to protein modifications such as glycosylation. In water, the protein is surrounded by a film of ordered water molecules and the hydrophobic patches of the protein are not exposed (Harris, in Harris and Angal 1989). For proteins to bind to a hydrophobic interaction column (HIC), the hydrophobic patches must be exposed. This process is termed "salting out", and involves adding a high molarity salt which solvates the water molecules surrounding the protein causing disorder in the water film and exposing areas of hydrophobicity (Harris, in Harris and Angal 1989).

Once proteins are bound to a HIC, they can be removed in several ways: i) reducing the ionic strength, ii) increasing the pH, iii) reducing the temperature, and iv) displacement with component that has a stronger attraction for HIC or makes the protein more hydrophilic. The number and size of the hydrophobic patches on the surface of the proteins determines their avidity of binding to the HIC, and therefore determines the rate of desorbtion from the column. The larger the number or size of the hydrophobic patches, the higher the binding avidity to the HIC.

A small change is sequence between either *Limulus* CRP or SAP molecular aggregate forms, or modifications such as glycosylation, could alter the surface hydrophobicity, and therefore change the interaction with the HIC matrix. *Limulus* pentraxins were absorbed onto a phenyl hydrophobicity interaction column and eluted using a decrease in ionic strength using elution buffer.

Materials:

Equipment:

- 0.2µm pore acetate filter (Sartorius)

- GE Healthcare ÄKTA Explorer FPLC system (GE Healthcare)

- 1ml HiTrap Phenyl HP Hydrophobicity Interaction Column (GE Healthcare)

Chemicals:
- Ammonium sulphate (Sigma)

- Calcium chloride dihydrate (Sigma and VWR).

- Sodium chloride (VWR)

- Tris(hydroxymethyl)methylamine (VWR)

Procedure:

The FPLC was used to analyse *Limulus* CRP and SAP components. The GE Healthcare ÄKTA Explorer FPLC system was used with a HiTrap Phenyl HP Hydrophobicity Interaction Column (HIC) (1ml). Binding buffer (see table 2.2) was used to equilibrate the column, and 350µg of protein made up to 1ml with binding buffer was loaded onto the column. Binding buffer contained a high concentration of ammonium sulphate to "salt out" the protein. Protein was washed onto the column using binding buffer, then eluted using 0-100% gradient of elution buffer (see table 2.2). Elution buffer contained no ammonium sulphate therefore changing the ionic strength and decreasing the hydrophobicity of the protein surface, thereby reducing protein-column hydrophobic interactions. The column was then washed through with water to remove excess elution buffer that would affect protein binding in subsequent use of the column.

Table 2.2. Table showing the buffers used for FPLC of *L. polyphemus* pentraxins. Binding buffer was degassed but not filtered before use. Elution buffer was filtered and de-gassed before use.

Buffer for FPLC	Components	pH
Binding buffer	20mM Tris, 10mM $CaCl_2$, 3.5M $(NH_4)_2SO_4$	8
Elution buffer	20mM Tris, 10mM $CaCl_2$, 100mM NaCl	8

2.1.5 Removal of elution buffer from pentraxins.

Materials:

- Dialysis tubing with molecular weight cut off of 12-14kDa (Medicell)

- Ethylenediaminetetraacetic acid (EDTA) (Sigma)

- Calcium chloride dihydrate (Sigma and VWR).

- Sodium chloride (VWR)

- Tris(hydroxymethyl)methylamine (VWR)

Procedure:

If the *Limulus* CRP or SAP was to be re-applied to the PE column, the ligand with which they were eluted was removed as otherwise they would not re-bind as their binding sites were already occupied. Ligands also needed to be removed from the proteins for crystallisation trials and FPLC analysis. As the pentraxins bind via calcium ions, removal of calcium ions from the solution of protein detached the ligand. However, the pentraxins also require calcium to keep a stable conformation and bind new ligands, and so calcium ions were re-added to the protein solution. This was performed using dialysis where the protein solution in dialysis tubing was immersed in EDTA buffer (see table 2.3) in 100 fold excess for 24 hours with one change of buffer. This was sufficient enough to allow the EDTA into the dialysis tubing and chelate the calcium ions holding the pentraxin and ligand together. The unbound ligands migrated out of the dialysis tubing due to a gradient, as did the calcium ions bound to EDTA. Re-addition of calcium ions to the protein was done via dialysis against calcium wash buffer (see table 2.3) in the same manner as with EDTA buffer. This allowed calcium ions to migrate into the dialysis tubing along a gradient where their concentration was lower and interact with the proteins.

Table 2.3. Table showing the buffers used for dialysis of proteins eluted using affinity chromatography. EDTA dialysis buffer was used to remove ligand from the proteins, and calcium wash buffer was used to replenish lost calcium. Dialysis was performed for 24 hours with one change of buffer.

Dialysis buffer	Components	pH
Calcium wash buffer	50mM Tris, 10mM $CaCl_2$, 150mM NaCl	7.4
EDTA buffer	50mM Tris, 10mM EDTA, 150mM NaCl	7.4

2.1.6 Concentrating pentraxins

Materials:

- Amicon Ultra 15ml and 4ml centrifugal filters molecular weight cut off 10kDa (Millipore).

Online programs:

- http://www.expasy.ch/tools/protparam.html

Procedure:

Once the CRP and SAP were dialysed to remove the elution ligand, they were concentrated in order to give an appropriate concentration for visualisation of protein components on a gel, or use of the protein in crystallisation trials. Concentration was achieved by centrifugation of the protein solution through a membrane with a molecular weight cut off of 10kDa. This allows the buffer solution to pass through the membrane, but prevents the proteins from doing so. The concentrators used were mostly Amicon Ultra 15ml and 4ml centrifugal filter devices at speed of 2424 rpm.

Once concentrated, the protein concentration was measured using UV absorbance spectrophotometry which measured the absorbance unit for full scale deflection (AUFS) of a solution at 280nm and 320nm. Wavelength 280nm measures the concentration of protein in solution according to the light absorbance of the aromatic rings of the amino acids tyrosine and tryptophan, and disulphide bonds. Wavelength 320nm measures the background absorbance, as proteins do not absorb light at this wavelength. AUF_{320nm} was subtracted from the AUF_{280nm}, then divided by the extinction coefficient of the *Limulus* CRP which is approximately 1.5 (Abs 0.1%) or *Limulus* SAP which is approximately 1.4 (Abs 0.1%), to give the concentration of protein in solution. The extinction coefficient was worked out by entering the existing sequences of either *Limulus* CRP 1.1, 1.4, and 3.3., or *Limulus* SAP into an online "ProtParam" program from Expasy (http://www.expasy.ch/tools/protparam.html). This program works out the extinction coefficient of the native protein in solution by calculating the total absorption at 280nm from tyrosine, tryptophan and cysteine disulphide bonds as determined from the amino acid sequence (Gasteiger *et al.*, 2003).

2.1.7 Introduction to gel electrophoresis

Gel electrophoresis involves the separation of components through a gel under an electric field according to their electrophoretic mobility. Two types of gels can be used for electrophoresis: agarose and acrylamide. Agarose is a linear polysaccharide of alternating galactose and 3,6-anhydrogalactose with hydrogen bond cross linking, and is mainly used for DNA or RNA electrophoresis as at low percentages its pores are large relative to proteins (Gaal *et al.*, 1980; Andrews 1986). The pores in acrylamide gels can be made much smaller and are therefore more often used for protein separation. Acrylamide is polymerised to polyacrylamide in the presence of N,N'-methylenebisacrylamide (bis–acrylamide) which is used as a cross linking agent (Shi and Jackowski, in Hames 1998). Unlike agarose which sets into a gel when it cools, acrylamide needs to be polymerised. Ammonium persulphate is decomposed by N, N, N', N'-tetramethylene diamine (TEMED) to give free radicals which promote polymerisation of acrylamide monomers with each other and bis-acrylamide (Shi and Jackowski, in Hames 1998). Acrylamide composition is defined by the total percentage concentration of both acryalamide and bisacryalamide (T%), and the percentage of cross-linker bisacrylamide relative to the total (C%) (Gaal *et al.*, 1980). The resolving power of the gel depends partly on both T% and C% as well as buffers and pH (Shi and Jackowski, in Hames 1998). With both agarose and acrylamide, the pore size can be varied as larger pores allow large and small molecules to travel through them easily and quickly, whilst smaller pores allow only the smaller molecules to travel through them easily and quickly.

Proteins can be isolated from each other using gel electrophoresis due to charge, size, or both. To isolate proteins due to size, all the protein components must have the same charge and so sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) is used. SDS is an anionic detergent that coats proteins so that they all have the same charge (Shi and Jackowski, in Hames 1998). It also acts as a protein denaturant along with β-

mercaptoethanol which reduces disulphide bridges that hold the tertiary structure of the protein together (Shi and Jackowski, in Hames 1998). Therefore, proteins are both denatured, and negatively charged so that they run through the gel according to their size. SDS-PAGE can involve both a resolving gel in which the proteins separate due to size, and a stacking gel in which the proteins are concentrated into a band before they enter the resolving gel to give greater resolution (Shi and Jackowski, in Hames 1998). Proteins are mixed with sample buffer that contains a dye to visualize the electrophoresis front, and sucrose or glycerol which helps the sample to sink to the bottom of the well. To separate proteins according to both their size and overall charge, native PAGE can be used which works in the same way as SDS-PAGE only without SDS, meaning that protein charge affects protein migration through the gel as well as the protein size.

Both SDS-PAGE and native gels can be formed as gradient gels where the percentage of acrylamide and therefore pore size, varies from low percentage (i.e. larger pore size) to higher percentage (i.e. smaller pore size). This allows a greater size range of proteins to be separated as larger proteins separate before smaller proteins separate (Shi and Jackowski, in Hames 1998). It also allows proteins of similar size to be more highly resolved than in a non-gradient gel as the proteins reach a pore size limit where they cannot move any further and therefore stack up in a band (Andrews 1986).

Materials for gels:

- Sodium dodecyl sulphate (Sigma)
- N, N, N', N'-tetramethylene diamine (TEMED) (Sigma)
- Ammonium persulphate (VWR)
- Tris(hydroxymethyl)methylamine (VWR)
- Acrylamide and bisacrylamide (Sigma).

Loading buffer:

- Glycine (VWR and Sigma)

- Glycerol (Sigma)

- β-mercaptoethanol (Sigma)

- Bromophenol blue (VWR)

- Molecular weight markers (Sigma)

Procedure:

Protein subunit composition of the PC and PE elutions of the *Limulus* plasma was analysed using SDS-PAGE in either a reduced state or a non-reduced state. A typical SDS gel was composed of a 12.5% acrylamide resolving gel, and a 4% acrylamide stacking gel (see table 2.4). Gradient SDS gels used were made of 5%-12.5% acrylamide. Before the protein elution samples were loaded onto the gels they were mixed with the sample buffer (see table 2.4) in a one to one ratio. If the samples were to be reduced, β-mercaptoethanol was added to the sample buffer at 5% volume. Samples for SDS-PAGE were heated at 95-100°C for approximately 2 minutes to further denature proteins just prior to addition onto the gel. Gels were placed in an electrophoresis tank in their respective buffers (see table 2.4), and a voltage of 200V run through them for approximately 40-50 minutes.

Materials for staining:

- Brilliant blue R250 (Sigma)

- Methanol (VWR)

- Ethanol (VWR)

- Acetic acid (VWR)

- Silver stain kit (Biorad)

Table 2.4. Table showing the components of SDS-PAGE gels. The volumes shown were enough to make two 0.75mm thick gels.

Gel electrophoresis. SDS-PAGE	Components
4% stacking gel	0.67ml acrylamide/bisacrylamide (30%T 2.6%C) 3ml deionised water 1.25ml 0.5M Tris HCl pH 6.8 50µl 10% SDS 5µl TEMED 50µl 10% Ammonium persulphate
12.5% resolving gel	3.33ml acrylamide/bisacrylamide (30%T 2.6%C) 2.55ml deionised water 2.00ml 1.5M Tris HCl pH 8.8 80µl 10% SDS 5µL TEMED 50µl 10% Ammonium persulphate
Running buffer	1.515g Tris 7.21g Glycine 0.5g SDS
Sample buffer	0.125M Tris-HCl pH 6.8 4% SDS 20% Glycerol 0.02% bromophenol blue (if reducing, add 5% volume β-mercaptoethanol)

Procedure:

Two different types of staining technique were used: Coomassie Brilliant Blue and Silver Stain. Coomassie Brilliant Blue requires between 5 and 10µg of protein in each lane of the gel in order to work. Silver stain is 10-50 times more sensitive than Coomassie Brilliant Blue Stain.

To stain with Coomassie Brilliant Blue, the gel was submerged in Coomassie Blue Stain (see table 2.5) for a minimum of one hour on a shaker. The proteins were stained by the Coomassie Blue in the solution and were fixed into the gel by the methanol which tightened the gel pore sizes so that proteins cannot leach out. To de-stain, the Coomassie Blue Stain was drained off, and the gel immersed in Coomassie Blue De-stain solution (see table 2.5) which was changed several times over the period of 2-3 hours until the background was clear and the protein bands visible.

Table 2.5. Table showing the components of the stains used on both native-PAGE and SDS-PAGE gels. The silver stain was supplied in kit form and stock solutions provided.

Coomassie Brilliant Blue stain	Components
Coomassie Brilliant Blue Stain	45% methanol, 10% acetic acid, 0.25% Brillaint Blue R250
Coomassie Blue De-Stain	30% ethanol, 10% acetic acid
Silver Stain	
Fixative	40% methanol, 10% acetic acid
Oxidiser	10 fold stock diluted with deionised water
Silver Reagent	10 fold stock diluted with deionised water
Developer	32g/litre

The Silver Stain method was adapted for mini gels (7cmx8cmx0.75mm) using the Bio-Rad Silver Stain Kit which was adapted from Merril *et al.,* (1981). Firstly gels were immersed in fixative (see table 2.5) for minimum of 30 minutes which tightened the pores in the gel so that the proteins did not leach out into the surrounding solution. Then the fixative was drained off and the gel immersed in oxidiser (see table 2.5) for 5 minutes. Oxidiser was removed from the gel by large volume water washes for not more than 15 minutes so as not to remove it from the proteins. Gels were then immersed in silver reagent (see table 2.5) which deposits silver nitrate on the proteins. A quick water rinse after 20 minutes removed most of the silver reagent, and then gels were immersed in developer which reacted with the silver reagent bound to the proteins in the gel. The developer is a reducing agent which reduces the silver nitrate to silver metal ions. Therefore, where silver nitrate was deposited on the proteins, it was reduced to silver showing the protein bands as a dark grey colour. The developer was changed every 5 minutes or so until the bands appear on the gel at the right intensity. Development of the gels was then stopped using 5% acetic acid, which also tightened the pores in the gel so that the proteins did not leach out of the gel.

2.2 Introduction to crystallography

Crystallography involves the growing of crystals from a solution of protein and using such crystals to obtain diffraction data which gives an electron density map from which protein structure can be determined.

Crystal growth is achieved by super-saturation of a protein in solution achieved by a change of precipitant concentration (such as PEG), protein concentration, pH or temperature (Wood, in Harris and Angal 1990). Under these conditions, proteins aggregate slowly and reach a critical size from which crystal growth proceeds spontaneously (Wood, in Harris and Angal 1990). Each protein crystal is a lattice of protein molecules that contact each other through weak forces such as hydrogen bonding, ion-pairs and hydrophobic interactions (Wood, in Harris and Angal 1990). To obtain the structure of the proteins in the crystal, X-ray diffraction is used as X-rays have a wavelength small enough to be diffracted by protein atoms unlike light which has a wavelength too large to resolve individual atoms.

Freezing crystals before diffraction data collection is called cryo-crystallography and has a number of advantages such as increasing molecular order, reduction of radiation damage to the crystal, and reduction of water X-ray scattering. Whilst diffraction data is being collected, the crystals are kept in a stream of cold nitrogen gas which keeps them frozen. Suitable crystals to be used for X-ray diffraction are soaked with cryoprotectant which removes water present in the crystal to prevent ice formation and cracking of the crystal. Diffraction data from the crystal is collected and used in structural studies.

For *Limulus* pentraxin crystal trials, super-saturation is achieved slowly using vapour diffusion. Purified protein is mixed with "mother liquor" that contains a precipitant, and kept in a sealed environment with a reservoir of mother liquor. Water from the protein-mother liquor mixture evaporates and diffuses into the mother liquor reservoir where the concentration of buffer and precipitant is higher. The conditions used in the mother liquor are refined to get the best crystal growth.

Materials:

Equipment:

- 24 well Linbro plate (MP Biomedicals GMBH)

- Coverslips (Molecular Dimensions)

- Microbridges (Crystal Microsystems)

- 0.2μm pore acetate filter (Sartorius).

Chemicals:

- Chemicals used in crystallisation trials (Sigma).

Procedure:

The sitting drop vapour diffusion method was adopted for all the crystal trays set up. This involved setting microbridges into wells on a 24 well Linbro plate (4x6 wells) (see figure 2.4).
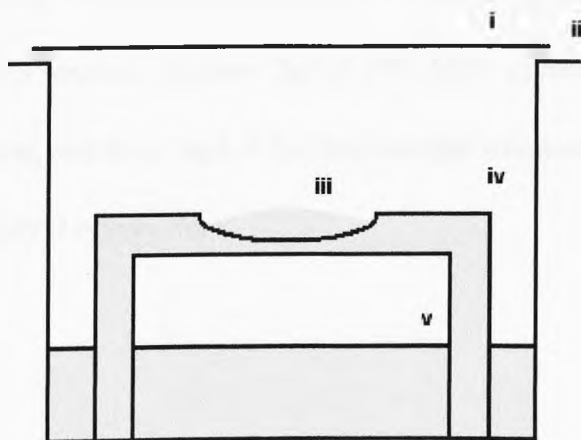
Figure 2.4. Diagram of sitting drop crystallisation method used. i- coverslip, ii- vacuum grease, iii- sitting drop, iv- microbridge, v- mother liquor. 1ml of mother liquor was placed in the well. 2μl of mother liquor and 2μl of protein solution was placed on the microbridge in a sitting drop. The well was sealed with vacuum grease and coverslip.

In the bottom of the well was the mother liquor solution for the crystal trial, usually composed of a buffer for example TRIS, at a selected pH; a precipitant for example polyethylene glycol, at a selected percentage; sometimes also calcium chloride ($CaCl_2$) to

provide the proteins with calcium ions, and a ligand such as PC or PE. All the solutions used in the crystal trays were filtered using a 0.2µm cellulose acetate membrane to remove any bacterial or fungal spores that could grow in the well and affect crystal growth. On the microbridge, 2µl of mother liquor solution from the well was mixed with 2µl of purified protein solution in calcium wash buffer. Each well was sealed with a 0.22mm thick coverslip and sealing grease that kept the well airtight preventing evaporation of solution.

All crystal trays were kept at room temperature and crystal development monitored using a light microscope. When crystals were used for diffraction they were soaked in an appropriate cryoprotectant with the same components of the mother liquor in the well, but also with increasing percentages of methane pentane diol (MPD) (5, 10, 15, 20%) or PEG 400 (5, 10, 20, 30%) to prevent ice formation. Cryobuffers were added in steps for example: 2µl of 5% MPD cryobuffer added to the microbridge solution, then after 5-10 minutes 2µl of 10% MPD cryobuffers added to the microbridge solution, and so on until 20% MPD cryobuffer is reached. Another 2µl of 20% MPD cryobuffer was then added to the microbridge solution, and then 10µl of the microbridge solution is removed and replace with 10µl of the 20% MPD cryobuffer.

Chapter 3. Isolation and characterisation of *Limulus* CRP

3.1. Introduction to *Limulus* CRP

3.1.1 Discovery of *Limulus* CRP

In 1968, studies (Marchalonis and Edelman 1968; Fernandez-Moran *et al.*, 1968) were carried out to isolate and characterise the haemagglutinating fraction of *Limulus polyphemus* haemolymph which had been previously been shown to agglutinate mammalian erythrocytes by Noguchi in 1903 (referenced in Armstrong and Quigley 1987). Isolation of the haemagglutinin was achieved by ultracentrifugation and gel filtration, which showed that approximately 3-10% of the haemolymph of *Limulus polyphemus* was haemagglutinin (Fernandez-Moran *et al.*, 1968). This haemagglutinin was shown to be unrelated to haemocyanin and had an approximate molecular weight of 399±9kDa, dissociating into two different kinds of subunit of approximately 22.5±0.6kDa in weight (Marchalonis and Edelman 1968). The presence of calcium chloride greatly increased the haemagglutinating activity (Marchalonis and Edelman 1968). Analysis by electron microscopy (Fernandez-Moran *et al.*, 1968) showed that the *Limulus* haemagglutinin was a uniform ring structure of apparent hexagonal shape approximately 100Å in diameter. It was also shown to have stacking conformation.

This *Limulus* haemagglutinin termed "Limulin" was isolated by sepharose chromatography from *Limulus* haemolymph, and further purified by sepharose chromatography to an isolated agglutinating fraction (Roche and Monsigny 1974). The agglutinating fraction contained protein (Limulin) of molecular weight 335kDa which dissociated into subunits of ~19kDa in the presence of SDS and β-mercaptoethanol, suggesting that the protein was an octodecamer of covalently bound subunits. Despite this single sized band ~19kDa on reduced SDS-PAGE, isoelectric focusing of Limulin gave three close bands at pH5 ±0.1 suggesting the presence of three isoforms of Limulin with

microheterogeniety, i.e. "isolectins" (Roche and Monsigny 1974). The same study also showed that Limulin had 3.6% total sugar content, and had its lectin ability and haemagglutinating ability inhibited by *N*-acetylneuraminic acid, *N*-glycolylneuraminic acid and *N*-acetylglucosamine as well as human orosomucoid which had high sialic-acid content. Inhibition was higher with sialic-acid bound glycoprotein than with free sialic acids.

Later, Limulin was suggested to be a type of C-reactive protein in *Limulus* which had calcium dependent PC binding ability and bound to pneumococcal C-polysaccharide and PC-BSA (Robey and Liu 1981). The *Limulus* C-reactive protein (CRP) was isolated from whole haemolymph using its affinity for PC-sepharose, and had an approximate molecular weight of 500kDa. Approximately 4mg of CRP per ml of pooled haemolymph was isolated suggesting a continual high concentration present in the haemolymph of the horseshoe crab (Robey and Liu 1981). Immunological cross-reactivity and SDS-PAGE indicated that Limulin and *Limulus* CRP were identical, which was further supported by amino acid analysis of each protein showing highly similar percentage amino acid composition. SDS-PAGE of Limulin/*Limulus* CRP identified two subunits of approximate size 18kDa and 24kDa, of which amino acid analysis showed significantly different amino acid composition (Robey and Liu 1981). From these results it was suggested that each *Limulus* CRP was a doubly stacked hexamer of complexes of both subunits. Addition of PC to *Limulus* CRP did not reduce its haemagglutinating activity in the study, but addition of sialic acid did. This indicated that PC and sialic acid binding sites were not the same on *Limulus* CRP, and that the sialic acid binding site was required for haemagglutination.

Later studies (Armstrong *et al.,* 1994) showed that a sialic-acid and PE –binding CRP subform in *Limulus* which was termed "Limulin" had haemagglutinating properties, and it may have been the causative factor in the haemagglutination observed in the *Limulus* CRP fraction described by Robey and Liu (1981). "Limulin" is closely related to *Limulus* CRP

in size (300kDa native, 24-28kDa subunit), N-terminal peptide sequence and PE binding

properties (Quigley *et al.*, 1994). Limulin is less than 1% of total *Limulus* pentraxin

(Quigley *et al.*, 1994) and therefore is a minor CRP component with fetuin and sialic acid

binding properties as well as agglutination ability, which distinguish it from other *Limulus*

CRP. At such a small percentage of pentraxin, Limulin was not found in earlier studies.

3.1.2 Sequence analysis of *Limulus, Tachypleus,* and *Carcinoscorpius* CRP.

In a study to investigate *Limulus* CRP at both protein and nucleotide level (Nguyen *et

al.*, 1986a, b), *Limulus* CRP was cleaved either chemically or enzymatically to produce

peptide fragments whose amino acid sequences were analysed, and also used for creating

DNA probes that were degenerate and complementary to *Limulus* CRP genes. An EMBL-3

phage library was created with genomic DNA from *Limulus* amebocytes, and screened

with three non-overlapping probes giving results suggesting multiple CRP genes, of which

three clones 1.1, 1.4, and 3.3 gave strong hybridisation signals. These three clones were

used to sequence three different CRP genes which all coded for a signal peptide of 24

amino acids followed by 218 amino acids. The gene was not separated by an intron, unlike

the human CRP gene that has an intron positioned in the area of gene coding for mature

protein (Lei *et al.*, 1985). Therefore, each *Limulus* CRP gene has a continuous reading

frame which was suggested to streamline constitutive expression (Nguyen *et al.*, 1986 a).

Limulus CRP from clone 1.1 and 3.3 was 1347 nucleotides long whilst *Limulus* CRP 1.4

was shorter at 1325 nucleotides long. All three *Limulus* CRP genes have high sequence

identity. *Limulus* CRP 3.3 and 1.1 amino acid sequences are most similar with only three

codon differences, 2 of which code for different amino acids, but 1.4 has 39 codon

differences to 3.3 and 1.1, 23 of which code for different amino acids. DNA from the

Limulus amebocyte was digested and probed with a DNA probe corresponding to 240-540

residues of *Limulus* CRP 1.4 gene sequence, which gave multiple DNA fragments

indicating multiple CRP genes of CRP 1.4 in the *Limulus* genome. This, combined with multiple hybridisation during clone screening, indicates many CRP genes in the *Limulus* genome.

The microheterogeneities found in the nucleotide sequences of 1.1, 1.4, and 3.3, were also present in the isolated peptides (Nguyen *et al.*, 1986b), confirming the presence of at least three *Limulus* CRP expressed genes. Each of the three peptides corresponding to each *Limulus* CRP gene sequenced, shared an identical amino acid terminal sequence up to residue 44, and an identical carboxyl amino acid sequence from 206-218. *Limulus* CRP cleavage gave additional peptides (Nguyen *et al.*, 1986b) which were not identical to the *Limulus* CRP 1.1, 1.4, or 3.3, but had extensive sequence homology indicating that there were more than just three *Limulus* CRP proteins. Analysis of the amino acid sequence of all three *Limulus* CRP proteins showed that 1.4 and 3.3 have six cysteines which form three intrasubunit disulphide bonds, and that 1.1 has seven cysteines which form three intrasubunit disulphide bonds leaving a single cysteine (Nguyen *et al.*, 1986b). The three disulphide bonds were found by thermolysin cleavage, to be between Cys38-Cys101, Cys88-Cys120, and Cys183-Cys217, in all three *Limulus* CRP, where Cys38-Cys101 is a conserved pentraxin disulphide bridge.

Cyanogen bromide cleavage of methionine residues that differ in position amongst the three *Limulus* CRP proteins, indicated that all three *Limulus* CRP proteins are present in equimolar amounts in pooled haemolymph which, along with failure to isolate them individually by mercury-sepharose chromatography, indicated that they might be aggregated, perhaps into the hexagonal structure of two subunits of each CRP as suggested by Robey and Liu (1981). This however, would not take into account the other types of *Limulus* CRP proteins that may be present as suggested by these studies. As pooled haemolymph was used, there is also a possibility that individual horseshoe crabs possess only one or two of the CRP forms and not all three. Each *Limulus* CRP was glycosylated

(Quigley *et al.*, 1994), and the site was established as Asn123 which occurs in the *Limulus* CRP once as Asn-Ala-Thr.

Amino acid sequence comparison of *Limulus* CRP protein with human CRP protein and other pentraxins suggested that *Limulus* and human CRP diverged 800 million years ago, and with approximately 25% sequence similarity it is also possible they share the same common ancestral gene (Nguyen *et al.*, 1986b).

It has also been found that *Tachypleus tridentatus* (Japanese horseshoe crab) has pentraxin-like proteins in its haemolymph. These were isolated using affinity chromatography against fetuin and PE agarose (Iwaki *et al.*, 1999). These proteins fell into three groups *Tachypleus tridentatus* (t)CRP1, tCRP2, and tCRP3 which all had different properties. Each tCRP represented a group of like genes that were found using degenerate primers based on the N-terminus and C-terminus amino acid sequences of the three tCRP proteins found from affinity chromatography. Each group contained several proteins (tCRP1= 8 proteins, tCRP2 = 7 proteins, tCRP3 = 7 proteins) which were grouped according to their high amino acid sequence similarity to the three initial proteins found: tCRP1, tCRP2, and tCRP3.

tCRP1 bound PE-agarose but not fetuin and did not have haemolytic or haemagglutinating properties making it similar to *Limulus* CRP. tCRP2 bound both PE-agarose and fetuin, and had calcium dependent haemolytic and haemagglutinating properties similar to the lectin Limulin which possesses all of these properties (Quigley *et al.*, 1994; Armstrong *et al.*, 1996). tCRP3 however, had calcium dependent haemolytic and haemagglutinating properties, and bound fetuin, but did not have the pentraxin characteristic PE-agarose binding. tCRP3 had a higher calcium dependent haemolytic and haemagglutinating activity than tCRP2, and may well be the dominant haemolytic protein in the haemolymph of *Tachypleus tridentatus*. Despite differences in binding properties

and activity, the three tCRPs had similar native sizes, subunit sizes, N-terminal sequences, quaternary structure, amino acid sequence length, site of expression, glycosylation, and calcium dependent binding. Native sizes of each protein were tCRP1: 300kDa dissociating into two subunits 29 and 31kDa, tCRP2: 330kDa dissociating into three subunits 29, 31, 33kDa, and tCRP3: 340kDa dissociating into two subunits 31 and 33kDa.

Analysis of amino acid sequences showed not only identical N-terminus sequences of each tCRP subunit and a common potential N-linked glycosylation site at Asn123 similar to *Limulus* CRP, but also showed an additional possible glycosylation site at Asn58 on the tCRP3 sequence. Deglycosylation of subunits from each tCRP indicated that the difference in mass of the subunits was not due entirely to glycosylation. Electron microscopy showed that tCRP1 and tCRP3 were hexagonal rings just like *Limulus* CRP (Fernandez-Moran *et al.*, 1968), but with slightly larger dimensions (Iwaki *et al.*, 1999). Reverse transcription-PCR showed that all three tCRPs were mainly expressed in the haemocytes and hepatopancreas of *Tachypleus tridentatus*.

An LPS binding protein similar in sequence to *Limulus* CRP's and *Tachypleus* CRP's was also identified as a CRP-like protein in *Carcinoscorpius rotundicauda* (Ng *et al.*, 2004). Peptide sequences from mass spectrometry of the CRP-like protein were used to design degenerate primers to clone the CRP gene. Eight CRP sequences were found in *Carcinoscorpius rotundicauda* of which three were homologous to CRP1 group of *Tachypleus tridentatus*, and five were homologous to CRP-2 group of *Tachypleus tridentatus* (Ng, *et al.*, 2004). Tests with *Pseudomonas* infection in *Carcinoscorpius rotundicauda* showed that CRP is up regulated during infection and that the main site of CRP synthesis is the hepatopancreas (Ng, *et al.*, 2004).

3.1.3 Structural studies of *Limulus* CRP

Early studies into *Limulus* CRP which at the time was thought to be a haemagglutinin (Limulin), showed that under an electron microscope, it was an hexagonally shaped ring of approximately 100Å diameter and a core of 20-40Å in diameter with some apparent stacking of rings (Fernandez-Moran *et al.*, 1968). Further studies on the properties and characterisation of Limulin/*Limulus* CRP supported the hexagonal ring structure theory (Robey and Liu 1981). An investigation to obtain X-ray diffraction data from crystals of *Limulus* CRP was not successful in solving the structure, but was able to show that *Limulus* CRP crystals could produce diffraction to 2.8 and 2.9Å, and that the unit cells formed in the crystal suggest that *Limulus* CRP is a "high oligomer" such as a hexamer or pentamer (Myles *et al.*, 1990).

More detailed studies into *Limulus* CRP were carried out to analyse the subunit composition (Tennent *et al.*, 1993) and oligosaccharide composition (Amatayakul-Chantler *et al.*, 1993). Reduced SDS-PAGE of *Limulus* CRP, isolated from *Limulus* haemolymph using PC-affinity chromatography, showed three bands at approximately 25kDa mass (Tennent *et al.*, 1993), the top two of which were closely spaced and appeared as a single band in previous studies (Robey and Liu 1981). There were also faint bands at approximately 70 and 50kDa and a series of lower mass bands thought to be from cleavage of CRP subunits. On native PAGE *Limulus* CRP appeared as a single band around 300kDa in size. In addition to the individual subunits of approximate size 25kDa, this supports the theory that *Limulus* CRP is a hexamer.

Mass spectrometry gave more precise sizes of both subunits and oligosaccharides present in *Limulus* CRP, and showed that two subunits corresponded in size to *Limulus* CRP 1.4 of mass 24178Da (Tennent *et al.*, 1993). Other subunit sizes found were 24281Da and 24523Da which did not correspond to the sizes of *Limulus* CRP 1.1, 1.4 or 3.3 and so

were previously uncharacterised. Mass spectrometry of *Limulus* CRP subunits showed many different masses of subunits, each with different glycosylations, the three main oligosaccharides of which are shown in figure 3.1 (Amatayakul-Chantler *et al.*, 1993).

Figure 3.1. Proposed structures of three major oligosaccharides from *Limulus* CRP (adapted from Amatayakul-Chantler *et al.*, 1993).

Oligosaccharide B1 (biantennary):
Manα6Manα-6
             \> Manβ4GlcNAcβ4GlcNAc
Manα2Manα-3

Oligosaccharide C3 (biantennary):
Manα6Manα-6
             \> Manβ4GlcNAcβ4GlcNAc
     Manα-3

Oligosaccharide D1 (monantennary):
Manα3Manα6Manβ4GlcNAcβ4GlcNAc
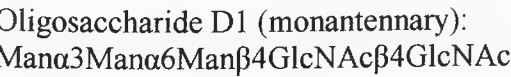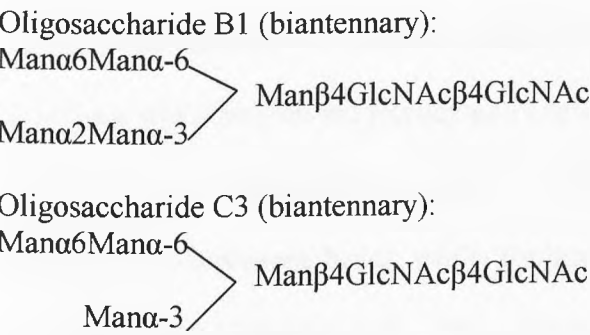
Mass spectrometry was performed on individual horseshoe crab blood rather than pooled haemolymph and showed a predominance of one of the three subunits identified by Nguyen *et al.*, (1986b) rather than apparent equimolarity of all three (Tennent *et al.*, 1993). This may imply expression of different *Limulus* CRP genes in individuals, or expression of genes according to required function. This may accompany variable glycosylation both within CRP subunits in an individual as seen by mass spectrometry results, or between individuals as shown by variable ratios of oligosaccharides characterised by mass spectrometry for individual and pooled haemolymph. Variation in concentration of *Limulus* CRP in individuals in addition to the variable subunit glycosylation and expression, suggests that *Limulus* CRP is a family of proteins with possible variable functions. Another difference between using pooled haemolymph and that from individuals was the concentration of CRP isolated per ml of haemolymph which was an average of 1.83mg/ml ranging from 0.275-6.64mg/ml, and so not consistent as previous reports based on pooled haemolymph suggest such as ~1mg/ml (Nguyen *et al.*, 1986a) and 4mg/ml (Robey and Liu 1981).

3.1.4 *Limulus* CRP binding properties.

Limulus CRP was first identified as having CRP-like PC-binding ability by Robey and Liu in 1981. Analysis and comparison of this PC binding ability (Robey and Liu 1983) indicated that each mole of *Limulus* CRP bound 0.5mole of phosphate in comparison to human CRP which bound 1 mole phosphate per mole CRP. One speculation is that due to subunit heterogeneity in *Limulus* CRP (Robey and Liu 1981), some subunits bind phosphate and others do not (Robey and Liu 1983). The lower affinity of *Limulus* CRP for PC compared to human CRP was shown by the binding affinity of 0.5mg of *Limulus* CRP per ml of PC-sepharose beads, whilst for human CRP 10mg of protein bound per ml PC-sepharose beads (Tennent *et al.*, 1993). However, phosphocholine competitively removed phosphate equally from human and *Limulus* CRP indicating that PC affinity was equal between the two proteins (Robey and Liu 1983). As phosphate binding to human and *Limulus* CRP was studied using designed spin-labelled phosphate compound, the proposed depth of the phosphate and hence phosphocholine binding site was estimated to be less than 5Å (Robey and Liu 1983). Therefore, if CRP was to bind PC, for example on the surface of disturbed cell membranes (Nakartes and Volanakis 1982) the PC must be extended from the surface by at least 5Å.

It was also suggested that *Limulus* CRP bound PE-sepharose with higher affinity than PC-sepharose, and that even at 100mM PC elution, not all CRP was removed from a PE-sepharose column (Tennent *et al.*, 1993). Later studies (Shrive *et al.*, 1999) used PE-agarose to isolate *Limulus* CRP from *Limulus* haemolymph with 10mM PC elution buffer, and the remaining protein bound to the PE-agarose was eluted with 10mM EDTA. N-terminal sequencing of the EDTA eluted protein showed the presence of some *Limulus* CRP and also a novel protein subsequently named *Limulus* SAP according to its affinity for PE agarose. This agreed with incomplete elution of *Limulus* CRP observed by Tennent *et al.*, (1993). It is possible that isolation of *Limulus* CRP from a PE column did not yield

complete elution with 100mM (Tennent *et al.*, 1993) because *Limulus* SAP was present
which only binds PE not PC.

3.1.5 Isolation of *Limulus* CRP

The presence of multiple CRP like proteins with variable ligand binding properties in
Tachypleus tridentatus, and the presence of multiple genes for *Limulus* CRP suggests that
there may be various *Limulus* CRP type proteins with different ligand binding properties.
Previous studies have already shown that *Limulus* CRP binds to both PC and PE, and that
affinity to these ligands may be different as homologous proteins such as tCRP3 do not
bind PE-agarose at all.

The purpose of this study was to find ligands that bound to *Limulus* CRP, and to attempt
to isolate different *Limulus* CRP molecular aggregate forms by their individual binding
affinity to these ligands. Previously known ligands PC and PE were used, PC as a 10mM
elution and PE as a 30mM gradient. It has been shown that human CRP binds to ribose-5-
phosphate and adenosine-5-monophosphate (AMP) (Lee *et al.*, 2002) and so these ligands
were used as elution buffers on *Limulus* CRP bound to a PE-agarose column. The sugars
N-acetylglucosamine, mannose, galactose, gluconic acid, galacturonic acid and glucuronic
acid were cheap ligands to try, and so these were used in elution buffers too. Salt elutions
were also performed to check that any CRP elution was due to ligand binding rather than
an ionic change caused by increased salt concentration.

All ligands mentioned were used as elution buffers to elute *Limulus* CRP bound to a PE-
agarose column monitored by UV absorbance at 280nm. Protein that was eluted was
analysed using SDS-PAGE to identify subunit forms. CRP that was eluted using PC,
ribose-5-phosphate and AMP was also analysed with a PhenylHiTrap hydrophobicity
interaction column on a FPLC system monitored using UV absorbance at 280nm and

215nm. The hydrophobicity interaction column separates out molecular aggregate forms of proteins according to their surface hydrophobicity, so the surface hydrophobicity of CRP eluted by ribose-5-phosphate and AMP was compared with that of whole *Limulus* CRP. If CRP isolated using a ligand such as ribose-5-phosphate or AMP contained different CRP molecular aggregate forms with different surface hydrophobicities, this would be noticeable on the UV trace of the FPLC.

CRP isolated using PC from whole plasma, and ribose-5-phosphate from whole CRP was used in several crystallisation trials in order to grow crystals to produce high quality diffraction data. This diffraction data could then be used to determine the *Limulus* CRP structure. Crystallisation trials were set up using conditions similar to those that had previously been successful in growing *Limulus* CRP crystals. CRP isolated using ribose-5-phosphate was used in order to obtain more ordered crystals by lowering the number of CRP isoforms.

3.2 Materials and Methods.

3.2.1 Isolation of *Limulus* CRP from *Limulus* plasma

The protocol for elution of *Limulus* CRP and *Limulus* SAP was based on that used by Shrive *et al.*, (1999) and Tharia *et al.*, (2002) with an initial elution of 10mM PC to elute CRP (Robey and Liu 1981), then 30mM PE to elute SAP (Tharia *et al.*, 2002) and a final elution of 10mM EDTA. Approximately 5-10mls of *Limulus* plasma (either whole or PEG-cut) was filtered with a 0.2μm cellulose filter to remove bacterial and fungal spores as well as residual PEG in the case of PEG-cut plasma, and diluted to 10-20mls volume with calcium wash buffer. Using the BIORAD Biologic LP system, the diluted plasma was washed through a 25ml volume PE-agarose column at a flow rate of 0.5ml/min, which had previously been equilibrated with calcium wash buffer. Protein elution from the PE-

agarose column was monitored using UV absorbance at 280nm and shown as an elution profile using LP Data View program on a BioRad Biologic HR controller. Once all plasma had been loaded, the column was washed through with calcium wash buffer until the UV_{280nm} reading was stable. Approximately 75mls (i.e. three column volumes) of 10mM PC in calcium wash buffer was washed through the column to elute *Limulus* CRP. The PE-agarose column was then equilibrated with calcium wash buffer, and SAP eluted with 30mM PE in calcium wash buffer (Shrive *et al.*, 1999). 10mM EDTA eluted any remaining calcium dependently bound proteins, and regeneration buffers A (500mM NaCl, 200mM Tris, 10mM EDTA), B (500mM NaCl, 50mM Tris, 10mM $CaCl_2$), and C (150mM NaCl, 50mM Tris, 10mM $CaCl_2$, 0.02% NaN_3) further cleansed the column of any remaining proteins using EDTA and high salt concentrations.

Isolated CRP was dialysed against EDTA buffer to remove the calcium dependently bound PC from the CRP, and then dialysed against calcium wash buffer to replace calcium lost by the previous dialysis. After dialysis, CRP was concentrated to a volume or concentration suitable for either re-application to the PE-agarose column, application to the FPLC, or use in crystallisation trials. Concentration was performed using Amicon Ultra 15ml and 4ml centrifugal filter devices with a molecular weight cut off of 10kDa at a speed of 2424rpm. CRP solution absorbance was measured at 280nm and 320nm against calcium wash buffer, and the concentration of CRP worked out as the absorbance at 320nm (background absorbance) subtracted from the absorbance at 280nm (absorbance of protein according to presence of aromatic rings, see section 2.1.5) divided by the extinction coefficient of *Limulus* CRP which is approximately 1.5 (Abs 0.1%). The extinction coefficient was worked out by entering the existing sequences of *Limulus* CRP 1.1, 1.4, and 3.3 into an online "ProtParam" program from Expasy.

3.2.2 Isolation of *Limulus* CRP by affinity chromatography

Ligands used to isolate CRP forms were ribose 5-phosphate, adenosine 5-monophosphate (AMP), PC, PE, galacturonic acid, glucuronic acid, N-acetylglucosamine, mannose, and galactose (all from Sigma). Calcium wash buffer was added to approximately 6mg *Limulus* CRP to give a volume of between 10-20mls that was then re-applied to the 25ml PE-agarose column which had been equilibrated with calcium wash buffer. The flow rate of buffers through the column was 0.5ml/min. After CRP application, the column was washed through with calcium wash buffer to stabilise the UV_{280nm} reading by washing through any CRP that had not bound to the column. The column was then washed with calcium wash buffer that contained one of the previously mentioned ligands (e.g. AMP) at a specific concentration (i.e. 5mM, 10mM or 30mM) for a volume of approximately 75mls or until the absorbance reading at UV_{280nm} was stable.

Gradients of buffers ribose-5-phosphate, AMP and PE were prepared by mixing buffer of the required highest concentration, for example 30mM AMP, with calcium wash buffer. This was performed using a single pump that mixed volumes of the two buffers from 0% of ligand buffer to 100% ligand buffer over 75mls. Eluted protein was collected manually or using the fraction collector. After ligand addition to the column and protein elution, the column was re-equilibrated with calcium wash buffer and then 10mM EDTA buffer was washed through to remove CRP that had not been eluted with the ligand buffer. In-between elutions approximately 50-75ml of flow through was collected from the column to observe the amount of CRP disassociating from the column without ligand presence. This was found to be of an average of approximately 0.01mg of protein per mg CRP eluted in total from the column. The column was then subjected to regeneration involving the regeneration buffers A, B, and C which contained EDTA, high salt concentrations and sodium azide respectively, to remove remaining proteins on the column and prevent fungal and bacterial growth on the column.

CRP eluted with ligand buffer was subjected to dialysis with EDTA to remove the ligand that was bound calcium dependently to the CRP, and then with calcium wash buffer to replace calcium lost in the previous dialysis. CRP eluted using EDTA was dialysed against calcium wash buffer to replace calcium ions lost during elution. CRP was then concentrated to suitable concentrations for crystallisation trials or application to a hydrophobicity column, using Amicon Ultra 15ml and 4ml centrifugal filter devices at a speed of 2424rpm. The concentration of CRP was determined as previously mentioned in section 3.2.1.

Isolation of *Limulus* CRP using ligands AMP and ribose-5-phosphate was repeated 2-3 times, whilst isolation using other ligands was performed only once.

3.2.3 Isolation of *Limulus* CRP by Fast Protein Liquid Chromatography (FPLC)

Limulus CRP was further separated using a FPLC system with a hydrophobicity interaction column that separates proteins by their hydrophobic properties. Approximately 350μg of CRP was made up to 2ml with binding buffer that contained 20mM Tris, 10mM $CaCl_2$, 3.5M $(NH_4)_2SO_4$. A 1ml HiTrap Phenyl HP hydrophobicity interaction column (HIC) was equilibrated with binding buffer. CRP was applied to the column in the binding buffer at a flow rate of 0.5ml/min. CRP was then eluted from the column using 40mls of a 0-100% gradient of elution buffer containing 20mM Tris, 10mM $CaCl_2$, and 100mM NaCl. CRP elution was monitored using UV absorbance at 280nm which monitors UV absorbance by aromatic rings of the amino acids tyrosine and tryptophan, and 215nm which monitors UV absorbance by peptide bonds and is more sensitive than the measurement at 280nm. The HIC was then washed with water, then ethanol for storage. FPLC use was performed with technical assistance from Dr Paul Nield.

Analysis of *Limulus* CRP using FPLC was performed once for each sample.

3.2.4 Analysis of *Limulus* CRP using gel electrophoresis

After isolated *Limulus* CRP had been dialysed and concentrated, it was analysed using gel electrophoresis to show the different sizes of CRP subunits. SDS-PAGE was used which involved denaturing CRP using SDS and occasionally reducing it using β-mercaptoethanol. SDS-PAGE gels were made from 12.5% acrylamide resolving gel, and a 4% acrylamide stacking gel. Each complete gel was 0.75mm thick and had 10 wells placed in the stacking gel. Both CRP samples and the markers were mixed with sample buffer that contained SDS and β-mercaptoethanol for reduced SDS-PAGE gels. CRP samples for SDS-PAGE and markers were heated at 95-100°C for approximately 2 minutes just prior to addition onto the gel. A prepared gel was then placed in an electrophoresis tank with appropriate running buffer, and the samples and markers loaded into the wells. The amount of CRP loaded into each well (except that which contained 10μl of size marker) was approximately 1μg if the gel was to be stained with silver stain, or approximately 5-10μg if the gel was to be stained with Coomassie blue. Gels were left for approximately 40 minutes with a voltage of 200V passed through them.

To visualise protein bands within the gels after electrophoresis, Coomassie blue and silver stain were used. Gels to be stained with Coomassie blue were removed from the electrophoresis tank and submerged in Coomassie blue stain for a minimum of an hour. Coomassie blue stain was then drained off, and the gel immersed in Coomassie blue de-stain solution which was changed several times over a period of 2-3 hours until the background was clear and the protein bands visible. Gels to be stained with silver stain were removed from the electrophoresis tank and immersed in fixative (40% methanol, 10% acetic acid) for a minimum of 30 minutes. After this, gels were stained with silver stain kit (Bio-Rad) which involves immersing the gel in oxidiser for 5 minutes, water washes for a maximum of 15 minutes, immersion in silver reagent for 20 minutes, and then development using immersion of gels in developer for periods of 5 minutes until developed

to the right degree. Gels were then immersed in 5% acetic acid for approximately 15 minutes.

3.2.5 Crystallisation trials of *Limulus* CRP

Isolated *Limulus* CRP was used in crystallisation trials involving a range of different buffers. Whole *Limulus* CRP isolated from plasma, or the fraction of CRP isolated using ligand binding affinity for ribose-5-phosphate, were used. Components to be used in each crystallisation tray were filtered with a 0.2μm pore acetate filter to remove bacterial and fungal spores. The exception to this was polyethylene glycol (PEG) at high molecular weight as this was too viscous to filter. Components of the mother liquor were placed in the wells of 24 well Linbro plates up to a total volume of 1ml. 2μl of mother liquor was taken and placed on a microbridge within the well. 2μl of *Limulus* CRP in calcium wash buffer was added to the mother liquor on the microbridge, and the well sealed with vacuum grease and a coverslip. Crystallisation trays were kept at room temperature and crystal growth observed by microscope. Crystals used to obtain diffraction data were cryprotected with the addition of cryobuffers to the well to replace mother liquor at increasing percentages of MPD of 5, 10, 15, and 20%. Addition of cryobuffers to crystals, and any subsequent use to obtain diffraction data was performed in collaboration with Jenny Patterson, Amy Shaw, and Prof Trevor Greenhough at Daresbury Synchrotron Radiation Source.

3.3 *Limulus* CRP results

The properties of *Limulus* CRP isolated from *Limulus* plasma are identified as both a subunit size in SDS-PAGE and molecular aggregate surface hydrophobicity by FPLC. Sections 3.3.2 to 3.3.8 analyse *Limulus* CRP isolated using isocratic and gradient elutions of ribose-5-phosphate and AMP. Analysis is performed on subunit size by SDS-PAGE, and molecular aggregate surface hydrophobicity by FPLC. Section 3.3.9 details elution of *Limulus* CRP using PC and analysis of molecular aggregate surface hydrophobicity using FPLC. Elutions of *Limulus* CRP using ligands galacturonic acid, glucuronic acid, N-acetylglucosamine, mannose, galactose, and gluconic acid is shown in section 3.3.10. Section 3.3.11 details elution of *Limulus* CRP using a PE gradient and the effect this has on subunit size of CRP from each fraction of the gradient. The elution of *Limulus* CRP using sodium chloride to check whether this affects CRP elution from a PE-agarose column is detailed in section 3.3.12. Section 3.3.13 details crystal trials performed using *Limulus* CRP.

3.3.1 Isolation of whole *Limulus* CRP from *Limulus* plasma.

Whole *Limulus* CRP was isolated from *Limulus* plasma using the methods described in section 3.2.1. A typical elution of pentraxins from *Limulus* plasma involved elution of CRP first then SAP second, whilst remaining calcium bound proteins were removed from the column using EDTA. A typical elution profile for CRP shows a sharp peak with a gradual descending limb occasionally with a shoulder (see figure 3.2). A PC elution provided on average 2.5mg of CRP per ml of plasma, which is high in comparison to approximately 0.3mg of SAP per ml of plasma.
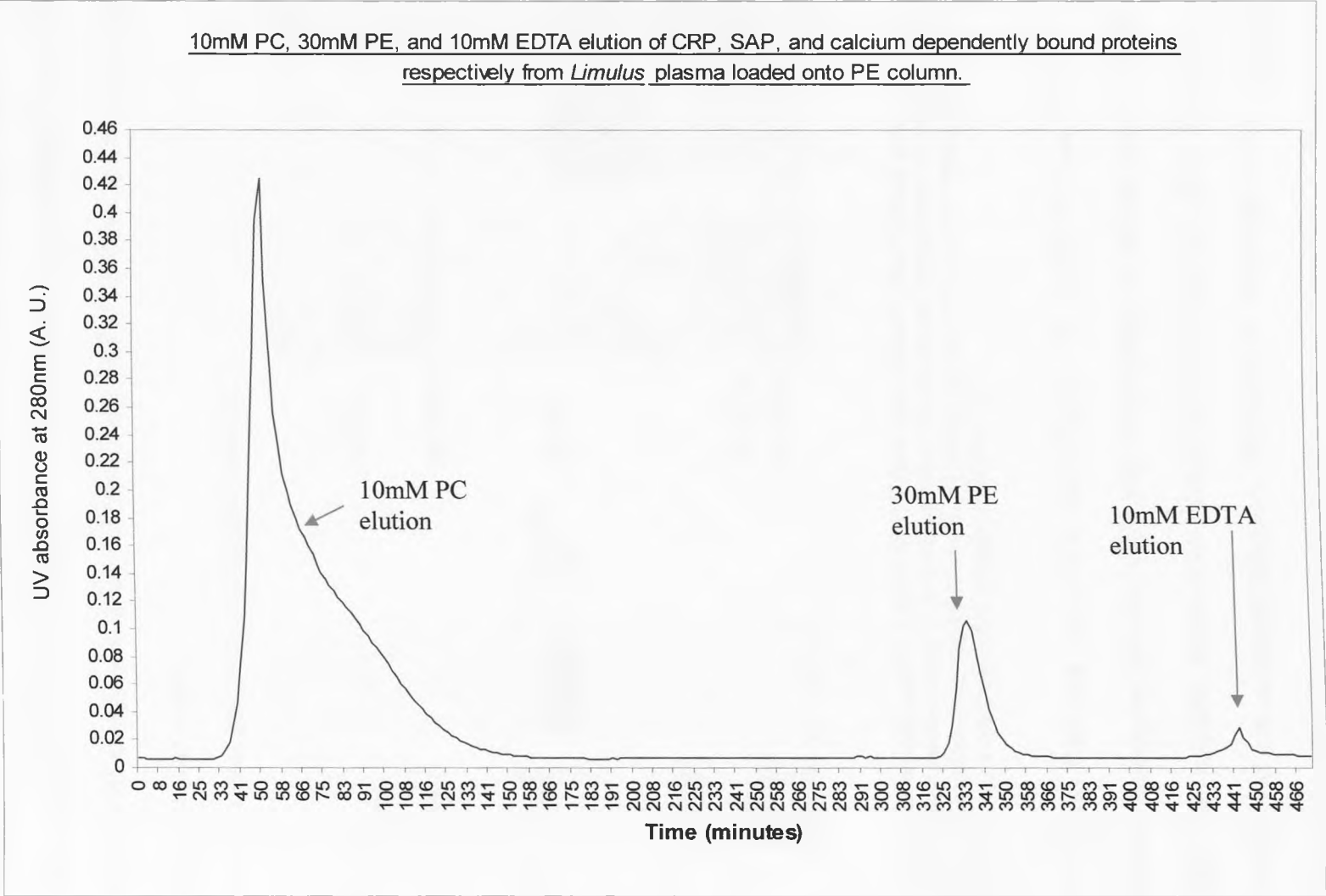
Figure 3.2. Elution profile of *Limulus* plasma with sequential elutions of 10mM PC, 30mM PE, and 10mM EDTA.

A typical SDS-PAGE gel showing the subunit sizes of *Limulus* pentraxins isolated using 10mM PC, 30mM PE and 10mM EDTA elutions is shown in figure 3.3. The PC elution shows three bands ~29kDa in size, the bottom two of which are closely spaced, and a minor band at approximately 66kda in size. The EDTA elution shows different sized bands including some of similar size to the CRP (PC) and SAP (PE) bands and those of lower and higher molecular weight.

Figure 3.3. Silver stained SDS-PAGE gel of *Limulus* pentraxins.
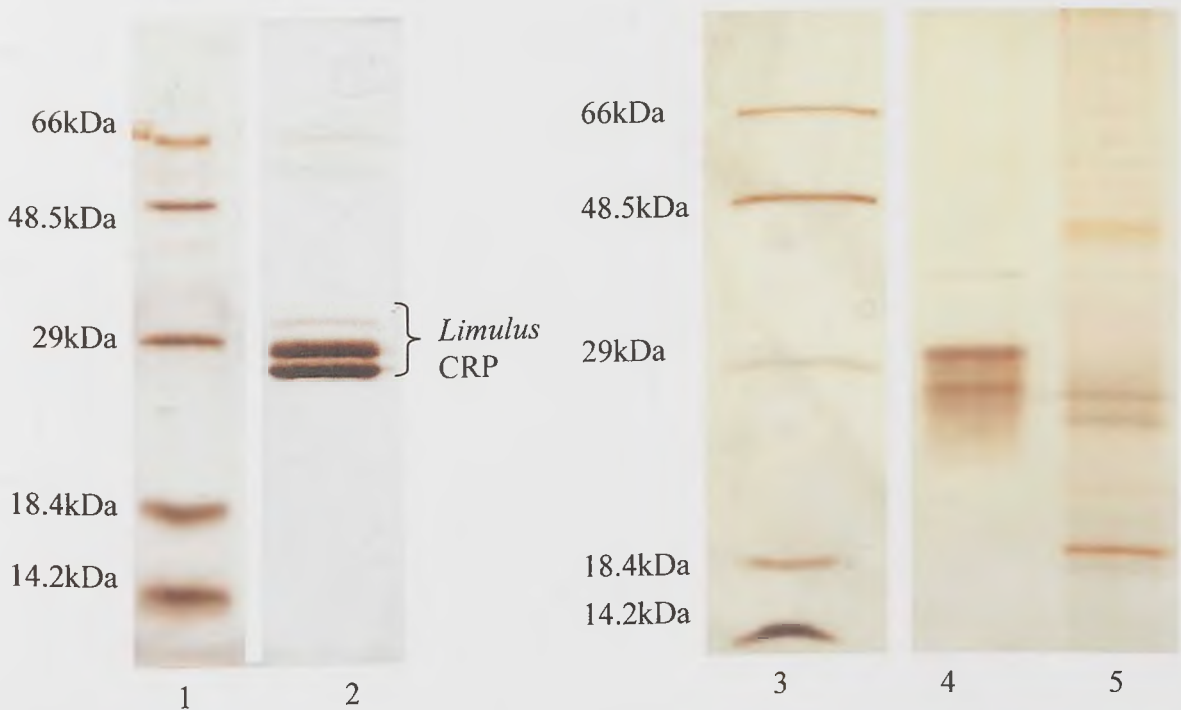
Lane 1: molecular weight markers. Lane 2: CRP isolated from *Limulus* whole plasma using 10mM PC. Lane 3: molecular weight markers. Lane 4: SAP isolated from *Limulus* plasma using 30mM PE elution after elution of CRP. Lane 5: calcium dependent bound remaining proteins eluted by EDTA after 10mM PC and 30mM PE elutions. Note: lanes 1 and 2 are from a separate gel to lanes 3-5.

Whole *Limulus* CRP that was eluted from plasma with 10mM PC was applied to a hydrophobicity interaction column (HIC), and eluted using a decreasing concentration of 3.5M ammonium sulphate (shown as an increase in 0-100% elution buffer). The elution of CRP from the HIC was monitored using UV absorbance at 280nm and 215nm (see figure 3.4). The elution profile shows four distinct peaks, the second and third of which appear to have shoulders on their descending limbs. Overall, there are four major peaks, but six peaks including the shoulders. This clearly shows the presence of at least six molecular aggregate forms of CRP, with different surface hydrophobicity properties, isolated from

Limulus plasma. It is possible that the smaller peaks correspond to trace impurities within the CRP fraction, but when large amounts (15-20mg) of CRP are applied to the HIC these small peaks are still present (data not shown). This suggests that the small peaks may not be representative of trace impurities but are instead CRP. Identification of protein components present in each peak is underway.
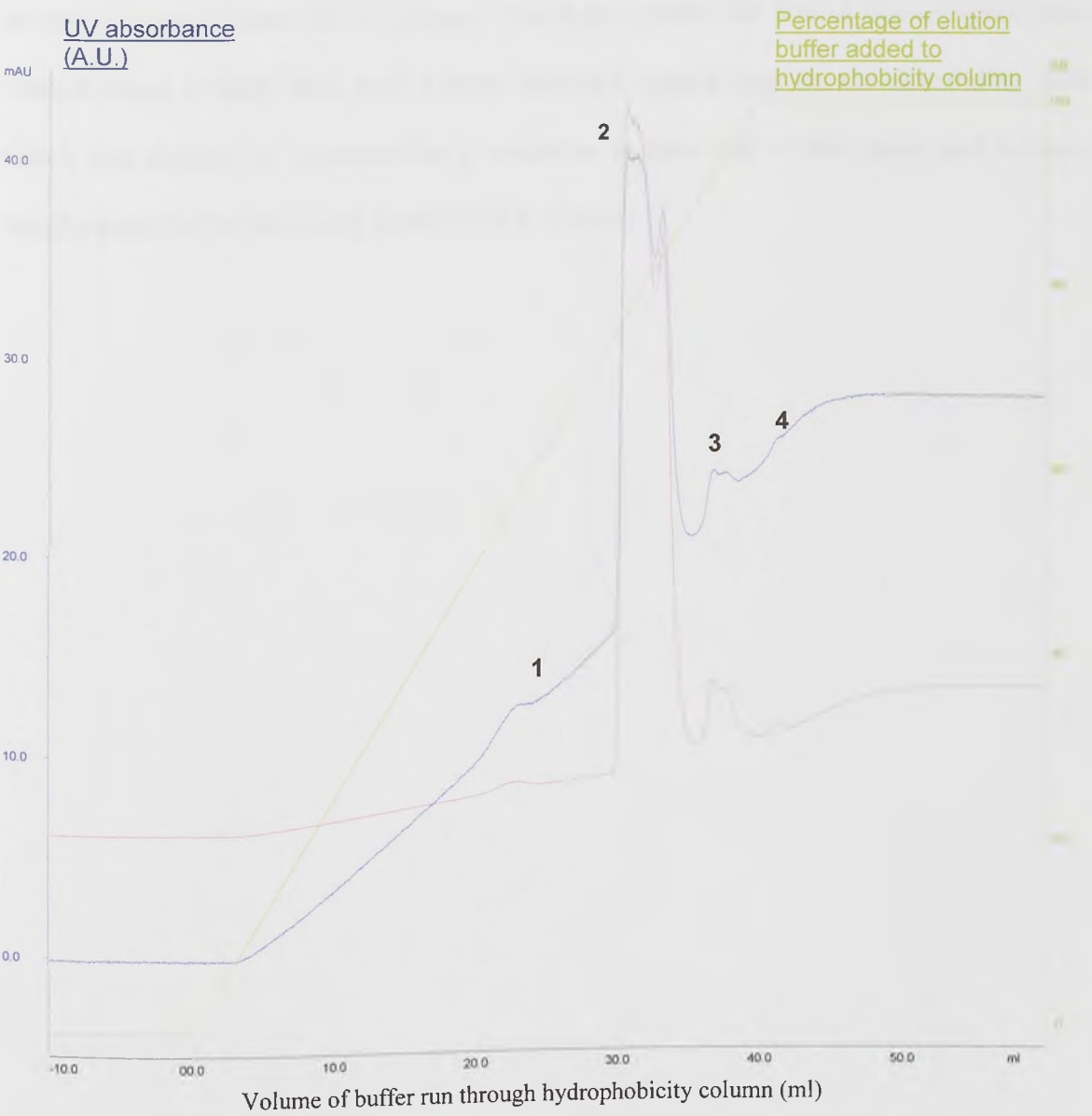
Figure 3.4 FPLC elution profile of *Limulus* CRP isolated from *Limulus* plasma using 10mM PC. UV absorbance at 280nm in shown in blue, UV absorbance at 215nm is shown in magenta, percentage of elution buffer is shown in green. The major peaks are labeled 1-4.

3.3.2 Isolation of *Limulus* CRP using isocratic ribose-5-phosphate elutions

Limulus CRP that had been isolated from *Limulus* plasma using 10mM PC was re-applied to a PE-agarose column as described in section 3.2.2. CRP was then eluted using

an isocratic elution of 10mM or 30mM ribose-5-phosphate buffer, and remaining CRP

removed from the column using 10mM EDTA. Ribose-5-phosphate was chosen as a ligand

because of its binding affinity to human CRP (Lee *et al.*, 2002). The elution profiles are

shown in figures 3.5 and 3.6 for 10mM and 30mM elutions respectively. The elution

profile for 10mM ribose-5-phosphate elution shows a sharp peak and a much smaller peak

for the following 10mM EDTA elution. The elution profile for 30mM ribose-5-phosphate

elution shows a sharp peak with a large extended second peak on the descending limb

which was thought to be caused by a reduction in flow rate at that point, and a much

smaller peak for the following 10mM EDTA elution.
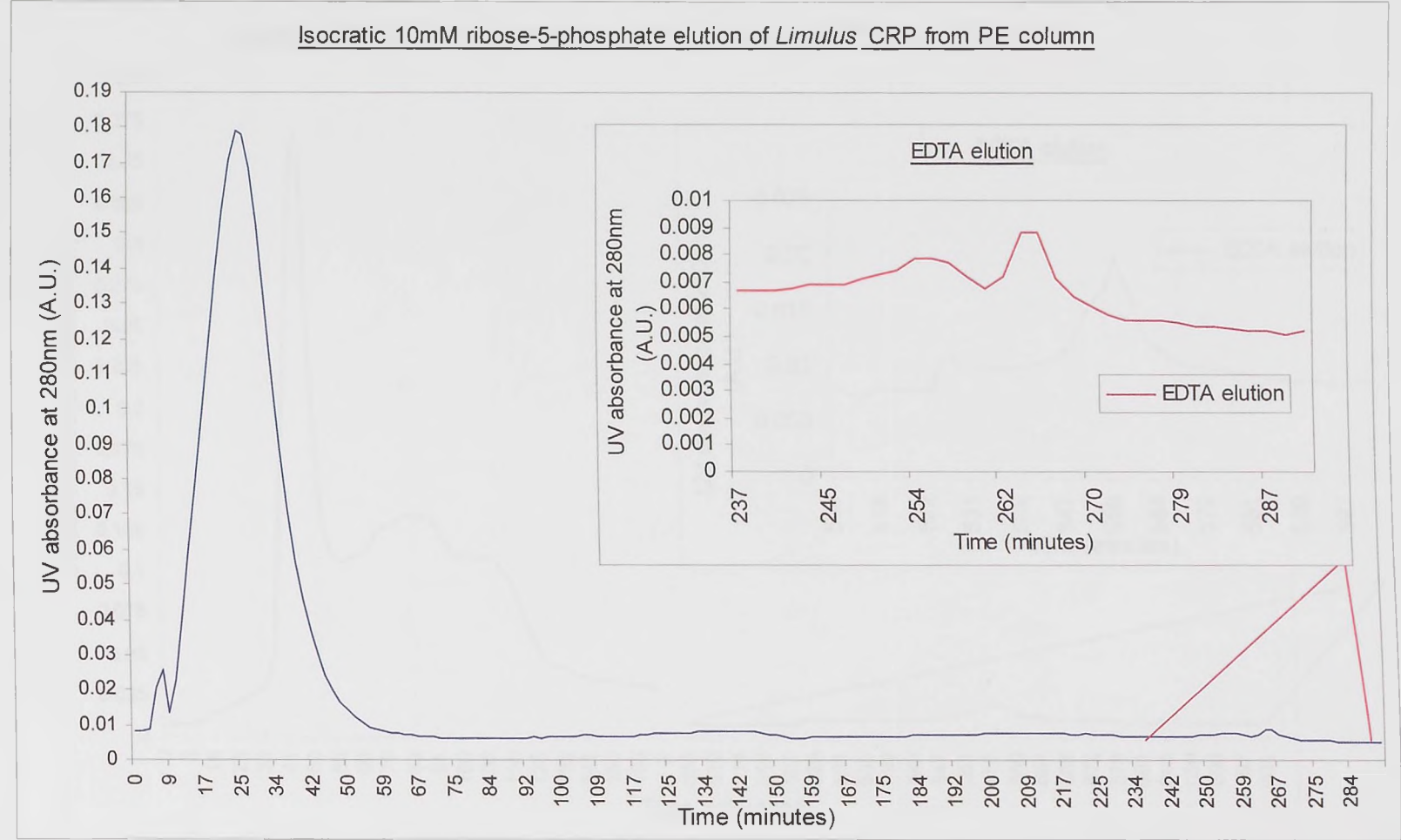
Figure 3.5. Elution profile of *Limulus* CRP eluted using 10mM ribose-5-phosphate then 10mM EDTA.
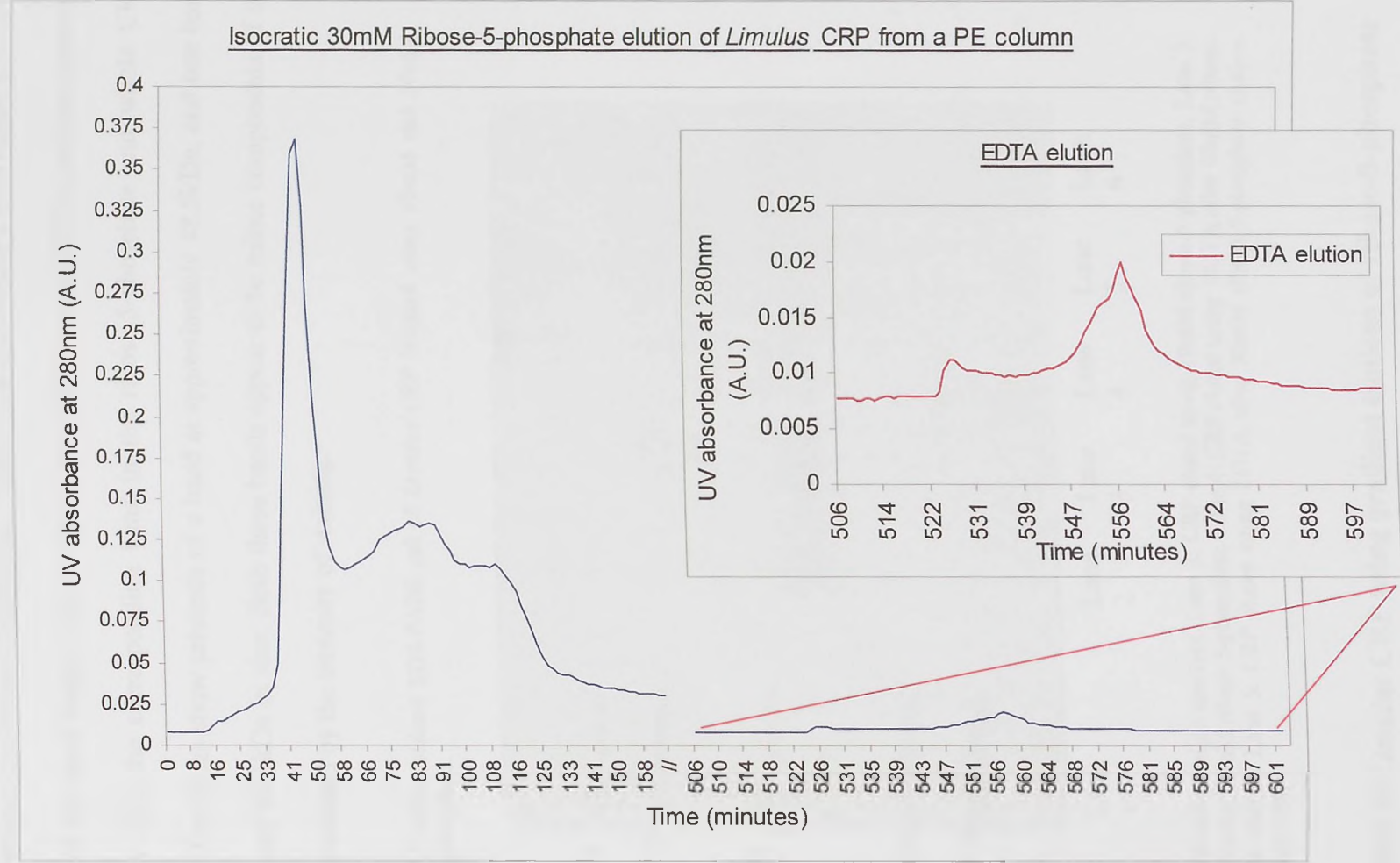
Figure 3.6. Elution profile of *Limulus* CRP eluted using sequential elutions of 30mM ribose-5-phosphate and 10mM EDTA. // indicates a break in time between elutions.

SDS-PAGE analysis of *Limulus* CRP isolated using 10mM or 30mM ribose-5-phosphate

isocratic elutions, and the CRP removed after using 10mM EDTA showed no apparent

difference between the ribose-5-phosphate eluted CRP and whole CRP (figure 3.7). The

larger sized of the three bands ~29kDa size in the CRP from the EDTA elutions appeared

to be darker than the corresponding band in the ribose-5-phosphate and whole CRP

elutions. All five lanes show presence of a band at approximately 48.5kDa, and one band

between 29 and 48.5kDa in size. Both these bands appear to be minor components of the

fractions as indicated by the intensity of staining.

Figure 3.7 Silver stained SDS-PAGE gel of *Limulus* CRP isolated using 10mM and 30mM
ribose-5-phosphate.
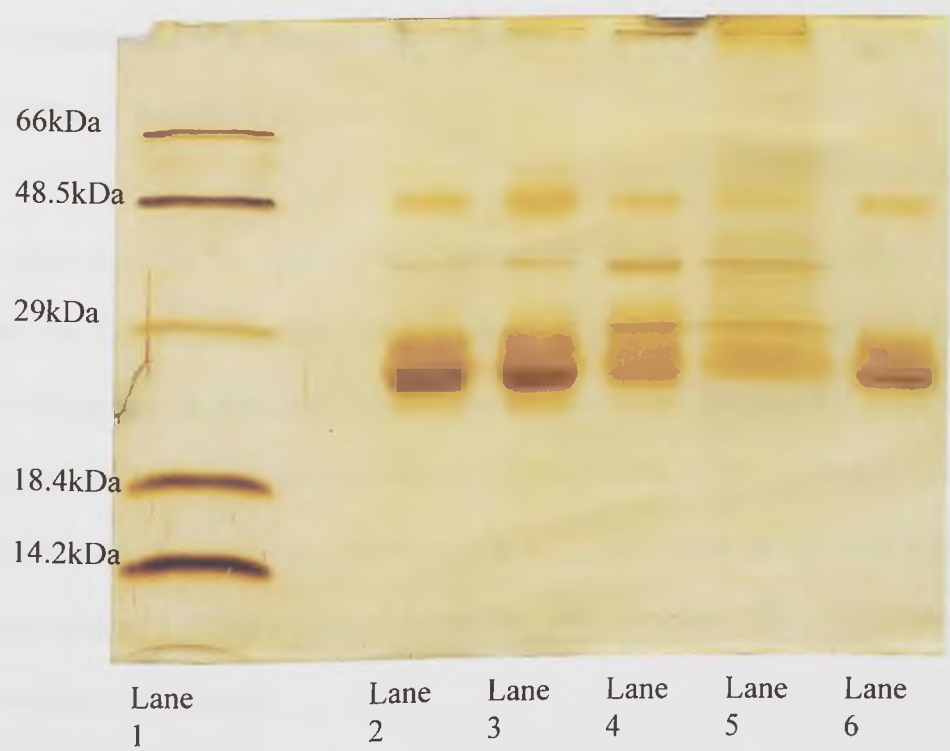
Lane 1: molecular weight markers. Lane 2: CRP eluted using 10mM ribose-5-phosphate. Lane 3:
CRP eluted using 30mM ribose-5-phosphate. Lane 4: CRP eluted using EDTA after 10mM ribose-
5-phosphate elution. Lane 5: CRP eluted using EDTA after 30mM ribose-5-phosphate elution.
Lane 6: whole CRP.

3.3.3 Isolation of *Limulus* CRP using gradient elutions of ribose-5-phosphate

Gradient elutions of 5mM, 10mM and 30mM ribose-5-phosphate were performed to

attempt to identify and isolate individual molecular aggregate forms of CRP with different

binding affinity to ribose-5-phosphate.

The elution profile for 5mM ribose-5-phosphate (figure 3.8) shows the presence of two joined peaks and a large extended peak on the descending limb for the ribose-5-phosphate elution, and a small sharp peak for the 10mM EDTA elution after it. Once 5mM concentration of ribose-5-phosphate had been reached, it was continued until the UV absorbance had dropped, and then the flow rate was lowered after which there was the large extended peak. It was thought that reduction in flow rate affected the UV absorbance as in figure 3.6. During the 5mM ribose-5-phosphate elution, the protein elution peaked at approximately 2.5mM ribose-5-phosphate. It should be noted that the 5mM ribose-5-phosphate elution occurred over a longer period of time (333 minutes without additional large extended peak, 875 minutes with additional large extended peak) than the 10mM and 30mM ribose-5-phosphate elutions.

The elution profile for 10mM ribose-5-phosphate (figure 3.9) shows a sharp peak with a shoulder on the ascending limb and a gradual tailing off on the descending limb for the ribose-5-phosphate elution, and a small sharp peak for the 10mM EDTA elution after it. During the 10mM ribose-5-phosphate gradient, the protein elution peaked at approximately 5.5mM ribose-5-phosphate concentration. The shoulder present on the main peak occurred at approximately 3-3.5mM ribose-5-phosphate concentration. The elution peak covered approximately 144 minutes.

The elution profile for 30mM ribose-5-phosphate (figure 3.10) shows a sharp peak for the ribose-5-phosphate elution, and a small peak for the 10mM EDTA elution after it. The 30mM ribose-5-phosphate elution occurred over a shorter period of time (55 minutes) than the 5mM and 10mM ribose-5-phosphate elutions indicating a higher affinity of *Limulus* CRP for ribose-5-phosphate at 30mM than at 10mM, and a higher affinity at 10mM concentration than at 5mM.

During the gradient, the protein elution peaked at approximately 9mM ribose-5-phosphate.   Gradient elutions were only performed once, and so the concentration of ribose-5-phosphate which corresponded to the peak of protein being eluted may differ should the elutions be repeated many times. However, what is clear is that when the gradient elution is performed, the higher ligand (ribose-5-phosphate) concentration gradients elute protein quicker than the lower ligand concentrations probably due to binding affinity of CRP for ribose-5-phosphate and availability of ligand being increased.
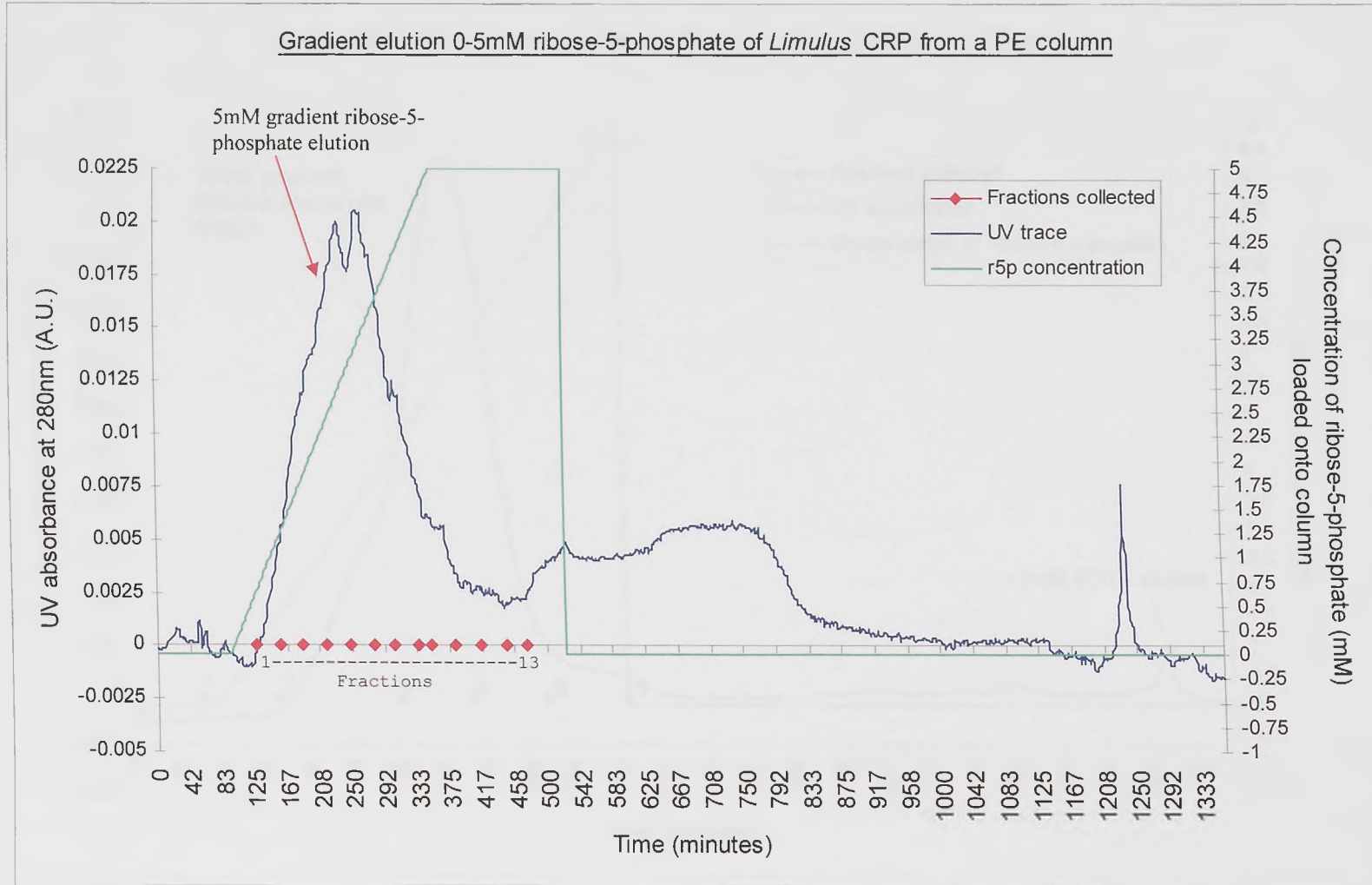
Figure 3.8 Elution profile of *Limulus* CRP eluted using a gradient of 5mM ribose-5-phosphate, then 10mM EDTA. The start of where fractions were collected is indicated.
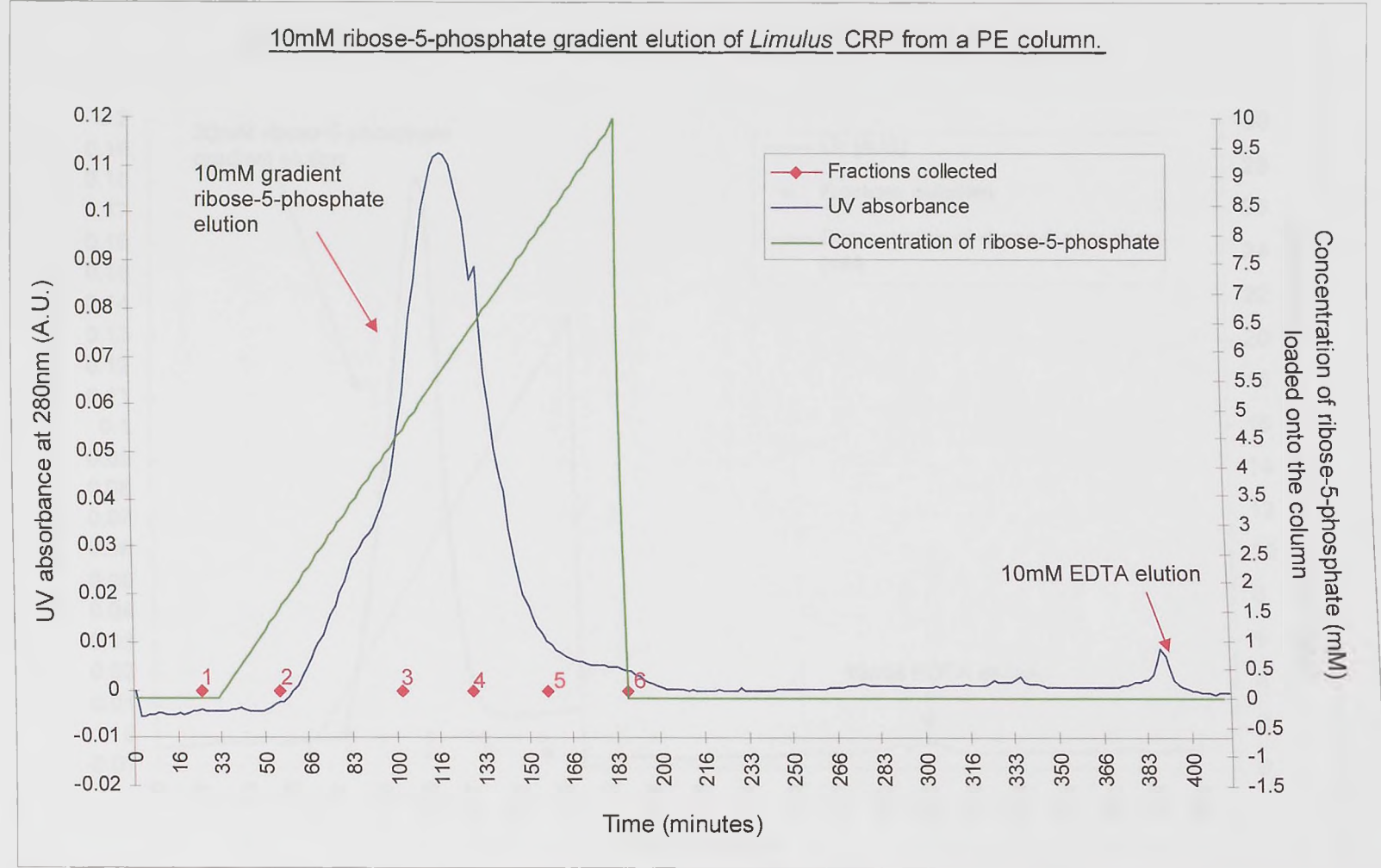
Figure 3.9 Elution profile of *Limulus* CRP eluted using a 10mM gradient of ribose-5-phosphate followed by a 10mM EDTA elution. The start point of collection of fractions 1 to 6 is shown.
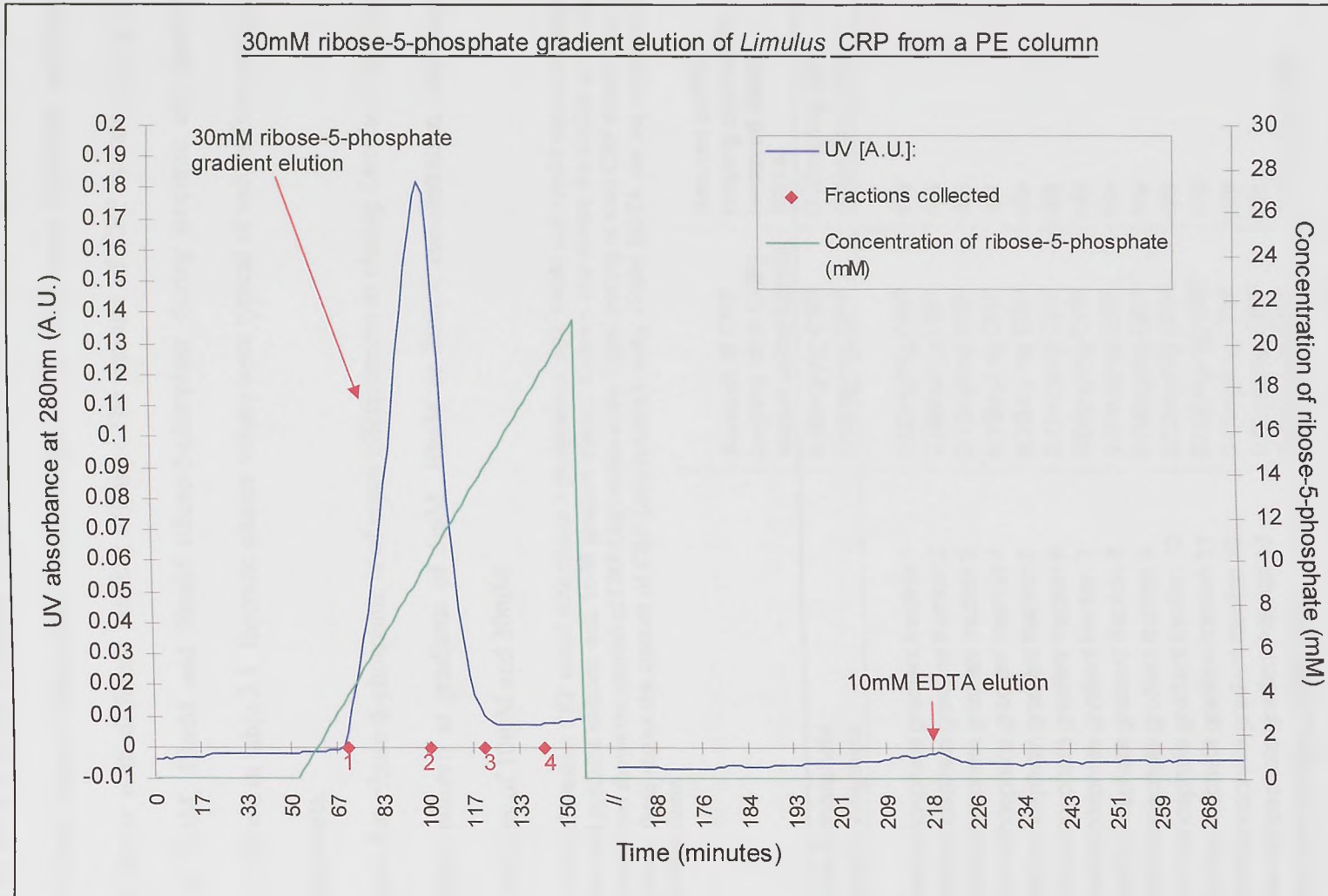
Figure 3.10. 30mM ribose-5-phosphate gradient elution of *Limulus* CRP from a PE column followed by a 10mM EDTA elution. The start points at which fractions 1-4 were collected are shown in red. // indicates a break in time between elutions.

None of the gradients were successful in isolating CRP molecular aggregate forms, although the 5mM gradient showed 2 peaks joined together and the 10mM gradient appeared to have a slight shoulder on its ascending limb which was not present in the isocratic 10mM ribose-5-phosphate elution (figure 3.5). These may represent molecular aggregate forms with different binding affinity. The amount of CRP isolated using the elutions of 5mM, 10mM and 30mM ribose-5-phosphate during isocratic and gradient elutions, is shown in table 3.1. Isocratic elution values were typical of values from multiple (2-3) experiments.

It appears that ribose-5-phosphate is of equal effectiveness at eluting *Limulus* CRP from a PE-agarose column at gradients of 5mM, 10mM or 30mM concentration, and using isocratic elutions of 10mM and 30mM.

Table 3.1. Table of *Limulus* CRP eluted with ribose-5-phosphate at both 10mM and 30mM concentrations in both isocratic and gradient elutions, and 5mM gradient elution. Column one shows the elution or fraction number. Column two shows the amount of CRP (mg) eluted using ligand per mg of total CRP eluted from the column. Column three shows the amount of CRP (mg) removed using 10mM EDTA per mg of total CRP eluted from the column.

Ligand	Amount of CRP (mg/mg total CRP) eluted using ligand	Amount of CRP (mg/mg total CRP) removed using EDTA
10mM ribose-5-phosphate	0.98mg/mg CRP	0.02mg/mg CRP
30mM ribose-5-phosphate	0.98mg/mg CRP	0.02mg/mg CRP
5mM ribose-5-phosphate gradient fraction 1	0.02mg/mg CRP	n/a
5mM ribose-5-phosphate gradient fraction 2	0.08mg/mg CRP	n/a
5mM ribose-5-phosphate gradient fraction 3	0.12mg/mg CRP	n/a
5mM ribose-5-phosphate gradient fraction 4	0.19mg/mg CRP	n/a
5mM ribose-5-phosphate gradient fraction 5	0.20mg/mg CRP	n/a
5mM ribose-5-phosphate gradient fraction 6	0.17mg/mg CRP	n/a
5mM ribose-5-phosphate gradient fraction 7	0.09mg/mg CRP	n/a
5mM ribose-5-phosphate gradient fraction 8	0.05mg/mg CRP	n/a
5mM ribose-5-phosphate gradient fraction 9	0.04mg/mg CRP	n/a
5mM ribose-5-phosphate gradient fraction 10	0.02mg/mg CRP	n/a
5mM ribose-5-phosphate gradient fraction 11	0.002mg/mg CRP	n/a
5mM ribose-5-phosphate gradient fraction 12	0.01mg/mg CRP	n/a
5mM ribose-5-phosphate gradient fraction 13	0.02mg/mg CRP	n/a
5mM ribose-5-phosphate gradient total	0.98mg/mg CRP	0.02mg/mg CRP
10mM ribose-5-phosphate gradient fraction 1	0.008mg/mg CRP	n/a
10mM ribose-5-phosphate gradient fraction 2	0.06mg/mg CRP	n/a
10mM ribose-5-phosphate gradient fraction 3	0.12mg/mg CRP	n/a
10mM ribose-5-phosphate gradient fraction 4	0.51mg/mg CRP	n/a
10mM ribose-5-phosphate gradient fraction 5	0.26mg/mg CRP	n/a
10mM ribose-5-phosphate gradient fraction 6	0.02mg/mg CRP	n/a
10mM ribose-5-phosphate gradient fraction 7	0.0021mg/mg CRP	n/a
10mM ribose-5-phosphate gradient total	0.98mg/mg CRP	0.02mg/mg CRP

30mM ribose-5-phosphate gradient fraction 1	0mg/mg CRP	n/a
30mM ribose-5-phosphate gradient fraction 2	0.50mg/mg CRP	n/a
30mM ribose-5-phosphate gradient fraction 3	0.455mg/mg CRP	n/a
30mM ribose-5-phosphate gradient fraction 4	0.01mg/mg CRP	n/a
30mM ribose-5-phosphate gradient fraction 5	0.0023mg/mg CRP	n/a
30mM ribose-5-phosphate gradient total	0.97mg/mg CRP	0.03mg/mg CRP

Analysis of subunit sizes of fractions of CRP from the 5mM, 10mM and 30mM ribose-5-phosphate gradients was performed using SDS-PAGE. The SDS-PAGE gels of the fractions from the 5mM ribose-5-phosphate gradient (figures 3.11 and 3.12) showed the presence of the triple bands of approximately 29kDa in size in fractions 2 to 7 and the EDTA elution. The middle sized band of the triple bands was the darkest in all the previous mentioned fractions and in whole CRP. Fraction 1 only appeared to show the presence of two bands with the lower molecular weight band absent. Fractions 8 and 9 also appear to have two bands, with the lower molecular weight band absent. Fractions 10, 11 and 13 only appeared to show the presence of one band: the middle sized band of the triplet, see in the other fractions. This is probably due to a much lower concentration of CRP present in these fractions, and darker staining of the middle sized band of the triple bands ~29kDa in size as seen in the other fractions, but it can also be said that some slight separation of CRP subunits has been shown.
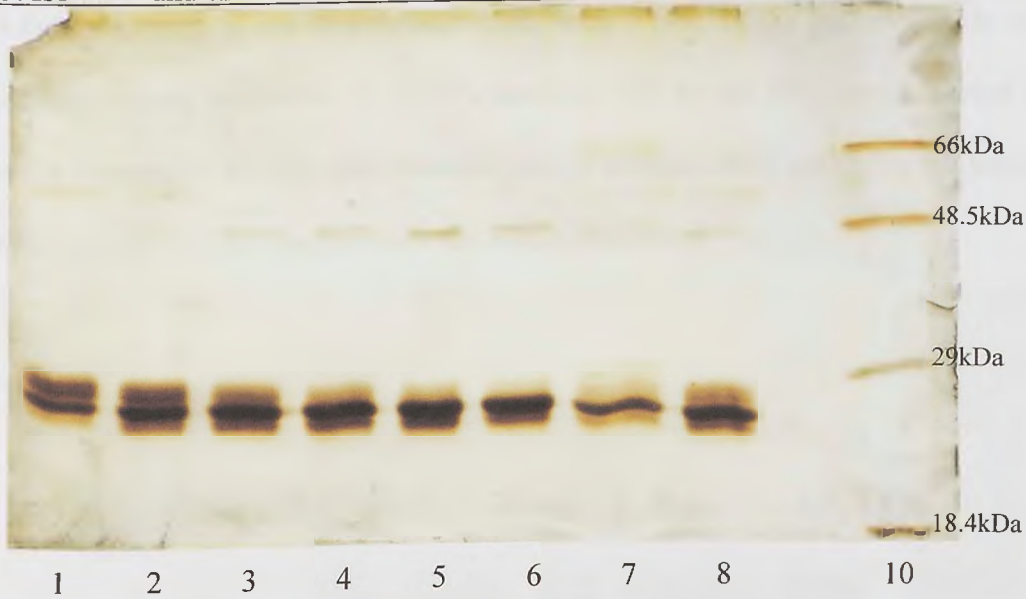
Figure 3.11 Silver stained SDS-PAGE gel of *Limulus* CRP fractions isolated using 5mM ribose-5-phosphate gradient. Lanes 1-6 show CRP from fractions 2-7 of 5mM ribose-5-phosphate gradient. Lane 7: CRP eluted using EDTA after 5mM ribose-5-phosphate gradient. Lane 8: whole CRP. Lane 10: molecular weight markers
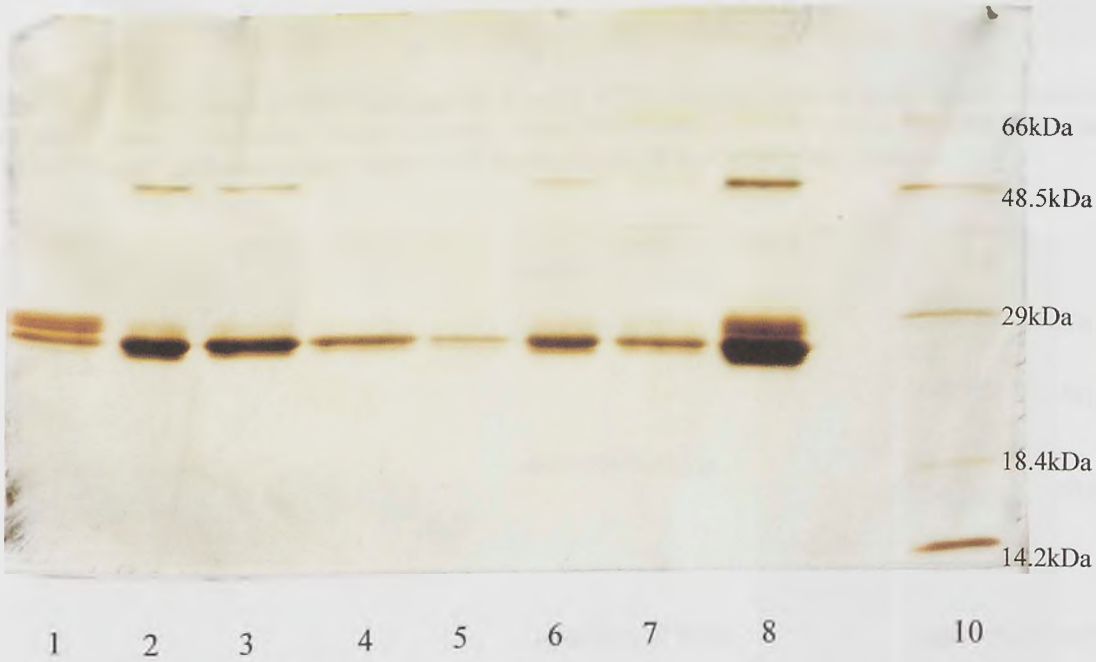
Figure 3.12 Silver stained SDS-PAGE gel of *Limulus* CRP fractions isolated using 5mM ribose-5-phosphate gradient. Lanes 1: fraction 1 from 5mM ribose-5-phosphate gradient. Lanes 2-7 shows fractions 8-13 from 5mM ribose-5-phosphate gradient. Lane 8: whole CRP. Lane 10: molecular weight markers.

The SDS-PAGE gel of fractions from the 10mM ribose-5-phosphate gradient (figure 3.13) and 30mM ribose-5-phosphate gradient (figure 3.14) showed no noticeable difference in band staining intensity and band absence between the fractions except for fraction 1 of the 10mM gradient which showed very slight staining probably due to the very small amount of protein loaded onto the gel. The EDTA fraction in both gels appeared

to have a lighter staining of the smallest molecular weight band of the triplet ~29kDa size

bands. Lighter staining occurs in the EDTA fractions due to the very small amount of

protein being present. In the four gels from all three gradients, there are higher molecular

weight bands present in CRP fractions and the EDTA fractions.
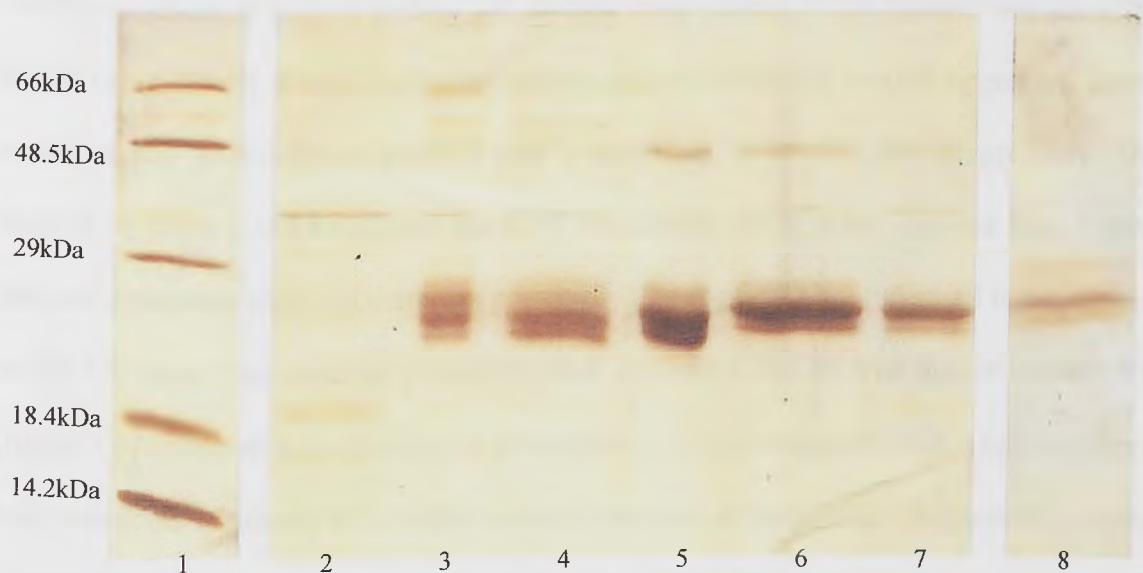
Figure 3.13. Silver stained SDS-PAGE gel of *Limulus* CRP fractions isolated using 10mM ribose-5-phosphate. Lane 1: molecular weight markers. Lanes 2-7: fractions 1-6 of 10mM ribose-5-phosphate gradient. Lane 8: CRP eluted using 10mM EDTA after 10mM ribose-5-phosphate gradient.
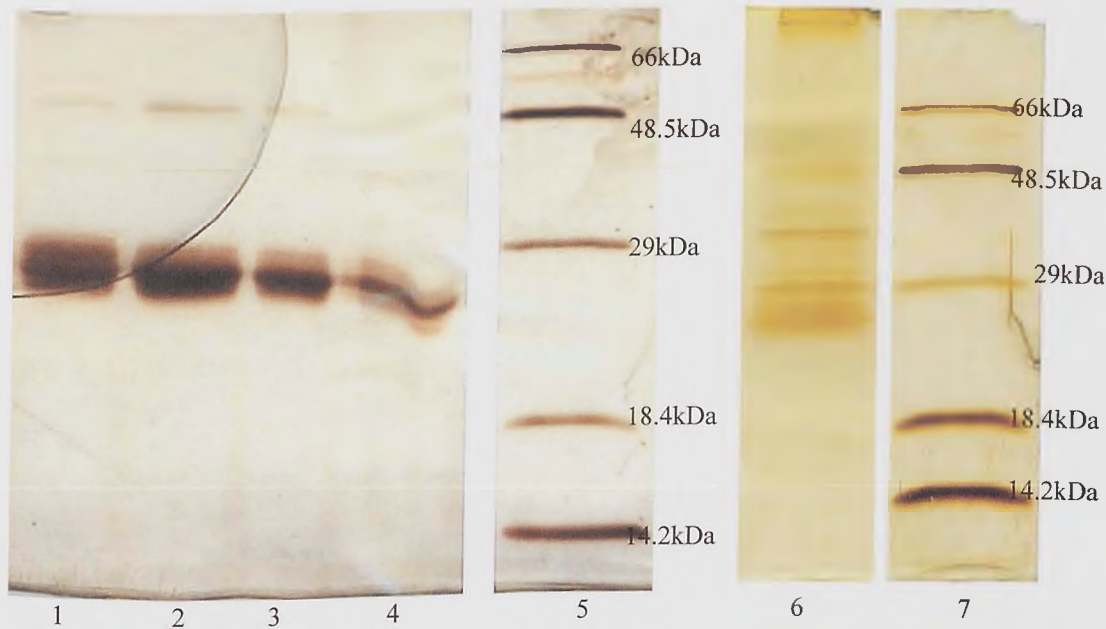
Figure 3.14. Silver stained SDS-PAGE gel of *Limulus* CRP isolated using 30mM ribose-5-phosphate. Lanes 1-4: fractions 1-4 of 30mM ribose-5-phosphate gradient. Lane 5: molecular weight markers. Lane 6: CRP eluted using 10mM EDTA after 30mM ribose-5-phosphate gradient. Lane 7: molecular weight markers. (note: lane 6 and 7 were taken from a separate gel to lanes 1-5).

110

3.3.4 FPLC analysis of ribose-5-phosphate eluted CRP

CRP isolated using 10mM ribose-5-phosphate was applied to a HIC and eluted using a 0-100% gradient of elution buffer. The elution trace (figure 3.15) showed that the CRP eluted using 10mM ribose-5-phosphate consisted of molecular weight aggregate forms corresponding to the major peaks 2 and 3 identified in whole CRP (figure 3.4). The absence of peaks 1 and 4 suggests that CRP that formed these peaks does not bind 10mM ribose-5-phosphate although there is a possibility that the peaks are too small to be noticed on the UV trace. The shoulder present on peak 2 is much smaller that that of whole CRP (figure 3.4) and there is no shoulder at all on peak 3. This indicates that CRP forming these peak shoulders is present at a lower concentration than in whole CRP suggesting a lower affinity for ribose-5-phosphate. Overall, ribose-5-phosphate appears to elute only three of the possible six molecular aggregate forms apparent in whole CRP as seen in figure 3.4.
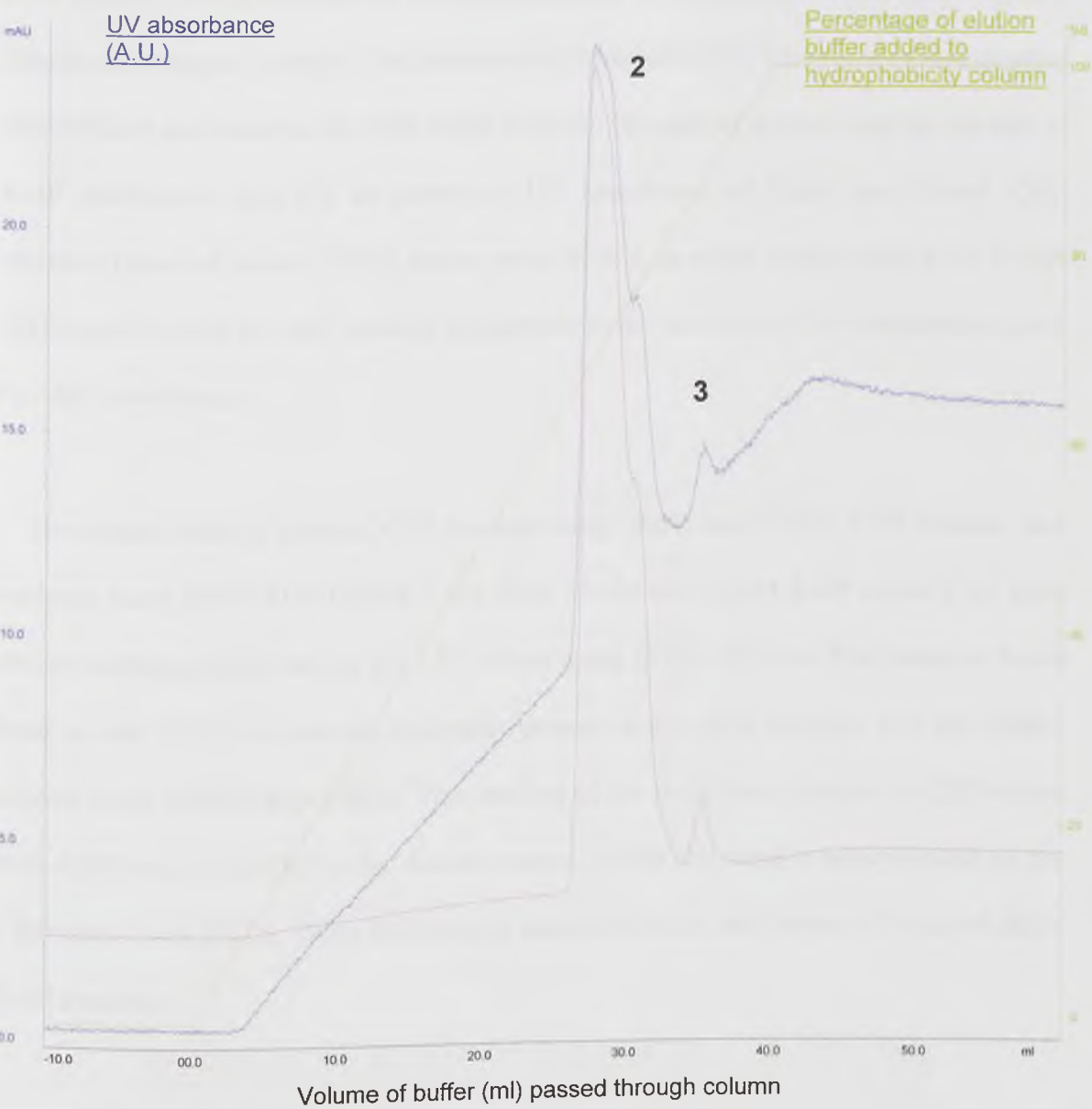
Figure 3.15. FPLC elution trace of *Limulus* CRP isolated using 10mM ribose-5-phosphate. UV absorbance at 280nm is shown in blue, UV absorbance at 215nm is shown in magenta, and the percentage of elution buffer is shown in green. The major peaks are labeled 2 and 3 according to numbering on figure 3.4.

3.3.5 Isolation of *Limulus* CRP using isocratic AMP elution

Whole *Limulus* CRP that had been isolated from *Limulus* plasma using 10mM PC was re-applied to a PE-agarose column as described in section 3.2.1. CRP was then eluted using 10mM or 30mM AMP buffer, and the remaining CRP removed from the column using 10mM EDTA buffer. AMP was chosen as a ligand because of its binding affinity to human CRP (Lee *et al.*, 2002). The elution profiles for the 10mM AMP elution and 30mM

AMP elution (not shown) showed a sharp increase in UV absorbance which then reached a plateau and dropped sharply. The increase and decrease in UV absorbance corresponded with addition and removal of AMP buffer from the PE-agarose column, and so was due to AMP interference with UV as proven by UV absorbance of 10mM and 30mM AMP solutions (data not shown). There are no peaks visible on either elution except the 10mM EDTA peaks which are very small in comparison to the increase in UV absorbance caused by AMP interference.

The subunit sizes of *Limulus* CRP isolated using 10mM and 30mM AMP elutions was analysed using SDS-PAGE (figure 3.16). Both 10mM and 30mM AMP eluted CRP gave similar staining profiles on the gel. CRP eluted using EDTA after the AMP elutions had a band at over 29kDa in size not apparently present in the AMP elutions, and had lighter stained bands smaller than 29kDa. The smallest of the three bands present in CRP eluted with AMP and whole CRP is the darkest stained, whilst this band is barely visible in the CRP eluted with EDTA. There are bands of around 48.5kDa and between 29 and 48.5kDa in all fractions.
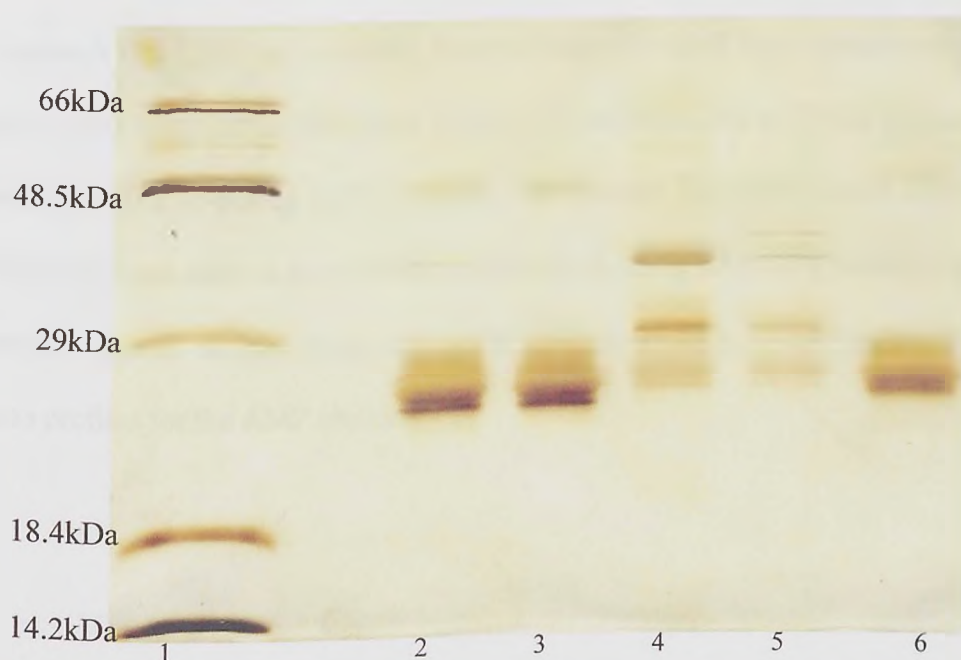
Figure 3.16 Silver stained SDS-PAGE gel showing fractions of CRP eluted using 10mM and 30mM AMP and EDTA. Lane 1: molecular weight markers. Lane 2: CRP eluted using 10mM AMP. Lane 3: CRP eluted using 30mM AMP. Lane 4: CRP eluted using EDTA after 10mM AMP elution. Lane 5: CRP eluted using EDTA after 30mM AMP elution. Lane 6: whole CRP.

3.3.6 Isolation of *Limulus* CRP using AMP gradient elution

Gradient elutions of 5mM, 10mM and 30mM AMP were performed to attempt to identify and isolate individual molecular aggregate forms of CRP with different binding affinity to AMP. The elution profiles of CRP eluted using a 5mM AMP gradient, 10mM AMP gradient, and 30mM AMP gradient (not shown), all showed a sharp increase in UV absorbance which then formed a plateau and fell sharply after the removal of AMP from the column. Despite the appearance of some shoulders on the gradient elutions of 10mM and 30mM AMP elution trace, the AMP interference with the UV reading was too great to show the correct protein elution trace, and the values of protein recovered from the column was compared to that eluted using gradients of 5mM, 10mM, and 30mM ribose-5-phosphate. Fractions were collected at similar time intervals to those collected for the ribose-5-phosphate gradient elutions. The amount of CRP eluted using 10mM and 30mM isocratic elutions and 5mM, 10mM and 30mM gradient elutions of AMP is shown in table 3.2. Isocratic elution values were typical values from multiple (2-3) experiments.

It appears that CRP has a slightly lower affinity for AMP than ribose-5-phosphate as it elutes slightly less CRP at 5mM and 10mM concentration, but at 30mM concentration it is equally effective at eluting CRP as ribose-5-phosphate. The total protein concentration of CRP eluted using AMP is very similar to that eluted using ribose-5-phosphate after dialysis to remove ligand, so confirming that AMP UV absorbance is responsible for abnormal elution profiles for the AMP elutions.

Table 3.2. Table of *Limulus* CRP eluted with AMP at both 10mM and 30mM concentrations in both isocratic and gradient elutions, and 5mM gradient elution. Column one shows the elution or fraction number. Column two shows the amount of CRP (mg) eluted using ligand per mg of total CRP eluted from the column. Column three shows the amount of CRP (mg) eluted using 10mM EDTA per mg of total CRP eluted from the column.

Ligand	Amount of CRP (mg/mg total CRP) eluted using ligand	Amount of CRP (mg/mg total CRP) removed using EDTA
10mM AMP	0.93mg/mg CRP	0.07mg/mg CRP
30mM AMP	0.99mg/mg CRP	0.01mg/mg CRP
5mM AMP gradient fraction 1	0.005mg/mg CRP	n/a
5mM AMP gradient fraction 2	0.10mg/mg CRP	n/a
5mM AMP gradient fraction 3	0.41mg/mg CRP	n/a
5mM AMP gradient fraction 4	0.25mg/mg CRP	n/a
5mM AMP gradient fraction 5	0.06mg/mg CRP	n/a
5mM AMP gradient fraction 6	0.02mg/mg CRP	n/a
5mM AMP gradient fraction 7	0.01mg/mg CRP	n/a
5mM AMP gradient fraction 8	0.01mg/mg CRP	n/a
5mM AMP gradient fraction 9	0.005mg/mg CRP	n/a
5mM AMP gradient fraction 10	0.005mg/mg CRP	n/a
5mM AMP gradient fraction 11	0.005mg/mg CRP	n/a
5mM AMP gradient fraction 12	0.005mg/mg CRP	n/a
5mM AMP gradient fraction 13	0.03mg/mg CRP	n/a
5mM AMP gradient total	0.95mg/mg CRP	0.05mg/mg CRP
10mM AMP gradient fraction 1	0.004mg/mg CRP	n/a
10mM AMP gradient fraction 2	0.11mg/mg CRP	n/a
10mM AMP gradient fraction 3	0.59mg/mg CRP	n/a
10mM AMP gradient fraction 4	0.12mg/mg CRP	n/a
10mM AMP gradient fraction 5	0.04mg/mg CRP	n/a
10mM AMP gradient fraction 6	0.02mg/mg CRP	n/a
10mM AMP gradient fraction 7	0mg/mg CRP	n/a
10mM AMP gradient total	0.89mg/mg CRP	0.11mg/mg CRP
30mM AMP gradient fraction 1	0.37mg/mg CRP	n/a
30mM AMP gradient fraction 2	0.29mg/mg CRP	n/a
30mM AMP gradient fraction 3	0.17mg/mg CRP	n/a
30mM AMP gradient fraction 4	0.07mg/mg CRP	n/a
30mM AMP gradient fraction 5	0.03mg/mg CRP	n/a
30mM AMP gradient fraction 6	0.02mg/mg CRP	n/a
30mM AMP gradient fraction 7	0.03mg/mg CRP	n/a
30mM AMP gradient fraction 8	0.02mg/mg CRP	n/a
30mM AMP gradient total	0.99mg/mg CRP	0.01mg/mg CRP

Analysis of subunit sizes of fractions of CRP from the 5mM, 10mM and 30mM AMP gradients was performed using SDS-PAGE. The SDS-PAGE gels of the fractions from the 5mM AMP gradient (figures 3.17 and 3.18) showed the presence of three bands ~29kDa in size in all fractions except fraction 1 and possibly fraction 2, which appeared to have only 2 bands. The smallest sized of the three bands appears as the darkest in all fractions except the EDTA fraction where the largest sized band of the three is darkest. Therefore, it

appears some slight separation of CRP has occurred should subunit composition represent

molecular aggregate CRP composition. In fractions 5 to 8 the middle band and largest band

of the three bands at ~29kDa appear lighter than in other fractions which may be due to gel

loading.
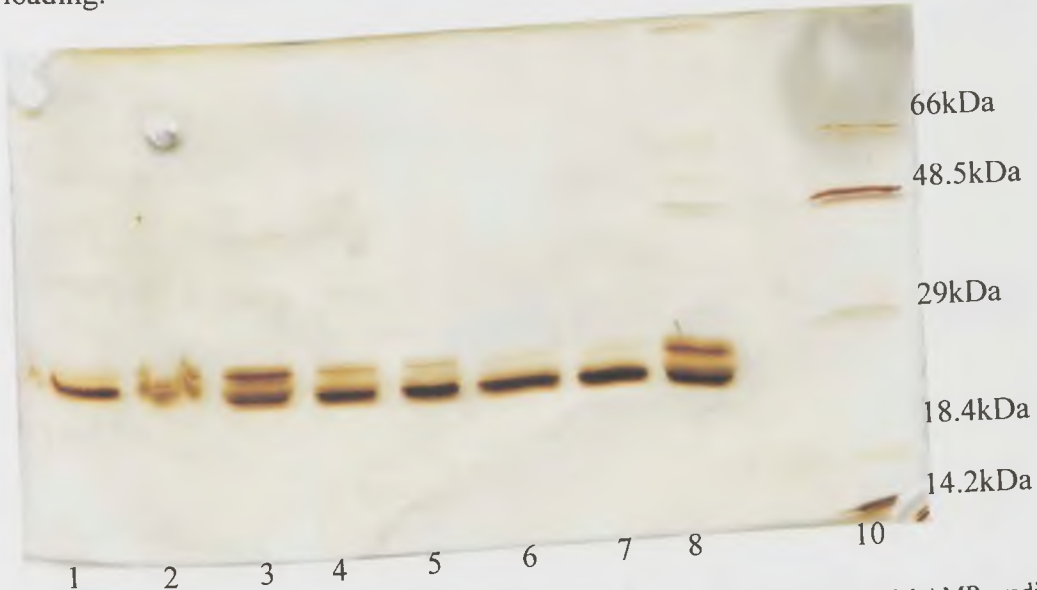
Figure 3.17. Silver stained SDS-PAGE gel of *Limulus* CRP isolated using 5mM AMP gradient. Lanes 1-7: fractions 1-7 of *Limulus* CRP from 5mM AMP gradient. Lane 8: whole CRP. Lane 10: molecular weight markers
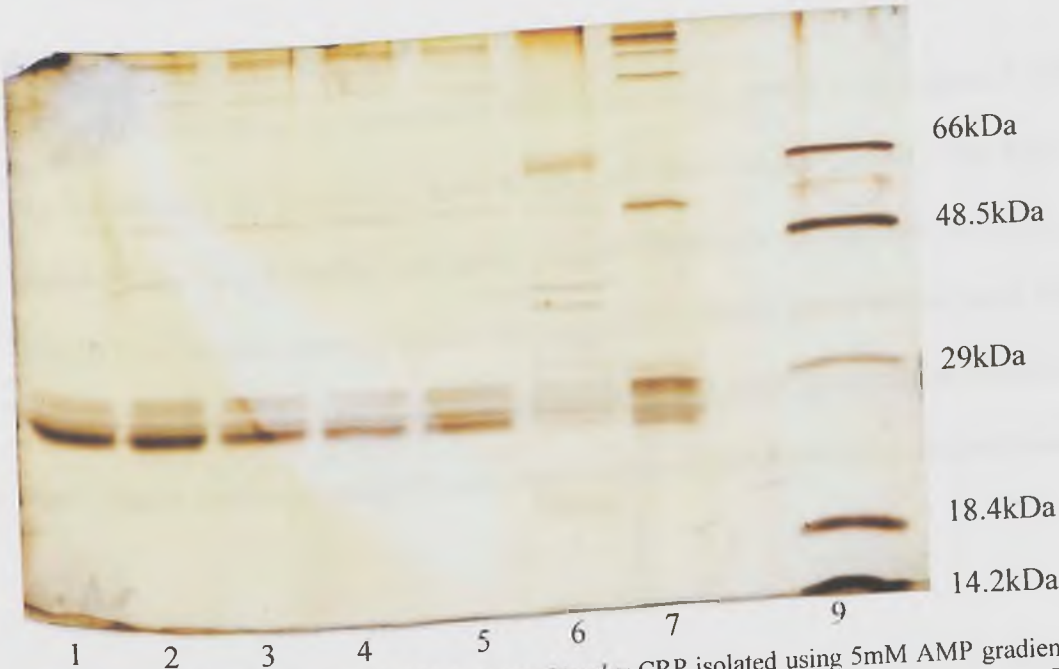
Figure 3.18. Silver stained SDS-PAGE gel of *Limulus* CRP isolated using 5mM AMP gradient. Lanes 1-6: fractions 8-13 of *Limulus* CRP from 5mM AMP gradient. Lane 7 CRP eluted using 10mM EDTA after 5mM AMP gradient. Lane 9: molecular weight markers.

The SDS-PAGE gel of fractions of CRP eluted using 10mM AMP (figure 3.19) show

darker staining of the middle of the three CRP bands ~29kDa for fractions 1 to 3. Fraction

1 showed faint bands indicating lower protein concentration probably due to its early position in the gradient. Both the smallest and middle sized bands ~29kDa were darkly stained in fractions 5 and 6. The EDTA fraction showed the middle of the three CRP bands as the darkest stained. This indicated some slight separation of CRP subunit forms.
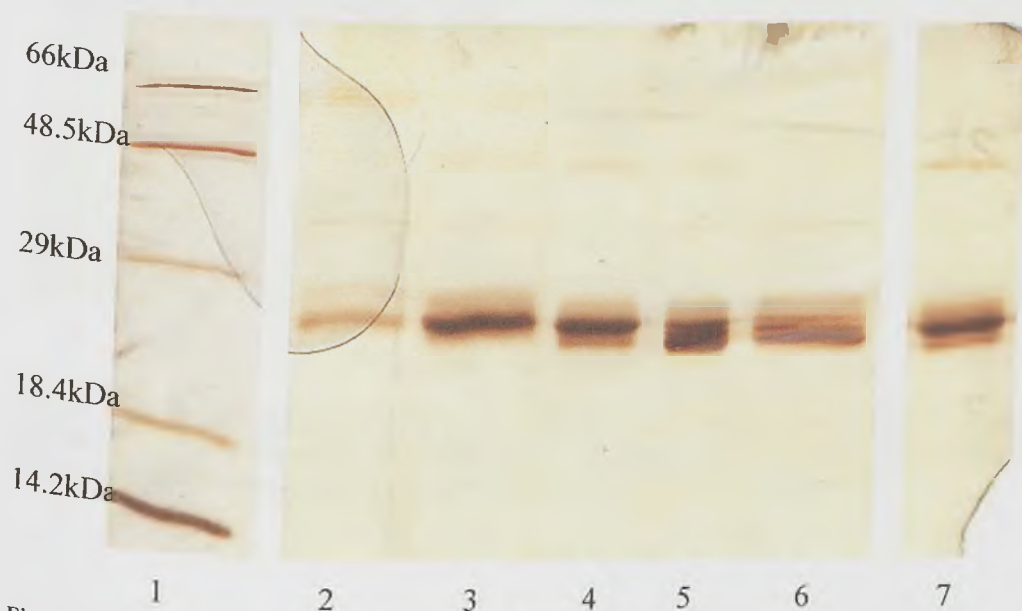
Figure 3.19. Silver stained SDS-PAGE gel of *Limulus* CRP isolated using 10mM AMP gradient. Lane 1: molecular weight markers. Lanes 2-6: fractions 1-5 of 10mM AMP gradient elution. Lane 7: CRP eluted with 10mM EDTA after 10mM AMP gradient elution.

The SDS-PAGE gel of fractions of CRP eluted using 30mM AMP (figure 3.20) showed similar staining for fractions 1 to 4, whilst fractions 5 and 6 showed only light staining probably due to gel loading and lower protein concentration than the previous fractions. The EDTA fraction showed darker staining of the middle sized of the three CRP bands ~29kDa similar to the EDTA fraction from the 10mM AMP gradient but staining was still faint. Higher molecular weight bands were seen in all gels from all three gradients.
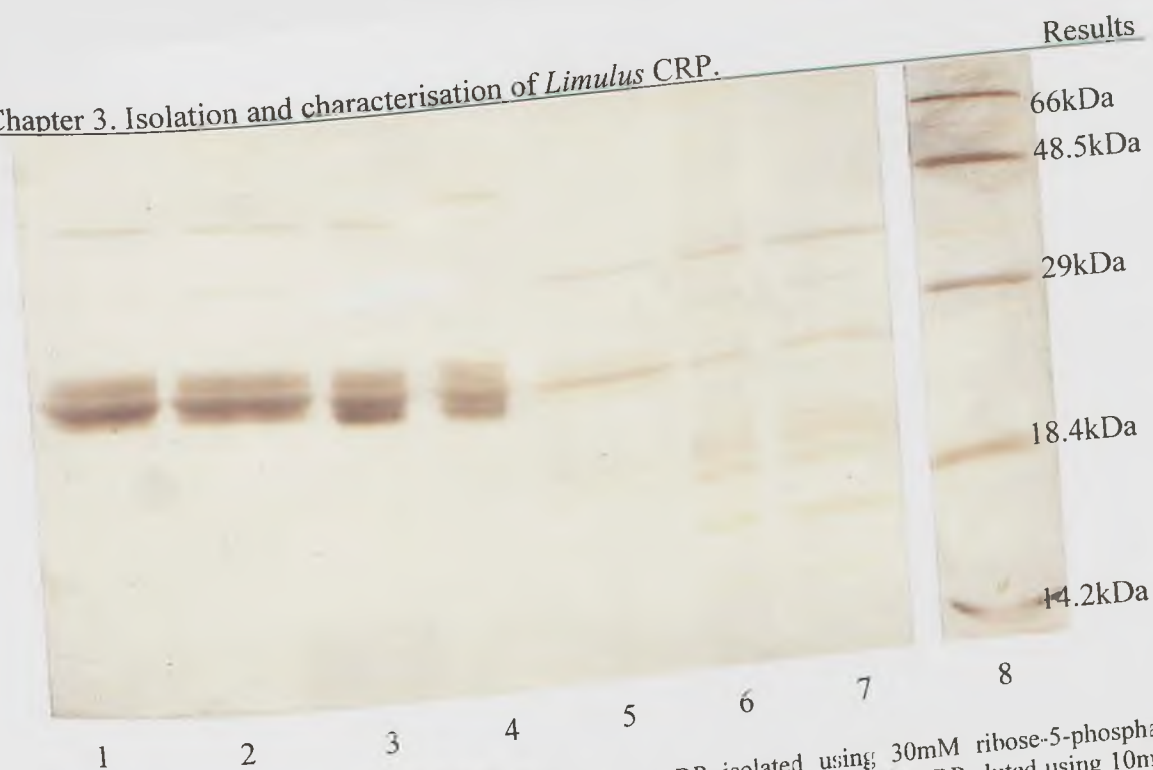
Figure 3.20. Silver stained SDS-PAGE of *Limulus* CRP isolated using 30mM ribose-5-phosphate gradient. Lane 1-6: fractions 1-6 of 30mM ribose-5-phosphate gradient. Lane 7: CRP eluted using 10mM EDTA after 30mM ribose-5-phosphate gradient. Lane 8: molecular weight markers.

3.3.7 FPLC analysis of *Limulus* CRP eluted using AMP

CRP isolated using 10mM AMP was applied to a HIC and eluted using a 0-100% gradient of elution buffer. The elution trace (figure 3.21) showed that the CRP eluted using 10mM AMP consisted of CRP of only three of the four molecular aggregate forms that were represented by peaks on the whole CRP trace (figure 3.4). Peak 4 was only small on the whole CRP trace and therefore its presence may be hard to detect on the trace for CRP eluted using AMP. The shoulder present on peak 2 appears smaller than it does in whole CRP, and the shoulder on peak 3 is absent. This indicates that CRP forming these peak shoulders is present at a lower concentration than in whole CRP suggesting a lower affinity for AMP. Overall, AMP appears to bind only four of the possible six molecular aggregate forms identified in whole CRP (figure 3.4).

118
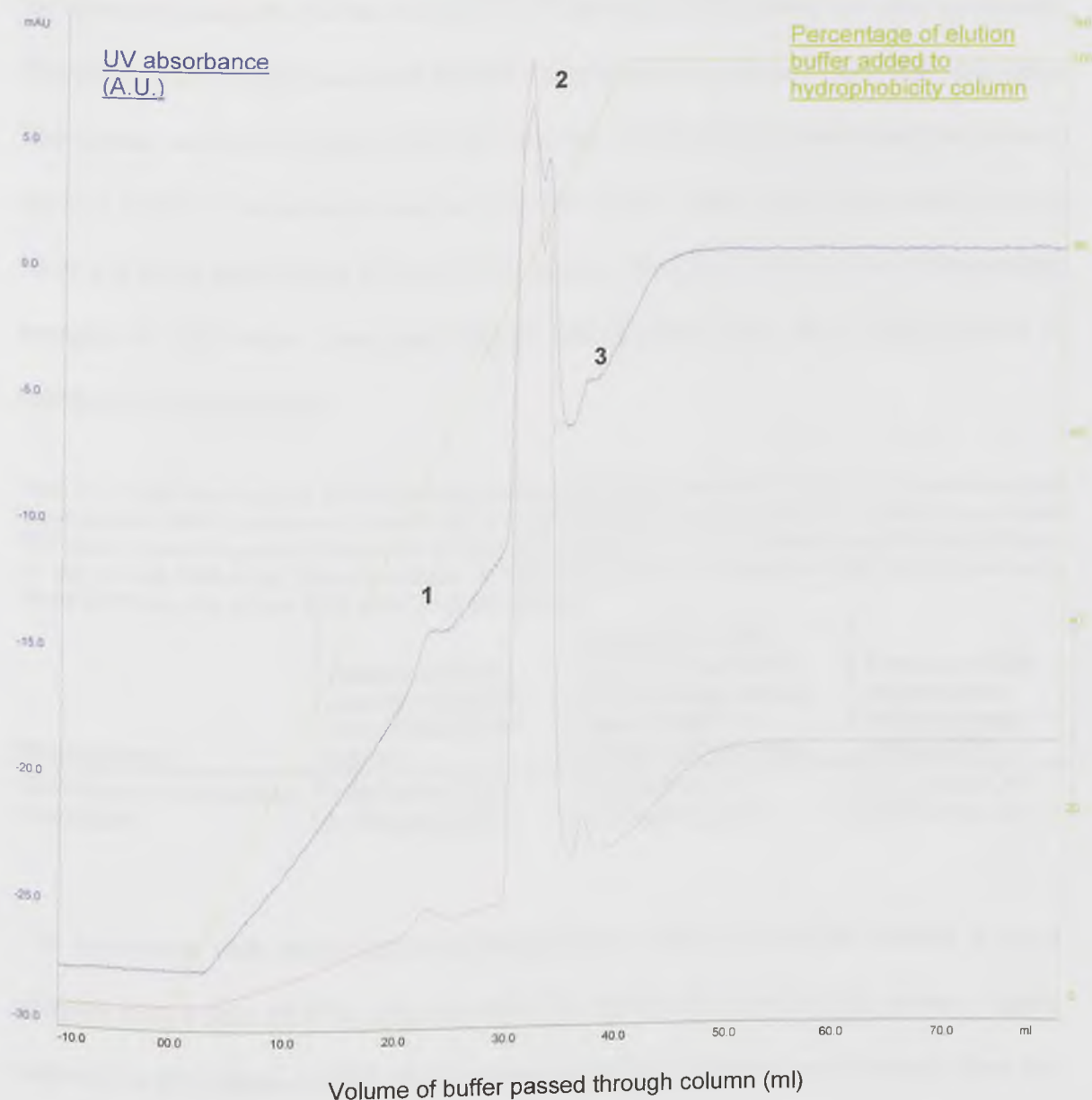
Figure 3.21. FPLC elution profile of *Limulus* CRP isolated using 10mM AMP. UV absorbance at 280nm is shown in blue, UV absorbance at 215nm is shown in magenta, and percentage of elution buffer is shown in green. The major peaks are labeled 1-3 according to figure 3.4.

3.3.8 Sequential elutions of *Limulus* CRP by 10mM ribose-5-phosphate and 10mM AMP

Sequential elutions of CRP from a PE-agarose column using 10mM ribose-5-phosphate and AMP were performed. As the elution profiles of CRP eluted using either 10mM or 30mM for each of the ligands were similar, and that 10mM concentration of ligand was sufficient to elute the majority of CRP, only 10mM elution buffers were used. The elution profile of 10mM ribose-5-phosphate followed by 10mM AMP (not shown) showed a peak

for ribose-5-phosphate elution and AMP UV absorbance interference as seen previously.

The 10mM EDTA peak was much smaller than the peak for ribose-5-phosphate and sharp.

The elution profile of 10mM AMP followed by 10mM ribose-5-phosphate (not shown)

showed AMP UV absorbance and no peak for 10mM ribose-5-phosphate elution, but a

small and sharp peak for the 10mM EDTA elution. The table 3.3 shows the corresponding

amounts of CRP eluted using each ligand. Values were taken from typical values of

multiple (2-3) experiments.

Table 3.3. Table showing total CRP eluted using ribose-5-phosphate and AMP. Column one shows the first ligand to elute CRP. Column two shows the amount of CRP (mg) eluted using the first ligand per mg of total CRP eluted from the column. Column three shows the amount of CRP (mg) eluted using the second ligand per mg of total CRP eluted from the column. Column four shows the amount of CRP (mg) eluted using 10mM EDTA per mg of total CRP eluted from the column.

Primary ligand	Amount of CRP (mg/mg total CRP) eluted using first ligand	Amount of CRP (mg/mg total CRP) eluted using second ligand (AMP or ribose-5-phosphate)	Amount of CRP (mg/mg total CRP) removed using EDTA
10mM ribose-5-phosphate	0.96mg/mg CRP	0.01mg/mg CRP	0.03mg/mg CRP
10mM AMP	0.93mg/mg CRP	0.01mg/mg CRP	0.06mg/mg CRP

In agreement with single isocratic elution data, ribose-5-phosphate seemed a more

effective eluant than AMP at concentrations of 10mM when used as the primary ligand

elution. The percentage of CRP eluted using the secondary ligand was identical when the

primary ligand was either ribose-5-phosphate or AMP which indicated that ribose-5-

phosphate and AMP eluted the same CRP as each other. The amount of CRP that

consistently ran through the column during times when no elution was proceeding was

approximately 1% of the total CRP eluted, which corresponds to the percentage of CRP

eluted with secondary ligand in table 3.3. This would mean that neither ribose-5-phosphate

nor AMP eluted any different CRP to each other. There is a possibility of error in

measurement of such small protein concentrations in the secondary elution of EDTA

elution, although the experiments were not repeated enough times to calculate errors.

SDS-PAGE analysis of CRP isolated using the sequential elutions of 10mM ribose-5-phosphate and AMP is shown in figure 3.22.

Figure 3.22. Silver-stained SDS-PAGE gel of *Limulus* CRP isolated using the sequential elutions of 10mM ribose-5-phosphate and AMP, and vice versa.
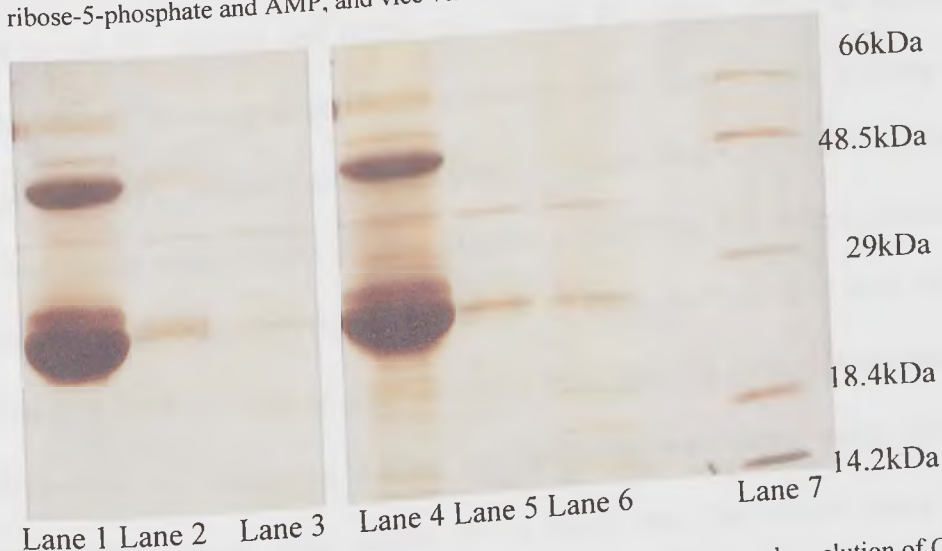
Lane 1: primary elution of CRP with 10mM ribose-5-phosphate. Lane 2: secondary elution of CRP with 10mM AMP. Lane 3: EDTA elution after secondary elution of AMP. Lane 4: primary elution of CRP with 10mM AMP. Lane 5: secondary elution of CRP with 10mM ribose-5-phosphate. Lane 6: EDTA elution after secondary elution of ribose-5-phosphate. Lane 7: molecular weight markers.

The gel shows no apparent difference in subunit size when CRP is isolated using ribose-5-phosphate or AMP as the first elution. There is a significant difference between the first and second elution which may be due to loading of the gel. However, the second elution using AMP shows two bands at approximately 28kDa, whilst the second elution using ribose-5-phosphate shows only one band at approximately 28kDa. The EDTA lane shows identical bands to the secondary elution when it is ribose-5-phosphate or AMP, and also shows additional minor bands at low molecular weights of approximately 18.4kDa. There is also a band present between 29 and 48.5kDa in size in all fractions, and a band at 48.5kDa in the primary elution lanes which may be present in the other lanes but fainter stained due to gel loading.

3.3.9 Analysis of *Limulus* CRP binding to 10mM PC

Whole *Limulus* CRP that had been isolated using a 10mM PC elution from plasma was re-applied to the PE column and eluted with 10mM PC again. This was performed to investigate CRP-PC binding as preliminary investigations showed poor binding of CRP when applied to a PC column (results not shown). Remaining CRP that was left on the column was eluted using 10mM EDTA. Despite all CRP on the column having being eluted initially with 10mM PC from plasma, not all CRP was re-eluted with 10mM PC. Only 0.89mg of CRP per 1mg of CRP eluted in total from the column was eluted using 10mM PC, whilst the remaining 0.11mg per 1mg CRP was eluted using EDTA. The proportion of CRP eluted using 10mM PC is lower than that eluted using ribose-5-phosphate (0.98mg/mg CRP) and AMP (0.93mg/mg CRP) (see tables 3.1. and 3.2). The elution profile of 10mM PC eluted CRP (figure 3.23) shows that the 10mM PC elution peak is broad and has a gradual descending limb similar to the 10mM PC peak from plasma but with two joined peaks, whilst the 10mM EDTA elution is sharp. As this elution was only performed once, it is possible that repetitions would show a higher amount of CRP-PC binding than shown here.
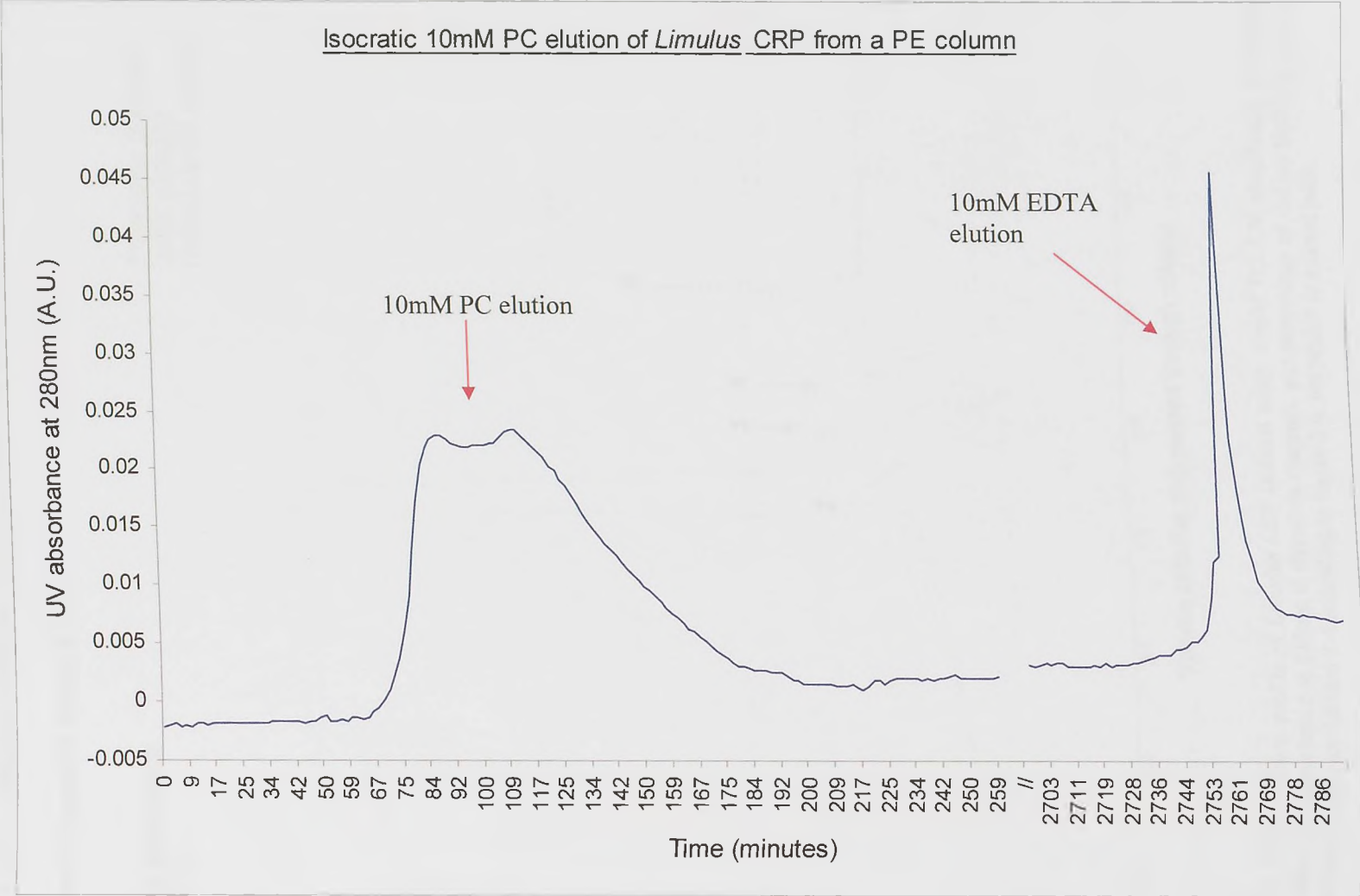
Figure 3.23 Elution profile of *Limulus* CRP eluted using sequential elutions of 10mM PC and 10mM EDTA. // *indicates a break in time between elutions.*

Limulus CRP that had been isolated using 10mM PC from CRP re-applied to the PE

column, and the remaining CRP eluted using EDTA were applied to a HIC, and eluted

using a 0-100% gradient of elution buffer as described in section 3.2.3. The elution profiles

are shown in figures 3.24 and 3.25.
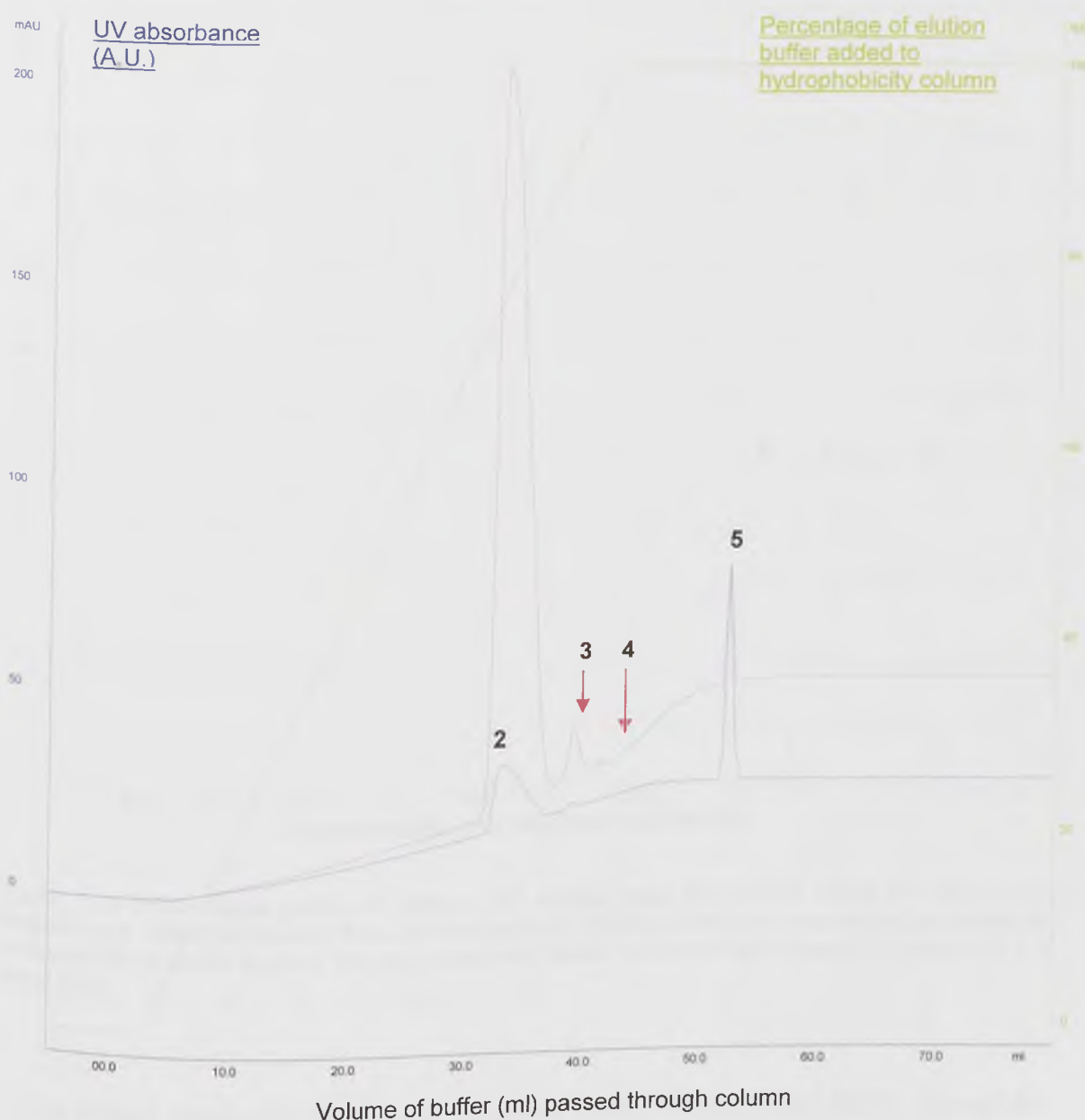
Figure 3.24. FPLC elution profile of *Limulus* CRP isolated using 10mM PC. UV absorbance at 280nm is shown in blue, UV absorbance at 215nm is shown in magenta, and percentage of elution buffer is shown in green. The major peaks are labeled 2-4 according to figure 3.4, but peak 5 is a novel peak.
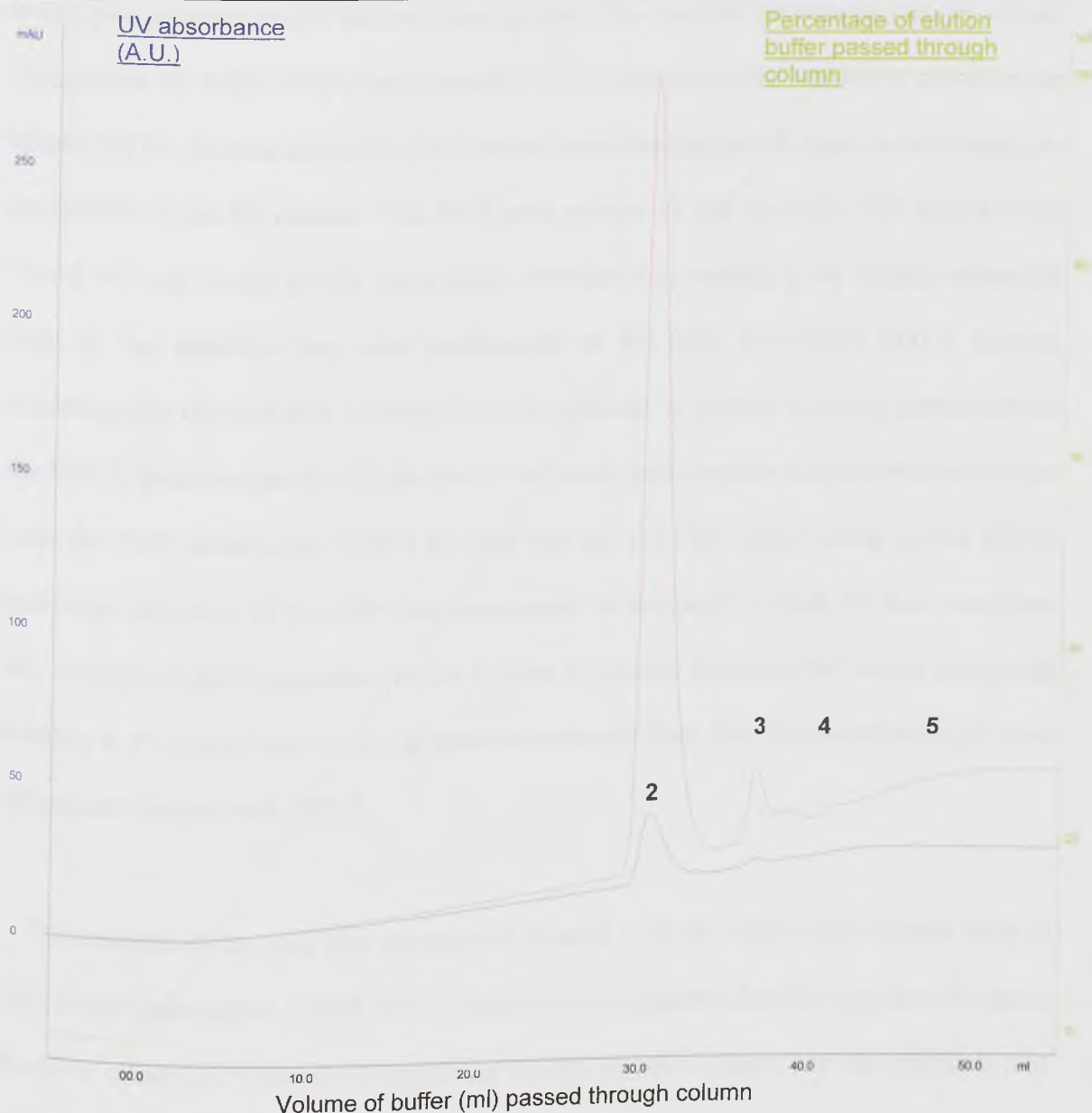
Figure 3.25. FPLC elution profile of *Limulus* CRP isolated using EDTA after 10mM PC elution. UV absorbance at 280nm is shown in blue, UV absorbance at 215nm is shown in magenta, and percentage of elution buffer is shown in green. The major peaks are labeled 2-4 according to figure 3.4, but peak 5 is a novel peak.

The elution profiles for CRP eluted using both 10mM PC and 10mM EDTA showed the presence of four major peaks, three of which were similar to the major peaks 2, 3 and 4 of whole *Limulus* CRP. Peak five does not correspond to peaks in any of the previous elution traces. It is much more pronounced on the trace for CRP eluted using 10mM PC than the CRP eluted using EDTA indicating that protein that composes this peak mostly binds 10mM PC but a small amount does not. Peak two is present on the trace for both the 10mM PC elution and the EDTA fraction indicating that some of the CRP corresponding

to this peak re-bound to PC but that some did not. The shoulder on peak two that is present on the trace for whole CRP is not present in CRP from either the 10mM PC eluted or the 10mM EDTA fraction indicating that protein corresponding to this peak is lost during re-application to the PE column. The third peak present on the trace for CRP eluted using 10mM PC and 10mM EDTA had a small shoulder that seemed to be slightly separated from it. The shoulder was more pronounced on the trace for 10mM EDTA fraction indicating that the CRP that corresponds to the shoulder is present in higher proportions in the EDTA fraction than the PC fraction. The fourth peak appears less pronounced on the trace for CRP eluted using 10mM PC than that for the CRP eluted using 10mM EDTA indicating that some of the CRP that corresponds to this peak re-binds PC but some does not. Changes in peak appearance on the elution traces may indicate CRP which has altered binding to PC which may be due to previous removal from the PE column using PC and subsequent dialysis with EDTA.

The absence of the peak that corresponds to peak 1 on the whole CRP elution trace in CRP eluted using either 10mM PC or 10mM EDTA, indicates that this protein was absent from the initial CRP that was re-applied to the PE-agarose column. It is also possible that dialysis with EDTA and re-application to a PE-agarose column changed the hydrophobicity of some CRP molecular aggregate forms, therefore showing new peaks and/or removing others seen previously e.g. peak 1.

3.3.10 Isolation of *Limulus* CRP using galacturonic acid, glucuronic acid, N-acetylglucosamine, mannose, galactose, and gluconic acid

Limulus CRP that had been isolated from plasma was re-applied to a PE column and eluted using 10mM galacturonic acid, glucuronic acid, N-acetylglucosamine, mannose, galactose, and gluconic acid. These were cheap sugars to purchase so were worth trying as ligands. N-acetylglucosamine, mannose, and galactose were sequential elutions, as were galacturonic acid and glucuronic acid. Remaining CRP on the column was eluted using 10mM EDTA. The amount of CRP eluted using each ligand and EDTA is shown in table 3.4

Table 3.4. Table showing the amount of CRP eluted using ligands shown in column one. Column two shows the amount of CRP (mg) eluted by ligand per mg of CRP eluted in total from the column. Column three shows the amount of CRP (mg) eluted by EDTA per mg of CRP eluted in total from the column

Ligand	Amount of CRP (mg/mg total CRP) eluted using ligand	Amount of CRP (mg/mg total CRP) removed using EDTA
10mM galacturonic acid	0.054mg/mg CRP	n/a
10mM glucuronic acid	0.0092mg/mg CRP	0.93mg/mg CRP
10mM N-acetylglucosamine	0.037mg/mg CRP	n/a
10mM mannose	0.0005mg/mg CRP	n/a
10mM galactose	0mg/mg CRP	0.96mg/mg CRP
10mM gluconic acid	0mg/mg CRP	1mg/mg CRP

Sequential elutions of ligands were performed because the initial ligands used did not elute a significant amount of protein as compared to background protein flow through the column which is approximately 0.01mg/mg of total eluted CRP. CRP eluted using early ligands of the sequential elutions may also have bound the later ligands, but as the amount of CRP eluted by any of the ligands was so small, sequential elutions with ligands in reverse order were not performed. Elution profiles showed that elution peaks were not visible for any of the ligands tested (data not shown), but EDTA elutions had sharp peaks similar to those of EDTA elutions for previous elution profiles.

3.3.11 Isolation of *Limulus* CRP using a PE gradient on a PE column

Whole *Limulus* CRP that had been isolated from plasma was re-applied to a PE column and eluted using a gradient of 30mM PE. Remaining CRP left on the column after the gradient elution, was removed using 10mM EDTA. The elution profile for the gradient (figure 3.26) was very similar to that of the CRP elution from whole plasma (figure 3.2) including the shoulder on the descending limb of the 30mM PE gradient elution, whilst the EDTA peak was small and sharp.

The 30mM PE gradient did not separate out any CRP peaks on the elution trace during the elution, and did not separate the shoulder from the main peak. Six fractions of the peak were taken as indicated on the elution profile, and the amounts of CRP present in each fraction including the EDTA fraction is shown in table 3.5.

Table 3.5. Table showing the amount of CRP eluted in the fractions collected from a 0-100% 30mM PE gradient. Column one shows the fraction of the elution. Column two shows the amount of CRP (mg) that the fraction contained per mg of total CRP eluted from the column.

Fraction of elution	Amount of CRP (mg/mg total CRP) eluted
Gradient fraction 1	0.003mg/mg CRP
Gradient fraction 2	0.47mg/mg CRP
Gradient fraction 3	0.211mg/mg CRP
Gradient fraction 4	0.054mg/mg CRP
Gradient fraction 5	0.023mg/mg CRP
Gradient fraction 6	0.0092mg/mg CRP
Total gradient fractions	0.775mg/mg CRP
EDTA elution	0.224mg/mg CRP
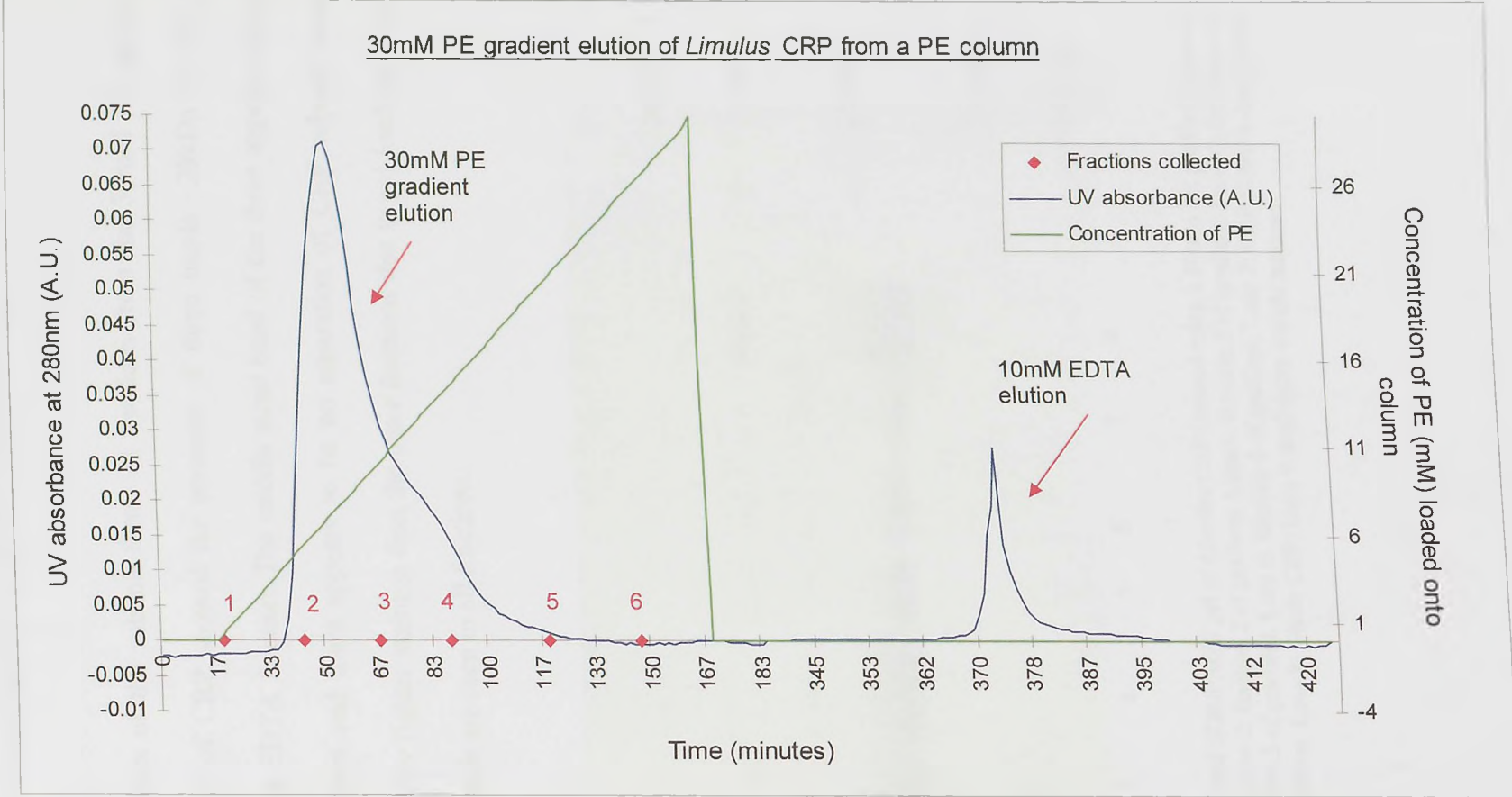
Figure 3.26 Elution profile of 30mM PE gradient elution of *Limulus* CRP, followed by 10mM EDTA. Start point at collection of fractions 1 to 6 is shown.

Overall, the PE gradient eluted less CRP than ribose-5-phosphate, AMP, and 10mM PC, indicating that CRP has a higher affinity for these ligands than PE. It could also mean a higher concentration than 30mM PE is needed to elute all CRP.

SDS-PAGE analysis of the subunit forms of fractions from the 30mM PE gradient elution (figure 3.27) of CRP showed the presence of three bands ~29kDa in size in fractions 1-6 and the EDTA fraction. The middle sized band of the three appears darkest stained in all fractions and there appears to be no separation of CRP subunit forms. Fractions 1 and 6 show lighter staining than the other fractions due to gel loading. Higher molecular weight bands are seen in all fractions.
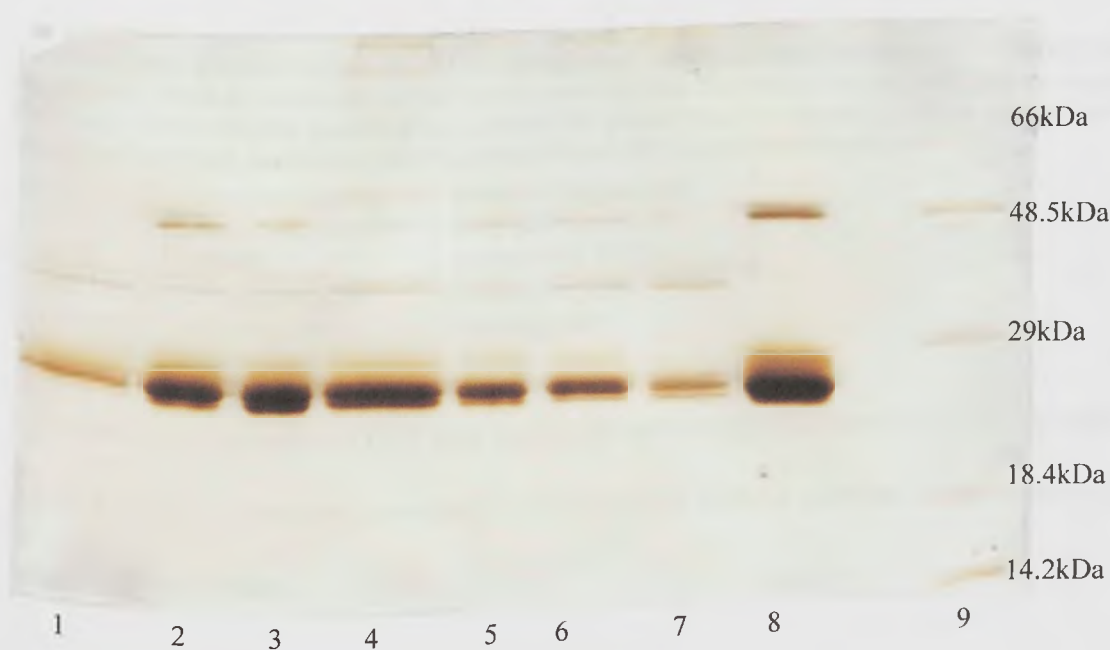
Figure 3.27. Silver stained SDS-PAGE gel of *Limulus* CRP isolated using a 30mm PE gradient. Lane 1: fraction 1 of gradient. Lane 2: fraction 2 of gradient. Lane 3: fraction 3 of gradient. Lane 4: fraction 4 of gradient. Lane 5: fraction 5 of gradient. Lane 6: fraction 6 of gradient. Lane 7: CRP eluted using EDTA after the 30mM PE gradient. Lane 8: whole CRP. Lane 9: molecular weight markers.

3.3.12 Isolation of *Limulus* CRP using sodium chloride on a PE column

Whole *Limulus* CRP isolated from *Limulus* plasma was re-applied to a PE column and eluted with 10mM and 30mM sodium chloride (NaCl). Remaining CRP was removed from the column using 10mM EDTA. Salt elutions were performed to check that CRP being eluted during 10mM and 30mM ribose-5-phosphate and AMP elutions was due to ligand and not due salt associated with the ligand. The elution profiles for both 10mM (figure 3.28) and 30mM (figure 3.29) sodium chloride showed no peak when NaCl was applied to the column, but a sharp peak with the 10mM EDTA elution. The amount of CRP eluted with 10mM and 30mM sodium chloride is shown in table 3.6.

Table 3.6 Table showing the amount of CRP eluted using 10mM and 30mM sodium chloride (NaCl) from a PE column. The first column shows the ligand concentration. The second column indicates the amount of CRP (mg) eluted with this ligand per mg of total CRP eluted from the column. The third column shows the amount of CRP (mg) eluted using EDTA per mg of total CRP eluted from the column.

Ligand	Amount (mg/mg of total CRP) of CRP eluted	Amount (mg/mg of total CRP) of CRP eluted with EDTA
10mM NaCl	0.01mg/mg CRP	0.99mg/mg CRP
30mM NaCl	0mg/mg CRP	1mg/mg CRP

Overall, NaCl elution of CRP was negligible indicating that ligands ribose-5-phosphate, AMP, PC and PE, were responsible for elution rather than an increase in salt concentration associated with ligand.
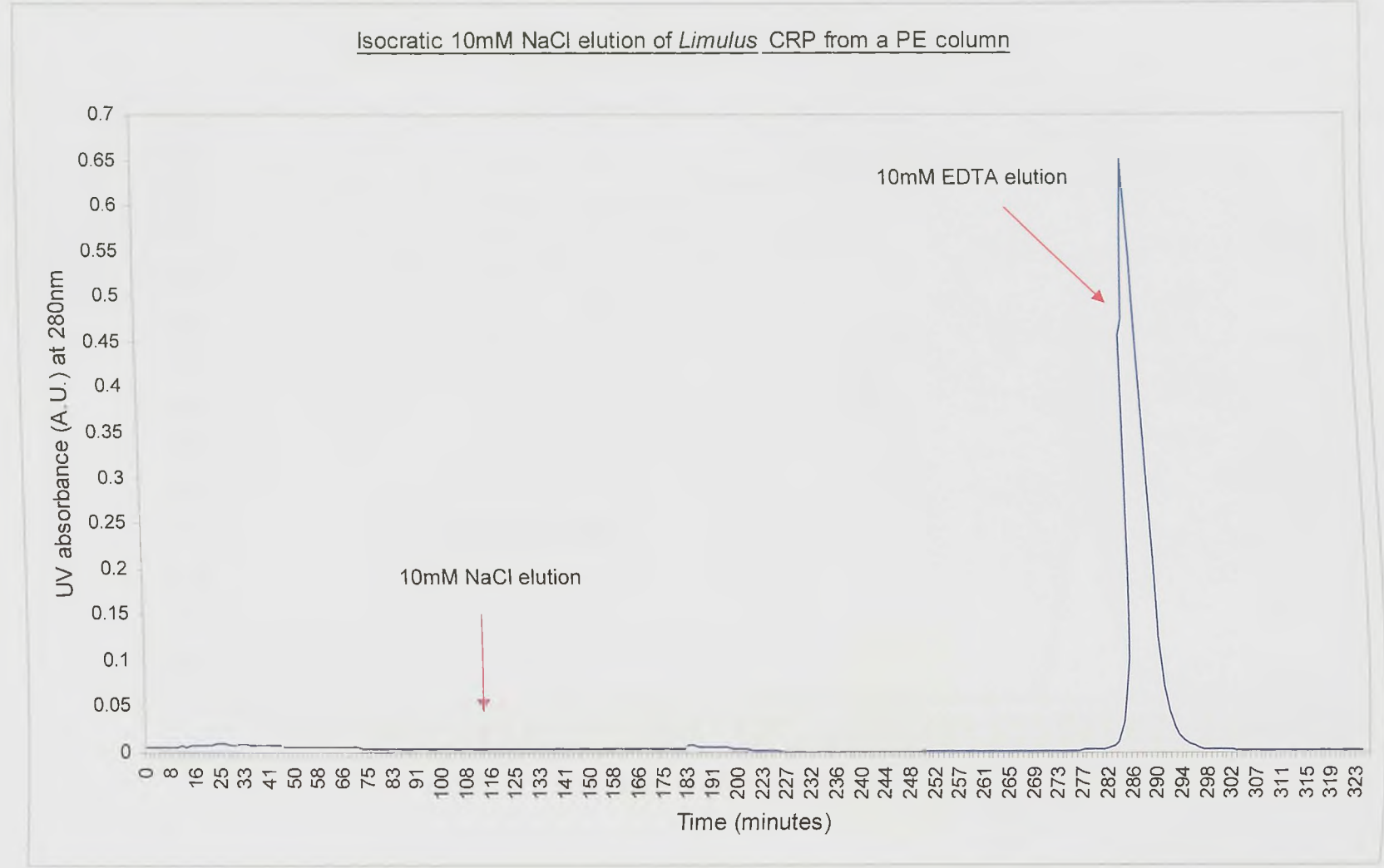
Figure 3.28. Elution profile of *Limulus* CRP with sequential elutions of *10mM NaCl and 10mM EDTA*.
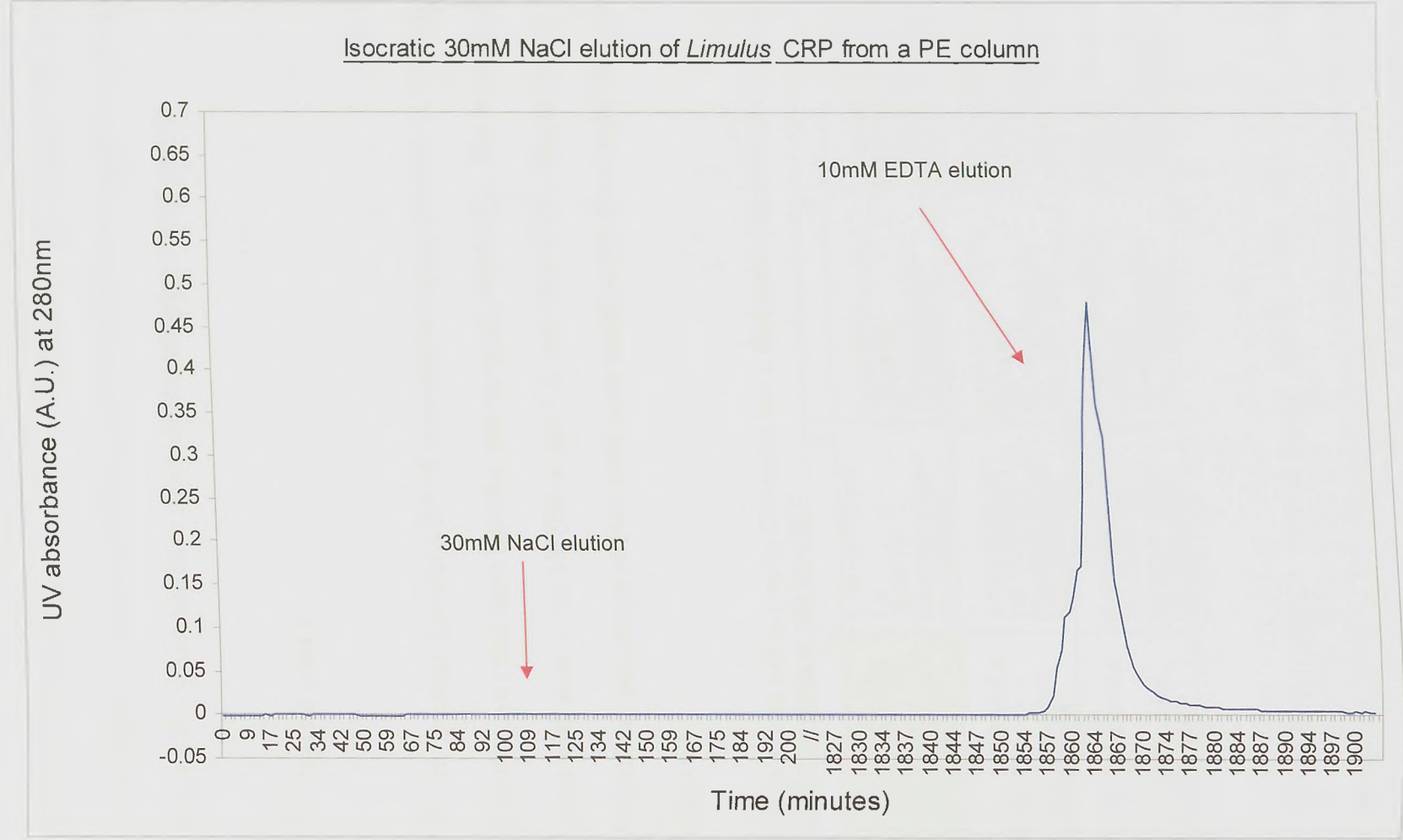
Figure 3.29. Elution profile of *Limulus* CRP with sequential elutions of 30mM NaCl and 10mM EDTA. // *indicates a break in time between elutions.*

3.3.13 Crystal trials of *Limulus* CRP isolated using 10mM PC and 10mM ribose-5-phosphate

Crystallisation trials were set up using *Limulus* CRP that had been isolated using 10mM PC or ribose-5-phosphate. Conditions for the trays were refined from earlier crystal trials involving *Limulus* CRP, which had been successful in achieving crystal growth.

Limulus CRP crystal trial 1

Trial 1 was set up using *Limulus* CRP of concentration approximately 11.7mg/ml in calcium wash buffer, isolated from *Limulus* plasma using 10mM PC. The non-variable conditions of the trial were 0.1M HEPES at pH 7.5, and 10mM PE. The percentage of ethylene glycol and the percentage of PEG 4000 were varied across the crystallisation tray as shown in table 3.7

Table 3.7 Conditions of each well in *Limulus* CRP crystal trial 1. Highlighted area shows well from which crystal has been taken for testing at SRS Daresbury.

	1	2	3	4
A	0.1M HEPES pH 7.5. 10mM PE. 4% ethylene glycol. 8% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 6% ethylene glycol. 8% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 8% ethylene glycol. 8% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 10% ethylene glycol. 8% PEG 8000.
B	0.1M HEPES pH 7.5. 10mM PE. 4% ethylene glycol. 10% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 6% ethylene glycol. 10% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 8% ethylene glycol. 10% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 10% ethylene glycol. 10% PEG 8000.
C	0.1M HEPES pH 7.5. 10mM PE. 4% ethylene glycol. 12% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 6% ethylene glycol. 12% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 8% ethylene glycol. 12% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 10% ethylene glycol. 12% PEG 8000.

Wells A1, B1 A4, B4 and C4 showed crystal growth whilst other wells showed formation of precipitate. Well A1 had the largest crystals with regular shapes (see figure 3.30) and a crystal from this well was used for testing at SRS Daresbury. Crystals that had formed in wells B1, A4, B4, and C4 were irregular in shape or thin slivers, and were not suitable for testing at SRS Daresbury.

134
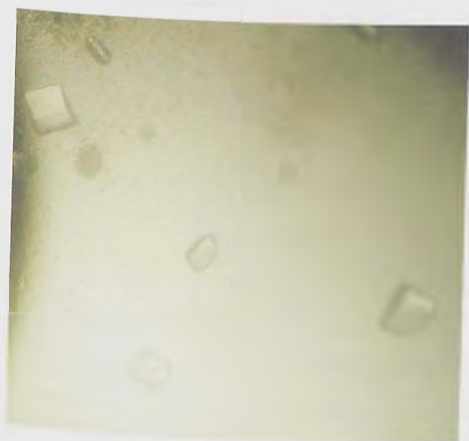
Figure 3.30 Photo of crystals grown in well A1 in *Limulus* CRP crystal trial 1

A crystal from well A1 (figure 3.30) was used to obtain diffraction data in September 2006 at Daresbury Synchrotron Radiation Source station 14.1. The crystal was cryoprotected using cryobuffers containing the same components as the mother liquor plus 10mM calcium and MPD at 5, 10, 15 and 20%. Cryobuffers were added to the sitting drop in the well at 5-10 minute intervals. 2µl of 5% MPD cryobuffer was added then 10%, then 15%, then 20% and another 20%. After cryobuffers were added, 10µl of liquid in the sitting drop was removed and replaced with 20% MPD cryobuffer. The protein crystal diffracted to 10Å.

Limulus CRP crystal trial 2

Limulus CRP crystal trial 2 was an extension of *Limulus* CRP crystal trial 1 with slightly different conditions using the same sample of *Limulus* CRP. The variable and non-variable conditions of the tray are based on the same buffer components as *Limulus* CRP crystal trial 1, except a wider range of percentages of ethylene glycol (see table 3.8).

Much crystal growth was observed for this crystal trial although in many wells the crystals formed were tiny and irregularly shaped often in clusters and therefore unsuitable for testing at SRS Daresbury. Wells B2, B4, A6, B6, and C6 had growth of larger crystals some with regular sides (see figure 3.31) whilst others had more irregular sides (see figure 3.32). The most suitable crystals from wells B4 and B6 were tested at SRS Daresbury.

Table 3.8 Conditions of each well in *Limulus* CRP crystal trial 2. Highlighted areas show wells from which a crystal has been taken for testing at SRS Daresbury.

	1	2	3		4	5	6
A	0.1M HEPES pH 7.5. 10mM PE. 3% ethylene glycol. 8% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 4% ethylene glycol. 8% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 5% ethylene glycol. 8% PEG 8000.		0.1M HEPES pH 7.5. 10mM PE. 9% ethylene glycol. 8% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 10% ethylene glycol. 8% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 11% ethylene glycol. 8% PEG 8000
B	0.1M HEPES pH 7.5. 10mM PE. 3% ethylene glycol. 10% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 4% ethylene glycol. 10% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 5% ethylene glycol. 10% PEG 8000.		0.1M HEPES pH 7.5. 10mM PE. 9% ethylene glycol. 10% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 10% ethylene glycol. 10% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 11% ethylene glycol. 10% PEG 8000.
C	0.1M HEPES pH 7.5. 10mM PE. 3% ethylene glycol. 12% PEG 8000	0.1M HEPES pH 7.5. 10mM PE. 4% ethylene glycol. 12% PEG 8000	0.1M HEPES pH 7.5. 10mM PE. 5% ethylene glycol. 12% PEG 8000		0.1M HEPES pH 7.5. 10mM PE. 9% ethylene glycol. 12% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 10% ethylene glycol. 12% PEG 8000.	0.1M HEPES pH 7.5. 10mM PE. 11% ethylene glycol. 12% PEG 8000

Figure 3.31 Photo of crystals from well B6 from *Limulus* CRP crystal trial 2.

Figure 3.32 Photo of crystals from well B4 from *Limulus* CRP crystal trial 2.

Petal shaped crystals from well B4 and regular shaped crystals from well B6 were tested for diffraction in July 2005 at Daresbury Synchrotron Radiation Source station 14.1. Cryobuffers used for both the crystals had the same components as the mother liquor plus MPD at 5, 10, 15, and 20%. Cryobuffers were added to the sitting drop in the well as mentioned previously for crystal trial 1. Diffraction snaps of the crystals were taken to determine whether they would produce good diffraction data, it was decided that they were too small for any data collection. Two crystals from B6 were tested for diffraction data collection in November 2005 at Daresbury Synchrotron Radiation Source station 14.1. The crystals within the well had already been cryoprotected from July 2005, however, no diffraction data was obtained as the crystals were not cryoprotected well which was probably due to the long period of time that they were kept in cryobuffer.

Limulus CRP crystal trial 3

Limulus CRP crystallisation trial 3 was a repeat of trial 2, but using *Limulus* CRP isolated by 10mM PC from PEG cut plasma instead of non-PEG cut plasma. This was done to see whether a possible reduction in the amount of haemocyanin affected crystal growth. The concentration of CRP was approximately 10mg/ml. As a repeat of trial 2, the variable and non-variable conditions in the crystallisation tray were exactly the same as those in *Limulus* CRP crystallisation trial 2 (see table 3.8).

Crystal growth was observed in wells B1, B2, A3, B3, and A4 but other wells showed formation of precipitate or clear protein solution. Crystals in wells B1, B2 and A4 showed clumping formation and most crystals were small and regular in shape. A small bi-pyramid was seen in well B2 (see figure 3.33) but this was too small to use for diffraction testing at SRS Daresbury. A larger regular shaped crystal that had formed in well B3 (see figure 3.34) was tested at SRS Daresbury in September 2006 at station 14.1. The cryobuffers used for this well had the same components as the mother liquor, and with MPD at 5, 10, 15 and 20%. Cryobuffers were added to the sitting drop in the well at 5-10 minute intervals as mentioned previously for crystal trial 1. The protein crystal diffracted to 12Å.

Figure 3.33 Photo showing crystals from well B2 in *Limulus* CRP crystallisation trial 3.

Figure 3.34 Photo showing a crystal from well B3 in *Limulus* CRP crystallisation trial 3.

Limulus CRP crystal trial 4

Limulus CRP crystal trial 4 used *Limulus* CRP at approximately 19mg/ml, isolated from PC eluted CRP using 10mM ribose-5-phosphate. The change in concentration of CRP was performed as very early work had produced some crystal growth with *Limulus* pentraxin at 18mg/ml. The trial was composed of two separate sets of conditions as shown in table 3.9 and 3.10. Condition 1 focused on using MES buffer instead of HEPES used in the previous trials, with a lower molecular weight PEG and calcium chloride presence. These conditions were based on trials performed previously that had shown successful crystal growth. Condition 2 focused on the conditions around well B4 in crystal trial 2, as promising crystal growth was observed at this condition.

Table 3.9 Variable and non-variable conditions of *Limulus* CRP crystal trial 4.

	Variable conditions	Non-variable conditions
Condition 1	mM CaCl$_2$ % PEG 6000	50mM MES pH 7.0
Condition 2	% ethylene glycol % PEG 8000	0.1M HEPES pH 7.5

Table 3.10 Conditions of each well of *Limulus* CRP crystal trial 4. Highlighted areas show wells from which crystal has been taken for testing at SRS Daresbury.

Condition 1				Condition 2		
	1	2	3	4	5	6
A	50mM MES pH 7. 8mM CaCl$_2$. 7% PEG 6000.	50mM MES pH 7. 10mM CaCl$_2$ 7% PEG 6000.	50mM MES pH 7. 12mM CaCl$_2$. 7% PEG 6000.	0.1M HEPES pH 7.5. 6% ethylene glycol. 8% PEG 8000	0.1M HEPES pH 7.5. 8% ethylene glycol. 8% PEG 8000	0.1M HEPES pH 7.5. 10% ethylene glycol. 8% PEG 8000
B	50mM MES pH 7. 8mM CaCl$_2$. 9% PEG 6000.	50mM MES pH 7. 10mM CaCl$_2$. 9% PEG 6000.	50mM MES pH 7. 12mM CaCl$_2$. 9% PEG 6000.	0.1M HEPES pH 7.5. 6% ethylene glycol. 10% PEG 8000	0.1M HEPES pH 7.5. 8% ethylene glycol. 10% PEG 8000	0.1M HEPES pH 7.5. 10% ethylene glycol. 10% PEG 8000
C	50mM MES pH 7. 8mM CaCl$_2$. 11% PEG 6000.	50mM MES pH 7. 10mM CaCl$_2$. 11% PEG 6000.	50mM MES pH 7. 12mM CaCl$_2$. 11% PEG 6000.	0.1M HEPES pH 7.5. 6% ethylene glycol. 12% PEG 8000	0.1M HEPES pH 7.5. 8% ethylene glycol. 12% PEG 8000	0.1M HEPES pH 7.5. 10% ethylene glycol. 12% PEG 8000

Only wells C2 and C3 had crystal growth in them for condition 1. The remaining wells had precipitate or clear protein solution. Crystals in C2 were flat and regular shaped (see figure 3.35) whilst crystals in C3 were tiny and rod shaped. For condition 2, wells B4, C5, B6, and C6 showed growth of mostly tiny crystals and
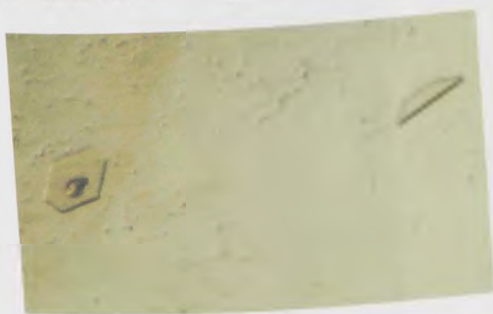
Figure 3.35 Photo of crystals in well C2 of *Limulus* CRP crystallisation trial 4.

irregular crystals. Wells with no crystal growth for condition 2 showed formation of precipitate. A crystal from well C2 (figure 3.35) was taken and used to obtain diffraction data in July 2007 at Daresbury Synchrotron Radiation Source station 10.1. Cryobuffers contained components of the mother liquor plus MPD at 5, 10, 15, and 20%. 30mM AMP was also included in the cryobuffers as a ligand to be soaked into the crystal so that if diffraction data was collected from the crystal it may include AMP bound to CRP. Cryobuffers were added to the sitting drop in the well as mentioned previously for crystal trial 1. The diffraction data showed scattered spots, and the odd spot at 3-5Å resolution but very weak.

Limulus CRP crystal trial 5

Limulus CRP crystal trial 5 was a repeat of *Limulus* crystal trial 4 but using *Limulus* CRP isolated from PC eluted CRP using 10mM ribose-5-phosphate, at a lower concentration of approximately 13.7mg/ml as the previous trials with CRP at a concentration of 19mg/mg did not show much crystal growth. The trial was repeated as one of the crystals (well C2) that grew from trial 4 gave some spots of good diffraction. The variable and non-variable conditions were the same as those for *Limulus* CRP crystallisation trial 4 (see table 3.10).

139

Figure 3.36 Photo showing a crystal from well A6 in *Limulus* CRP crystallisation tray 5.
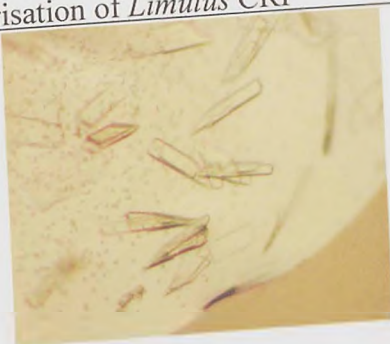
Figure 3.37 Photo showing crystals from well A5 in *Limulus* CRP crystallisation trial 5.

Figure 3.38 Photo showing crystals from well C3 in *Limulus* CRP crystallisation trial 5

Generally the wells of this tray contained either no crystals or precipitate. However, three wells A6, A5 and C3 had small regular shaped crystals and long slivers of crystals (see figures 3.36, 3.37, and 3.38). Crystals were taken from wells A6 (figure 3.36) and C3 (figure 3.38) for testing at SRS Daresbury. Crystals were taken from well A6 in both November 2006, and March 2007 at Daresbury Synchrotron Radiation Source station 14.1. The cryobuffer used for this well was composed of the mother liquor components plus MPD at 5, 10, 15, and 20%. Cryobuffers were added to the sitting drop of the well as mentioned previously for crystal trial 1. No diffraction data was obtained from either crystal. A bi-pyramid crystal was taken from well C3 and used for diffraction testing in July 2007 at Daresbury Synchrotron Radiation Source station 14.1. The cryobuffer used in this well had the same components as the mother liquor plus MPD at 5, 10, 15, and 20%. 30mM ribose-5-phosphate was included in the cryobuffer to soak into the crystals so that diffraction data was that of *Limulus* CRP ligand bound. The crystal started deteriorating after the second 20% MPD cryobuffer was added to the well, and parts diffracted to 8.2Å (see figure 3.39).
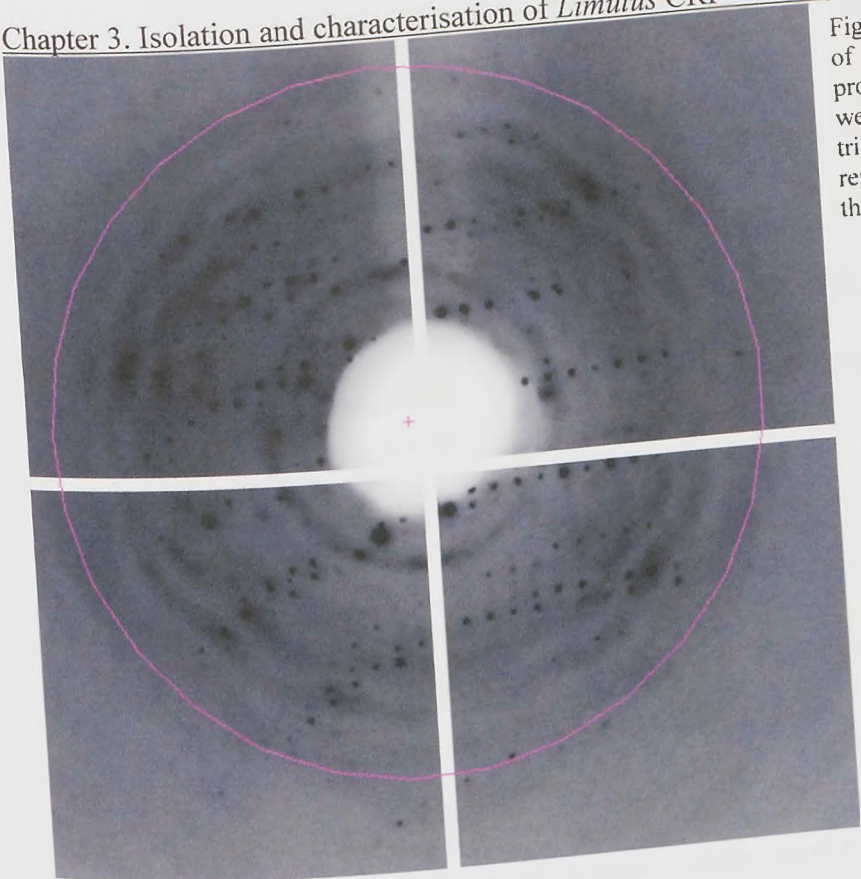
Figure 3.39 A zoomed in area of the diffraction pattern produced by a crystal from well C3 *Limulus* CRP crystal trial 5. The coloured circle represents the area at which there was diffraction to 8.2Å.

Limulus CRP crystal trial 6

This trial was a set of nine individual unique conditions to crystallise *Limulus* CRP, and was set up using *Limulus* CRP isolated from PC eluted CRP using 10mM ribose-5-phosphate at a concentration of approximately 13.7mg/ml as used previously. All nine sets of conditions contained 10mM ribose-5-phosphate in order to attempt to co-crystallise it with *Limulus* CRP. These conditions had been refined from previous trays undertaken by others. The conditions in each well are shown in table 3.11

Table 3.11 Table showing the conditions of each well in *Limulus* CRP crystal trial 6.

	1	2	3	4	5	6
A	10mM ribose-5-phosphate. 5% PEG 6000. 8% MPD. 0.1M HEPES pH 7.5.	10mM ribose-5-phosphate. 0.2M (NH₄)₂ SO₄. 30% PEG 8000	10mM ribose-5-phosphate. 0.2M (NH₄)₂ SO₄. 30% PEG 4000	10mM ribose-5-phosphate. 0.5M (NH₄)₂SO₄. 0.1M HEPES pH 7.5. 30% MPD.	10mM ribose-5-phosphate. 1.6M (NH₄)₂SO₄. 0.1M HEPES pH 7.5. 0.1M NaCl.	10mM ribose-5-phosphate. 10% PEG 8000. 8% ethylene glycol. 0.1M HEPES pH 7.5.
B	10mM ribose-5-phosphate. 0.17M Li₂SO₄. 12% PEG 2000. 0.1M Tris pH 7.0.	10mM ribose-5-phosphate. 10% PEG 6000. 10mM CaCl₂. 50mM MES pH 7.	10mM ribose-5-phosphate. 10% PEG 6000. 12mM CaCl₂. 10mM NaCl. 50mM MES pH7.0.	empty	empty	empty

141
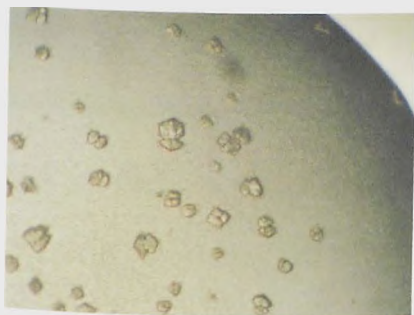
Figure 3.40 Photo of crystals in well B1 of *Limulus* CRP crystallisation trial 6

Crystals grew in well B1 only (see figure 3.40) and in no other well. None of the crystals were suitable for use in testing for diffraction at SRS Daresbury due to their small and irregular shape.

3.4 Discussion

3.4.1 Isolation and characterisation of *Limulus* CRP from plasma

Isolation of CRP from the plasma of *Limulus polyphemus* showed that CRP is in high concentration of 2.5mg CRP per ml plasma, which agrees with previous studies that showed concentrations between 1mg (Nguyen *et al.*, 1986a,b, Tennent *et al.*, 1993) and 4mg of CRP per ml of plasma (Robey and Liu 1981). CRP re-applied to a PE column was shown to have affinity to both PC and PE, although not all CRP was eluted with these ligands. *Limulus* CRP was shown to be composed of at least three subunits of different molecular weights as indicated by their presence on SDS-PAGE which agreed with previous results (Tennent *et al.*, 1993). The presence of three bands in SDS-PAGE indicates at least three different subunits perhaps which have different sequence and/or glycosylation to give them different molecular weights from one another. The presence of three bands of *Limulus* CRP during isoelectric focusing (Roche and Monsigny 1974) and at least three CRP genes (Nguyen *et al.*, 1986) indicate that there are at least three types of *Limulus* CRP subunits which may represent different molecular aggregate forms similar to those of *Tachypleus tridentatus* CRP like proteins (Iwaki *et al.*, 1999). Each CRP molecular aggregate form may have different binding properties like those from *Tachypleus tridentatus*, suggesting different ligands/targets, but may or may not be different in function.

The suggestion that there are multiple isoforms of *Limulus* CRP molecular aggregate, is further supported by surface hydrophobicity analysis. Interaction with a hydrophobic interaction column depends on surface hydrophobicity which can be affected by not only sequence and quaternary structure of protein, but also by post-translational modification such as glycosylation of the protein, and the extent to which this occurs. Analysis of hydrophobicity interactions of whole CRP indicated four major molecular aggregate forms with different surface hydrophobicity shown as peaks on the elution trace, with the possibility of a further two molecular aggregate forms shown as shoulders on these peaks.

3.4.2 *Limulus* CRP elution using ribose-5-phosphate and AMP

Limulus CRP was eluted by both ribose-5-phosphate and AMP from a PE column that it had bound calcium dependently inferring that either CRP binds to both ribose-5-phosphate and AMP, or that these chemicals affect CRP-PE interaction thereby displacing CRP from the PE-agarose column. Binding affinity of CRP appeared slightly higher for ribose-5-phosphate than AMP according to the proportion of CRP eluted using isocratic 10mM and 30mM elutions of both ligands compared to the proportion eluted with EDTA, which may be due to differences in structure between these ligands affecting access to the binding site (figure 3.41). AMP has a much larger side chain than ribose-5-phosphate which may make it more difficult to fit in the CRP binding pocket.

10mM concentrations for each ligand were sufficient to elute the majority of CRP indicating a high affinity of CRP for these ligands. An increase to 30mM concentration did not greatly affect the proportion of binding of CRP to either ribose-5-phosphate or AMP, but did decrease the elution times. Some CRP remained unbound even after 30mM ligand concentration indicating that some CRP did not have affinity for either ligand up to 30mM concentrations.

Figure 3.41. Structures of ligands ribose-5-phosphate and AMP (adapted from www.bmrb.wisc.edu).

Ribose-5-phosphate

Adenosine-5-monophosphate
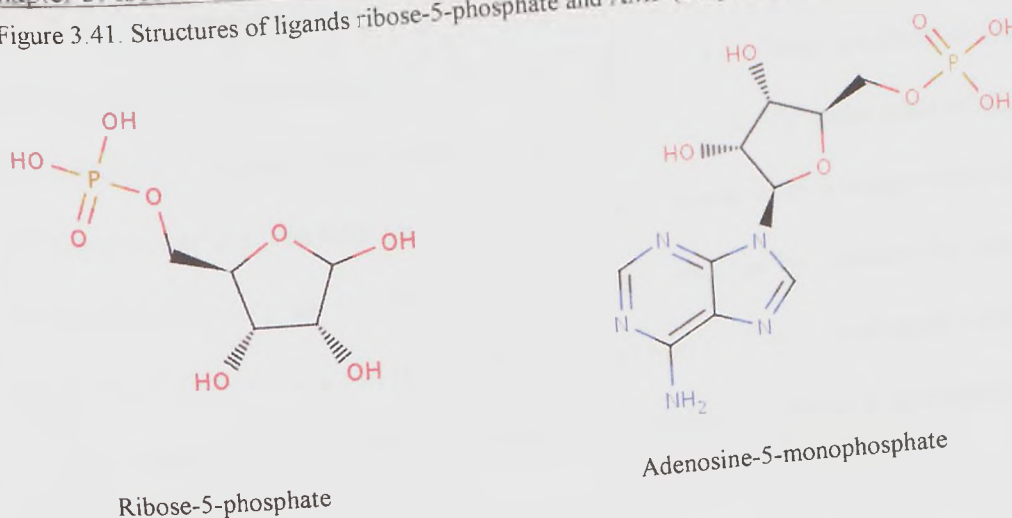
Gradient elutions of ribose-5-phosphate and AMP failed to isolate CRP isoforms as indicated by the elution profiles and SDS-PAGE analysis. It was apparent from the gradient elutions of ribose-5-phosphate that the higher the concentration of ligand, the quicker the protein was eluted from the column. This was not apparent on the elution traces using AMP because the ligand interfered with UV absorbance. The concentration of ribose-5-phosphate present on the column when the elution trace peaked indicated that the peak for 30mM ribose-5-phosphate gradient was at approximately 9mM concentration, whilst that for 10mM ribose-5-phosphate gradient was at approximately 5.5mM concentration, and that of the 5mM ribose-5-phosphate gradient was at approximately 2.5mM. It could be possible that if the gradient concentration was lowered to 2.5mM, the elution peak would occur below 2.5mM. It was not possible to predict where the elution peaks were in the AMP gradient elutions due to the absorbance by the ligand.

Therefore, the gradient elution using ribose-5-phosphate does not have to be complete at the highest concentration (i.e. 5mM, 10mM, 30mM) to show the greatest amount of elution of CRP from the PE-column. The elution peaks indicate where the protein is being eluted maximally from the column and then the elution trace drops when the amount of protein being eluted drops, in this case when the amount of CRP bound to the column has mostly been eluted. Therefore, during the 30mM ribose-5-phosphate gradient the concentration of

144

ligand used to elute the most CRP was 9mM, and during the 10mM ribose-5-phosphate gradient it was 5.5mM, and during the 5mM ribose-5-phosphate gradient it was 2.5mM. The reduction in ligand concentration is probably related to the time span of the gradient, and it is possible that running the 30mM or 10mM gradient over a longer time span should see the elution trace peak at 2.5mM or maybe even less. So, the greater the availability of ligand, the quicker the elution trace peaks. The presence of individual peaks indicates CRP molecular aggregate forms have uniform binding affinity for ribose-5-phosphate.

However, the SDS-PAGE analysis of isocratic and gradient elutions of ribose-5-phosphate and AMP appeared to show a difference in CRP subunit size compared to the CRP eluted using EDTA after them. The largest sized band of the CRP triplet bands (~29kDa) in the EDTA fraction is much more pronounced than in the fractions of CRP eluted using isocratic elutions of ribose-5-phosphate and AMP. This indicates a higher concentration of protein composing this band in the EDTA fraction and a lower concentration in the ribose-5-phosphate or AMP eluted fraction. This shows some specificity of *Limulus* CRP molecular aggregate forms that may be composed of same sized subunits, towards ribose-5-phosphate and AMP and therefore some partial separation of *Limulus* CRP. However, as the amount of protein eluted with EDTA was so small in comparison to the amount eluted with ribose-5-phosphate and AMP, the darker staining of a single band in the EDTA fraction does not represent much CRP at all and may be an anomaly of a highly concentrated protein sample. There is a possibility that CRP subunits may aggregate with like subunits creating molecular aggregate forms with specific affinity, some with specificity for ribose-5-phosphate or AMP and some with none. This would indicate that each *Limulus* CRP molecular aggregate form was coded for by individual genes which perhaps could be switched on according to requirements. There is also a possibility that individual CRP genes are switched on according to infection as suggested for CrOctin, a SAP-like protein present in *Carcinoscorpius rotundicauda* (Li *et al.*, 2007).

Sequential elution of CRP using AMP after ribose-5-phosphate appears to elute no more CRP than elution using ribose-5-phosphate after AMP. The different proportion of CRP that elutes with AMP compared to ribose-5-phosphate in isocratic elutions is so small that there may be no apparent additional binding in sequential elutions. SDS-PAGE analysis of CRP subunit sizes from sequential elutions does not show that AMP and ribose-5-phosphate elute different forms to each other. However, analysis using surface hydrophobicity of CRP molecular aggregate forms isolated using isocratic 10mM ribose-5-phosphate and AMP elutions shows otherwise. This difference may be due to a comparison of subunits against molecular aggregate forms, contamination in the CRP samples thereby affecting surface hydrophobicity studies, or be due to final CRP concentration measurement errors during sequential elutions.

In human CRP, studies have shown that CRP-PC binding is inhibited mostly by PC then PE, ribose-5-phosphate and lastly AMP (Lee *et al.*, 2002). Therefore, if *Limulus* CRP binds the same ligands with similar affinity to human CRP, then ribose-5-phosphate should elute more *Limulus* CRP than AMP. This, however, does not appear to be the case for *Limulus* CRP as both ribose-5-phosphate and AMP elute similar proportions of CRP from a PE agarose column.

3.4.3 Analysis of CRP molecular aggregate forms by surface hydrophobicity

Analysis of the surface hydrophobicity of CRP eluted using ribose-5-phosphate and AMP, indicated that CRP eluted with ribose-5-phosphate contained two of the four major CRP molecular aggregate forms shown as peaks on the elution trace in whole CRP, whilst CRP isolated using AMP contained three of the four major CRP molecular aggregate forms. CRP composing major peaks 2 and 3 on the elution trace for whole *Limulus* CRP are eluted using ribose-5-phosphate and AMP, but CRP composing peak 1 only appears to be eluted using AMP. As peak 1 on the elution trace of whole CRP is so small in

146

comparison to peaks 2 and 3, the proportion of protein corresponding to it is also small. This may represent a trace impurity like haemocyanin or α2-macroglobulin, or perhaps another protein present in the haemolymph of the horseshoe crab. It is also possible that this small peak represents Limulin, a CRP form at less than 1% of total pentraxin that binds sialic acid. However, quantification of the protein that composes each peak has not been performed, but is currently underway.

Fewer CRP molecular aggregate forms present in CRP eluted using ribose-5-phosphate and AMP, compared to whole CRP, may be representative of the incomplete elution of *Limulus* CRP by ribose-5-phosphate and AMP from a PE-agarose column. The CRP peaks missing on the elution profile for CRP eluted using ribose-5-phosphate and AMP compared to those present in whole CRP are small in comparison with the other peaks present, which could correspond to the small amount of CRP not eluted using ribose-5-phosphate or AMP. CRP eluted using AMP showed one more peak on the elution trace than CRP eluted by ribose-5-phosphate, which indicated that AMP and ribose-5-phosphate eluted different CRP molecular aggregate forms. However, the difference is only slight and the absence of these small peaks on the elution trace compared to whole CRP may result from absence of contaminating protein removed when CRP was re-applied to the column and not represent differences in CRP molecular aggregate forms binding to ribose-5-phosphate and AMP.

As the surface hydrophobicity analysis was performed only once for each CRP fraction, it is possible that the smallest of the peaks identified on the trace may not be continually present if the experiment was repeated several times. This would suggest that the peaks were representative of trace contamination of the CRP sample. However, reproduction of the surface hydrophobicity analysis on whole *Limulus* CRP at higher concentrations still showed the presence of all four major peaks which indicates that these peaks were not representative of trace contamination, but actually *Limulus* CRP.

3.4.4 *Limulus* CRP elution using PC and PE

Despite CRP being initially eluted with 10mM PC from plasma, not all CRP re-applied to a PE-column was eluted with 10mM PC. The proportion of CRP that was eluted using PC was less than that eluted using ribose-5-phosphate and AMP. This was also true for the 30mM PE gradient which eluted less CRP than PC, ribose-5-phosphate and AMP. Analysis of surface hydrophobicity of CRP molecular aggregate forms indicated the presence of major peaks 2, 3 and 4 from whole CRP in both the CRP eluted using PC and the following fraction of CRP eluted using EDTA. These peaks were similar indicating that PC did not specifically elute a single molecular aggregate form of CRP.

There was a fifth peak on the surface hydrophobicity elution traces that was not present on the elution traces of whole CRP, or CRP eluted using ribose-5-phosphate or AMP which is considerably larger on the trace for CRP eluted using PC than the following EDTA eluted fraction. The presence of CRP corresponding to a fifth peak and the absence of CRP corresponding to the first peak identified in whole CRP in both CRP eluted using PC and EDTA, indicates that the initial CRP that was eluted from the plasma was different in CRP molecular aggregate composition to that used for the 10mM PC elution. Pooled plasma was used for both isolations, but from two different dates: plasma collected in 2002 was used for initial CRP isolation, whilst plasma collected in 2005 was used to provide CRP for the 10mM PC elution of re-applied CRP. There is also the possibility that peak 1 and peak 5 correspond to contamination within the CRP fractions and that there are only three major peaks which may correspond to the three CRP sequences already found (Nguyen *et al.*, 1986b).

There is a possibility that elution of CRP using PC and subsequent dialysis using EDTA may have affected the PC binding site and therefore interaction with PC which may explain the reduced affinity of CRP for PC. This does not appear to occur to a large extent

with ribose-5-phosphate or AMP which have near total elution of CRP, unless the minority of CRP that is not eluted by them has structurally changed to not allow binding. This suggests a difference in possible binding area of CRP to PC compared to ribose-5-phosphate and AMP. It is possible that ribose-5-phosphate and AMP do not bind at the PC/PE binding site but near it, thereby displacing CRP from the PE column.

It is also possible that initial elution of CRP from the plasma using PC and dialysis with EDTA has affected CRP interaction with PE-agarose so that it binds more avidly than with PC, but does not affect ribose-5-phosphate or AMP binding. PE elution of CRP from a PE-agarose column gives a lower proportion of CRP eluted than using PC, suggesting that if CRP has higher affinity for PE, it is on the solid-phase not the mobile-phase. PC and PE conjugated to agarose do not differ structurally except for the spacer attached to the amine group of solid-phase PE. This may be related to a variation in *Limulus* CRP forms binding to PE, similar to the *Tachypleus* CRP forms of which tCRP1 and tCRP2 bind PE-agarose, but tCRP3 does not bind PE-agarose or PE-BSA, but its precipitation reaction with fetuin is inhibited by free PE (Iwaki *et al.*, 1999). This may be due to modification of the amino group of PE when attached to BSA or agarose.

3.4.5 *Limulus* CRP elution using galacturonic acid, glucuronic acid, N-acetylglucosamine, mannose, galactose, and gluconic acid

Negligible CRP was eluted with the sugars glucuronic acid and mannose, and no CRP seemed to be eluted with galactose and gluconic acid. CRP eluted using galacturonic acid and N-acetylglucosamine was approximately 5% and 4% of total CRP eluted. This amount is higher than background CRP run-off through column at approximately 1% and higher than the amount that would be eluted with Limulin which is less than 1% of total pentraxin but binds N-acetylglucosamine (Roche and Monsigny 1974). Therefore, *Limulus* CRP appears to have very slight affinity for ligands galacturonic acid and N-acetylglucosamine,

149

although this could be due to trace contaminants in the *Limulus* CRP fraction such as lectins. The low affinity of CRP to these ligands could be due to the absence of a phosphate group or an equivalently charged group, to interact with the calcium ions like PC and PE do.

Human CRP binds to carbohydrates with terminal galactosyl residues and those phosphorylated at carbon position 6 (Kottgen *et al.*, 1992; Culley *et al.*, 2000; Lee and Lee 2003). Phosphorylation was also shown to be important using inhibition studies on native-CRP binding (Lee *et al.*, 2002), and neo-CRP binding (Lee and Lee 2003). This indicates that a phosphorylated group on the ligand is required for binding which none of the above ligands possess.

The low affinity exhibited by these ligands meant that they were not further explored as CRP ligands.

3.4.6 Crystallisation trials of *Limulus* CRP

Crystallisation trials of whole *Limulus* CRP and CRP eluted using ribose-5-phosphate, both showed crystal growth, but crystals showed poor diffraction. This may be due to a number of different isoforms of CRP which affect crystal packing, or it may be due to trace contaminants or general poor diffraction. Surface hydrophobicity analysis of whole *Limulus* CRP showed at least four molecular aggregate forms of CRP in whole CRP, and at least two molecular aggregate forms in CRP eluted using ribose-5-phosphate. A reduction in the number of isoforms of CRP used for crystallisation trials may produce crystals with higher quality diffraction. So far this has not been possible as the ligands used isolated a large majority of CRP, and therefore contain many CRP molecular aggregate forms as shown by surface hydrophobicity analysis.

150

3.4.7 Overall conclusions

Limulus CRP is composed of at least three subunits as indicated by genetic analysis (Nguyen *et al.*, 1986b) and SDS-PAGE analysis. Surface hydrophobicity analysis indicates the presence of 4-6 molecular aggregate forms of *Limulus* CRP, although contamination in the form of other plasma proteins may be present. Ribose-5-phosphate and AMP elute a majority of CRP from a PE column, but the amount is not vastly different between them. Surface hydrophobicity analysis indicates that more CRP molecular aggregate forms (3 forms) are present in CRP eluted using AMP than in that eluted using ribose-5-phosphate (2 forms). Elution of CRP using ribose-5-phosphate and AMP does not significantly reduce the number of CRP molecular aggregate forms present as shown by surface hydrophobicity analysis, which may explain why crystallisation trials with this protein were not successful.

The majority of *Limulus* CRP binds to ribose-5-phosphate and AMP but virtually none binds to the ligands galacturonic acid, glucuronic acid, N-acetylglucosamine, mannose, galactose, and gluconic acid. If *Limulus* CRP has similar binding properties to human CRP, then the observed *Limulus* CRP binding to these ligands would be expected as human CRP binds well to ribose-5-phosphate and AMP but not very well, if at all, to the other ligands that have been tested due to binding occurring through the phosphate group.

Chapter 4. Isolation and characterisation of *Limulus* SAP

4.1 Introduction.

Limulus pentraxins that were isolated by their calcium dependent PE binding affinity were crystallised (Myles *et al.*, 1990). X-ray diffraction data was collected from multiple crystals, and one of these crystal forms showed the presence of a structure that was not *Limulus* CRP, but of a new structure *Limulus* SAP (Shrive *et al.*, 1999). *Limulus* SAP was found to be a novel PE binding protein in *Limulus* plasma, and is a fairly large constituent of the haemolymph at 8-19% total pentraxin or 0.12-0.17mg/ml haemolymph.

The amino acid sequence of *Limulus* SAP was at first unknown, so the structure was based on a polyalanine model. This showed the structure of *Limulus* SAP to be a doubly stacked hexadecamer of two cyclic octameric rings composed of identical protomers (Shrive *et al.*, 1999) (see figure 4.1.). Other pentraxin structures also show a cyclic arrangement of protomers such as human CRP (Shrive *et al.*, 1996) and SAP (Emsley *et al.*, 1994), as well as a common pentraxin fold which is also displayed by *Limulus* SAP (Shrive *et al.*, 1999). The *Limulus* SAP protomer has similar calcium-binding site topology to human CRP and SAP. It also possesses three disulphide bridges similar to *Limulus* CRP, and amino acid sequence analysis indicates they are in the same location in *Limulus* CRP (Nguyen *et al.*, 1986a, b; Shrive *et al.*, 1999; Tharia *et al.*, 2002). *Limulus* SAP subunits interact with each other non-covalently within the octameric ring and between the two rings (Shrive *et al.*, 1999). Arrangement of protomers within the *Limulus* SAP structure means that the proposed calcium binding site and the pentraxin helix are situated on the external and internal edges of the octamer (Shrive *et al.*, 1999) rather than on opposite faces as in human CRP and SAP indicating a possible difference in function. *Limulus* SAP also does not possess a Clq binding cleft which has been proposed on human CRP but not human SAP (Shrive *et al.*, 1999).

152

Figure 4.1. *Limulus* SAP from top view (a), and side view (b) (adapted from Shrive *et al.*, 1999). The yellow and green colours indicate the hydrogen bonded dimer.

(a)

(b)
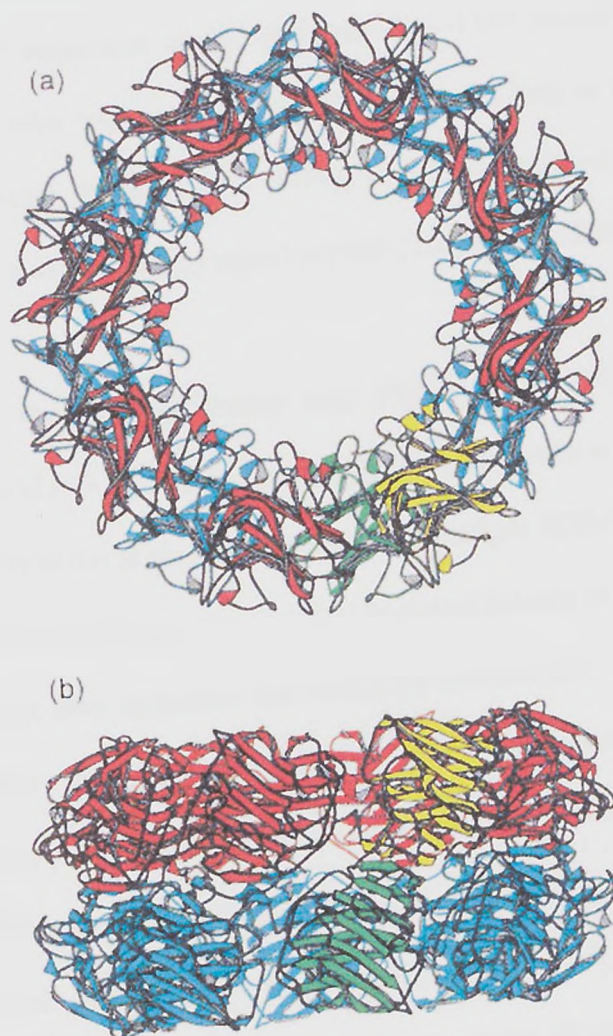
An amino acid sequence had been suggested using the electron density map of X-ray diffraction data from *Limulus* SAP crystals (Shrive *et al.*, 1999). A later study that provided the whole amino acid sequence of *Limulus* SAP derived from cDNA, showed that the predicted amino acid sequence was 70% identical to that derived from cDNA (Tharia *et al.*, 2002). Once the amino acid sequence of *Limulus* SAP had been determined, exact comparisons with sequences of other pentraxins could be performed. *Limulus* SAP had 32% amino acid sequence identity with *Limulus* CRP 1.1, 1.4, and 3.3, 31-35% amino acid sequence identity with *Tachypleus tridentatus* CRP, and 22% amino acid sequence identity with human SAP and CRP (Tharia *et al.*, 2002). This difference of primary structure of *Limulus* SAP with *Limulus* CRP proved that it was a novel protein, separate from the family of *Limulus* CRP but perhaps not so dissimilar in function due to the similar PE

153

binding affinity (Shrive *et al.*, 1999). The cDNA sequence of the *Limulus* SAP gene was 1335bp long, encoding an open reading frame of 702bp from nucleotide 37-738 with two polyadenylation signal sequences at 893-898 and 1313-1319 nucleotides (Tharia *et al.*, 2002). The mature *Limulus* SAP protomer is 217 amino acids long as estimated previously by the polyalanine model (Shrive *et al.*, 1999), and 23.8kDa in weight, but is initially synthesised with a 17 amino acid long signal peptide (Tharia *et al.*, 2002).

In agreement with the electron density map (Shrive *et al.*, 1999), the amino acid sequence derived from cDNA showed a potential glycosylation site at Asn81 (Tharia *et al.*, 2002). Variable glycosylation at this site may explain multiple SDS-PAGE bands seen for subunits of 26-30kDa mass (Shrive *et al.*, 1999) as determination of *Limulus* SAP cDNA sequence gave only one gene sequence that coded for *Limulus* SAP (Tharia *et al.*, 2002). Variable glycosylation suggests that different SAP forms may exist within the haemolymph of *Limulus polyphemus*, and these forms may have different properties and/or functions. Glycosylation can affect protein function either by absence or presence of glycosylated groups such as those in human transferrin (Hunt *et al.*, 1989) or by variation in glycosylated groups such as those on the Fc fragment of IgG (Raju 2008). The possibility of *Limulus* SAP variants with different sequence was also indicated by differences in peptide sequences from cyanogen bromide digestion with the derived amino acid sequence from cDNA (Tharia *et al.*, 2002).

To further analyse the homology of *Limulus* SAP to other pentraxins, phylogenetic analysis using the cDNA derived amino acid sequence was carried out (Tharia *et al.*, 2002). *Limulus* SAP was aligned with the known horseshoe crab sequences of *Limulus* CRP 1.1, 1.4, and 3.3, and *Tachypleus* CRPs, and there were indications that the *Limulus* SAP and CRP split occurred after the split of invertebrates from mammals (Tharia *et al.*, 2002). The phylogenetic analysis also indicated that *Limulus* and *Tachypleus* CRP diverged relatively recently compared to divergence of *Limulus* CRP and SAP. There is

perhaps a common ancestor of SAP and CRP that was present in both chordates and arthropods, as it appears that CRP and SAP divergence occurred after divergence of chordates and arthropods.

The aim of this research was to attempt to isolate different molecular aggregate forms of *Limulus* SAP of which subunits had been identified previously on SDS-PAGE (Shrive *et al.*, 1999), using ligand binding properties and surface hydrophobicity. *Limulus* SAP binds PE (Shrive *et al.*, 1999) but does not bind PC, unlike *Limulus* CRP which binds to both. A typical SAP purification is performed by elution from a PE-agarose column using 30mM PE, so a gradient of 0-30mM PE was used to attempt to isolate different molecular aggregate forms of SAP that could possibly bind PE with different affinity. As there is some sequence similarity with *Limulus* CRP (Tharia *et al.*, 2002; chapter 6) which had been shown to bind both ribose-5-phosphate and AMP (see chapter 3), these ligands were used at both 10mM and 30mM concentrations to attempt to isolate *Limulus* SAP molecular aggregate forms.

Crystallisation trials of *Limulus* SAP, isolated from the plasma of *Limulus*, were continued using conditions based on those which had previously been successful in producing crystals. These crystallisation trials were performed to obtain crystals suitable for collecting diffraction data. Diffraction data could be used to solve the structure of different SAP molecular aggregate forms, or SAP which had been soaked with ligand such as PE to find a ligand bound structure.

4.2 Materials and Methods

4.2.1 Isolation of *Limulus* SAP from *Limulus* plasma

Limulus SAP was isolated from *Limulus* plasma using its binding affinity for PE (Shrive et al., 1999). 5-10mls of *Limulus* plasma was diluted two-fold with calcium wash buffer and applied to a 25ml PE-agarose column using the BIORAD Biologic LP system. Flow rate through the column was maintained at 0.5ml/min for all elutions. Protein elution was monitored at UV_{280nm}. The column was washed with calcium wash buffer to remove all unbound proteins, and then *Limulus* CRP was eluted using 10mM PC in calcium wash buffer as described previously (see section 3.2.1). Following the elution of CRP, the PE-agarose column was equilibrated with calcium wash buffer, and once the UV_{280nm} reading had stabilised, SAP was eluted using 30mM PE in calcium wash buffer. The column was equilibrated again with calcium wash buffer, and then elutions of 10mM EDTA, regeneration buffer A (500mM NaCl, 200mM Tris, 10mM EDTA), regeneration buffer B (500mM NaCl, 50mM Tris, 10mM $CaCl_2$), and regeneration buffer C (150mM NaCl, 50mM Tris, 10mM $CaCl_2$, 0.02% NaN_3) were performed. Approximately 75ml of each elution buffer was applied to the column.

Isolated SAP in the PE elution buffer was dialysed against EDTA to remove the calcium-dependently bound PE, then dialysed against calcium wash buffer to replace calcium ions. After dialysis, SAP was concentrated using Amicon Ultra centrifugal filter devices with a molecular weight cut-off of 10kDa, at 2424rpm to an appropriate concentration or volume for re-application to the PE-agarose column, application to a hydrophobicity column, or crystallisation trials. The concentration of SAP in solution was measured using UV_{280nm} and UV_{320nm} against calcium wash buffer. The UV_{320nm} reading was the background reading for the protein sample, which was subtracted from the UV_{280nm} that was the UV absorbance of the amino acids tryptophan and tyrosine. The

protein reading was then divided by the extinction co-efficient of *Limulus* SAP, which is approximately 1.5 (Abs 0.1%), to give the correct concentration of SAP eluted from the column.

4.2.2 Isolation of *Limulus* SAP by affinity chromatography

SAP isolated from *Limulus* plasma was re-applied to a PE-agarose column, and attempts were made at separating SAP forms by their binding affinity to ribose-5-phosphate and AMP. SAP was diluted to approximately 0.2mg/ml with calcium wash buffer and applied to the PE-agarose column at a flow rate of 0.5ml/min. Any SAP that did not bind to the column (breakthrough) was washed through with calcium wash buffer. Protein elution from the column was measured by UV_{280nm}. Approximately 75mls of 10mM or 30mM ribose-5-phosphate or AMP in calcium wash buffer was then washed through the column, and the SAP elution collected. The column was re-equilibrated with approximately 75mls of calcium wash buffer and a 10mM EDTA buffer applied. The EDTA elution was collected, and the column regenerated using regeneration buffers A, B, and C. 10mM and 30mM NaCl in calcium wash buffer elutions were also performed as a check that increasing the salt concentration associated with ligands ribose-5-phosphate and AMP was not responsible for elution of SAP.

SAP eluted with ribose-5-phosphate and AMP was dialysed against EDTA to remove calcium-dependently bound ligand, then dialysed with calcium wash buffer to replace calcium ions. SAP eluted with EDTA was dialysed against calcium wash buffer to replace calcium ions. Dialysed SAP was concentrated to an appropriate concentration for crystallisation trials or use in FPLC, using Amicon Ultra centrifugal filter devices at 2424rpm. The concentration of SAP was measured at UV_{280nm} and UV_{320nm} against calcium wash buffer, and calculated as described previously.

Attempts at separating *Limulus* SAP forms was also performed using a 0-30mM PE gradient on *Limulus* plasma applied to a PE-agarose column. After a 10mM PC elution removing CRP, a 0-30mM gradient of PE was applied using a single pump and two buffer system (calcium wash buffer and 30mM PE in calcium wash buffer), which provided a 0-100% 30mM PE gradient over 75ml elution volume. The column was re-equilibrated with calcium wash buffer and elutions of 10mM EDTA, regeneration buffer A, B, and C were performed. The fractions of the gradient were collected and subjected to dialysis with EDTA and calcium wash buffer as described previously for *Limulus* SAP isolated from plasma. The concentration of SAP in each fraction was also calculated as described previously for *Limulus* SAP isolated from plasma.

Isolation of *Limulus* SAP using ribose-5-phosphate and AMP as both an isocratic and gradient elution was repeated 2-3 times, and consistently gave similar results.

4.2.3 Isolation of *Limulus* SAP using Fast Protein Liquid Chromatography (FPLC)

Limulus SAP was further analysed using FPLC. This basically involved adding SAP to a hydrophobicity column with which different SAP molecular aggregate forms would interact with different affinity depending on the surface hydrophobicity of the protein. The GE Healthcare ÄKTA Explorer FPLC system was used with a HiTrap Phenyl HP hydrophobicity interaction column (HIC) of 1ml volume. The column was first equilibrated using binding buffer (20mM Tris, 10mM $CaCl_2$, 3.5M $(NH_4)_2SO_4$) for approximately 20ml at a flow rate of 1ml/min. Approximately 350µg of SAP was added to binding buffer to make the volume up to 2ml which was then applied to the column at a flow rate of 0.5ml/min to allow hydrophobic interactions between the SAP and the column. The column was then washed through with approximately 40ml of binding buffer, and SAP eluted with a 40ml gradient (0-100%) of elution buffer (20mM Tris, 10mM $CaCl_2$,

100mM NaCl). The elution of SAP was monitored using UV absorbance at 280nm and 215nm. Once all the SAP had been eluted, the column was then washed with water and ethanol for storage.

FPLC was performed with technical assistance from Dr P. Nield.

Isolation of SAP using FPLC was performed only once for SAP eluted using PE, ribose-5-phosphate, AMP, and the subsequent EDTA eluted SAP.

4.2.4 De-glycosylation of *Limulus* SAP

Limulus SAP was de-glycosylated using an Enzymatic Protein De-glycosylation Kit (EDEGLY) (Sigma). The de-glycosylation reaction was composed of 100µg SAP, 10µl reaction buffer, and 1µl of each enzyme (PNGase F, *O*-glycosidase, α-2(3,6,8,9) neuraminidase, β (1-4) galactosidase, β-*N*-acetylglucosaminidase). The reaction proceeded at 37°C for 5 days, and was stopped by placing at 4°C. *Limulus* SAP was also deglycosylated under denaturing conditions. Denaturation involved using 100µg of *Limulus* SAP with 10µl of reaction buffer and 2.5µl of denaturation solution (Sigma). The reaction was heated at 100°C for 5 minutes then cooled to room temperature, and 2.5µl of TRITONX-100 solution (Sigma) added. The deglycosylation was then performed as previously mentioned. A fetuin (native and denatured) control de-glycosylation was also performed using the same protocol but with kit supplied fetuin, rather than *Limulus* SAP. *Limulus* SAP de-glycosylation was performed with technical assistance from Ian Burns. De-glycosylation was performed only once on *Limulus* SAP and fetuin.

4.2.5 Analysis of *Limulus* SAP using gel electrophoresis

This protocol is the same as that used for *Limulus* CRP (see 3.2.4 Analysis of *Limulus* CRP using gel electrophoresis).

4.2.6 Crystallisation trials of *Limulus* SAP

Once *Limulus* SAP had been isolated using 30mM PE from *Limulus* plasma, it was used in crystallisation trials. Trials were set up following conditions that had been previously used for SAP crystallisation, and had proved successful in producing crystals of SAP.

Limulus SAP was concentrated to approximately 10mg/ml for crystallisation. Stock solutions of buffers to be used in the crystallisation trials were filtered using a 0.2μm pore cellulose acetate filter to remove fungal spores and bacteria. The trays used in the crystallisation trials were 24 well Linbro plates, and the sitting drop method was used. This involved placing mother liquor which contained a mixture of buffers, salts and precipitant into the well up to the volume of 1ml. A microbridge was then placed in to the well and 2μl of mother liquor placed in to the small well on top of it. 2μl of SAP was then mixed with 2μl of mother liquor on the microbridge and the well sealed using vacuum grease and a coverslip.

Crystal trays were kept at room temperature and checked regularly to monitor crystal growth. The mother liquor surrounding crystals used to test for diffraction and data collection, was replaced by cryobuffers at increasing percentages (5, 10, 15 and 20%) of MPD or (5, 10, 15, 20, and 30%) PEG 400. Addition of cryobuffers to crystals, and any subsequent use to obtain diffraction data, was performed in collaboration with Jenny Patterson, Amy Shaw and Prof Trevor Greenhough at Daresbury Synchrotron Radiation Source.

4.3 Results

The subunit size of *Limulus* SAP as determined by SDS-PAGE and the surface hydrophobicity of molecular aggregated *Limulus* SAP, are analysed in section 4.3.1. The effect that deglycosylation has on both subunit size and molecular aggregate surface hydrophobicity of *Limulus* SAP are analysed in section 4.3.2. The results of studies used to isolate *Limulus* SAP using binding affinities to ligands PE, ribose-5-phosphate and AMP are detailed in sections 4.3.3 – 4.3.5. The affect of isolation of *Limulus* SAP on the subunit composition is analysed using SDS-PAGE, and the effect on molecular aggregate SAP surface hydrophobicity is analysed using FPLC. A change in subunit composition between SAP and that isolated using PE gradient, ribose-5-phosphate, or AMP, may infer the presence of different molecular aggregate SAP forms composed of identical subunits that have different binding properties. FPLC on both whole SAP and SAP isolated using a PE gradient, ribose-5-phosphate, and AMP indicates a change in the number of molecular aggregates of SAP that vary in surface hydrophobicity and hence binding properties. The affect of salt concentration on the elution of *Limulus* SAP from the PE-agarose column is analysed in section 4.3.6 to determine whether elution by ribose-5-phosphate and AMP is due to ligand or salt associated with ligand. Section 4.3.7 details crystal trials using SAP and their results.

4.3.1 Isolation of *Limulus* SAP from *Limulus* plasma

As previously described (3.3.1), SAP eluted from *Limulus* plasma was found to give a concentration of approximately 0.3mg/ml of plasma. Elution traces of sequential 10mM PC, 30mM PE, and 10mM EDTA (see figure 3.2) showed the peak corresponding to PE elution smaller than that of PC elution, and sharper. SDS-PAGE analysis shows that SAP is composed of three subunits shown as bands around 29kDa in size (figure 4.2). Lower molecular weight bands at 18.4kDa and 14.2kDa are molecular weight marker overspill.

Figure 4.2. Silver stained SDS-PAGE of *Limulus* pentraxins eluted with 10mM PC and 30mM PE
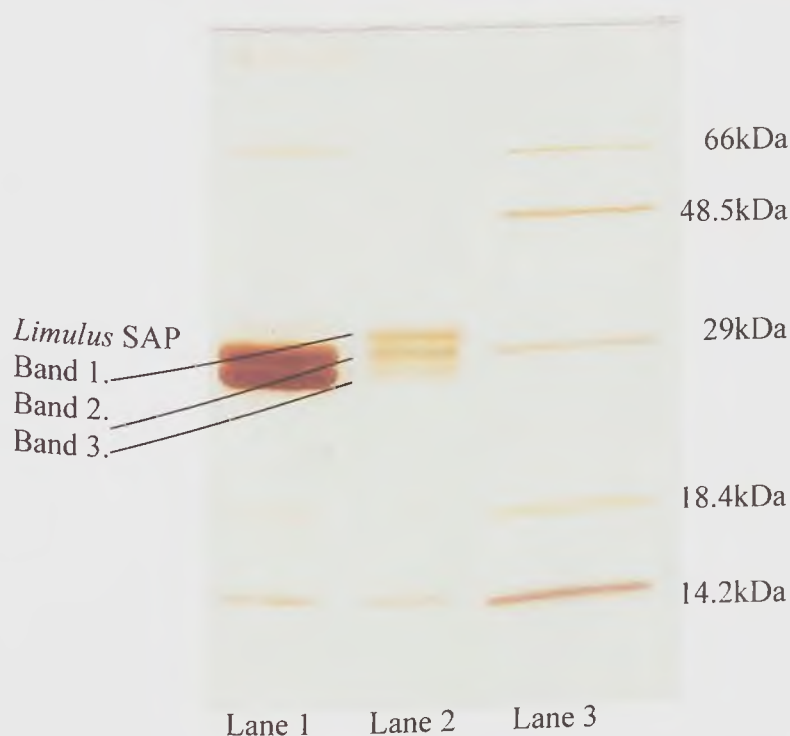
Lane 1   Lane 2   Lane 3

Lane 1: *Limulus* CRP isolated using 10mM PC. Lane 2: *Limulus* SAP isolated using 30mM PE following a 10mM PC elution. Lane 3: Molecular weight markers

Addition of whole SAP isolated using 30mM PE from *Limulus* plasma to a hydrophobicity interaction column (HIC), gave four identifiable peaks (figure 4.3) indicating the presence of at least four molecular aggregate SAP forms that vary in surface hydrophobicity. The first peak was small in comparison to the second and third peak, which appear merged together so that peak two is a shoulder on the ascending limb of peak three. Peak four is not entirely separated from peak three also appearing as a shoulder, and is the smallest of all the peaks.

Figure 4.3 Elution trace of *Limulus* SAP from a hydrophobicity interaction column (HIC) using an increasing concentration of elution buffer. The blue trace shows absorbance at 280nm, magenta shows absorbance at 215nm, and green shows concentration of elution buffer. The major peaks are labelled 1-4.

4.3.2 Deglycosylation of *Limulus* SAP

To identify whether the number of subunits seen on SDS-PAGE were related to glycosylation of *Limulus* SAP, a deglycosylation was performed. The deglycosylation was performed only once, but the procedure was carried out over the maximum time allowing the enzymes to deglycosylate the *Limulus* SAP completely, and a fetuin control was

performed to check that the enzymes were working. The effect of deglycosylation on SAP

subunit size was analysed using SDS-PAGE, and the effect on molecular aggregate SAP

surface hydrophobicity was analysed using FPLC. SDS-PAGE analysis (figure 4.4) of

deglycosylated SAP against denatured deglycosylated SAP and whole SAP indicated that

deglycosylation reduces the molecular weight of SAP subunits.

Figure 4.4. Coomassie blue stained SDS-PAGE gel of de-glycosylated *Limulus* SAP. Lane 1: Sigma low range molecular weight markers. Lane 2: deglycosylated *Limulus* SAP. Lane 3: denatured and deglycosylated *Limulus* SAP. Lane 4: *Limulus* SAP (non-deglycosylated or denatured). Lane 5: Fetuin (deglycosylated) Lane 6: Fetuin (deglycosylated and denatured) Lane 7: Fetuin (non deglycosylated and non denatured).

A fetuin control was performed to check that deglycosylation enzymes were working. The deglycosylated fetuin (lane 5) has lower molecular weight than the non deglycosylated fetuin (lane 7) indicating that glycosylated groups have been removed and that the enzymes were working. Non-deglycosylated SAP (lane 4) was composed of three bands: two approximately 30-32kDa in size, and a smaller band between 24-29kDa in size. These bands were also seen on figure 4.2 (lane 2). Deglycoslyated SAP (lane 2) also showed

three bands, of which the two heaviest in molecular weight appeared smaller in size than those in whole non-deglycosylated SAP. The second highest molecular weight band appeared more darkly stained on the gel in deglycosylated SAP indicating that it was in higher concentration than the other two bands. This may be due to deglycosylation of SAP composing the higher molecular weight band (at ~32kDa) in SAP, therefore reducing the molecular weight which appears as an increase in concentration (i.e. darker staining) of the second highest molecular weight band. SAP that was denatured and deglycosylated (lane 3) appears to have two darkly stained bands; one at a similar size to the second highest molecular weight band in non-deglycosylated SAP and deglycosylated SAP (~30kDa), and one at a size smaller than the smallest band of non-deglycosylated SAP and deglycosylated SAP. There is no third band visible although this may be due to dark staining of the two visible bands. There is also a band present at 66kDa in size which may be haemocyanin or contamination.

Deglycosylated SAP was subjected to analysis by FPLC and the UV trace is shown in figure 4.5. The elution trace shows only three peaks which correspond to peaks 1, 3 and 4 of the non-deglycosylated SAP trace (figure 4.3). The shoulder on the ascending limb of the largest peak (peak 3) is missing which indicates that the shoulder present on peak 3 in non-deglycosylated SAP is removed by deglycosylation and may become part of peaks 1, 3 or 4. The other two peaks (peaks 1 and 4) are not removed by SAP deglycosylation, indicating a difference between surface hydrophobicity of SAP molecular aggregate forms that is not due to glycosylation. Such an example could be sequence differences, as there is a possibility of multiple SAP genes present similar to the gene multiplicity of *Limulus* CRP, although at present only one sequence has been found. There could also be other structural modifications such as phosphorylation or sulphation. It is also possible that one or both of the two smaller peaks (peaks 1 and 4) are impurities within the *Limulus* SAP sample such as haemocyanin, *Limulus* CRP, or minor components of the haemolymph.
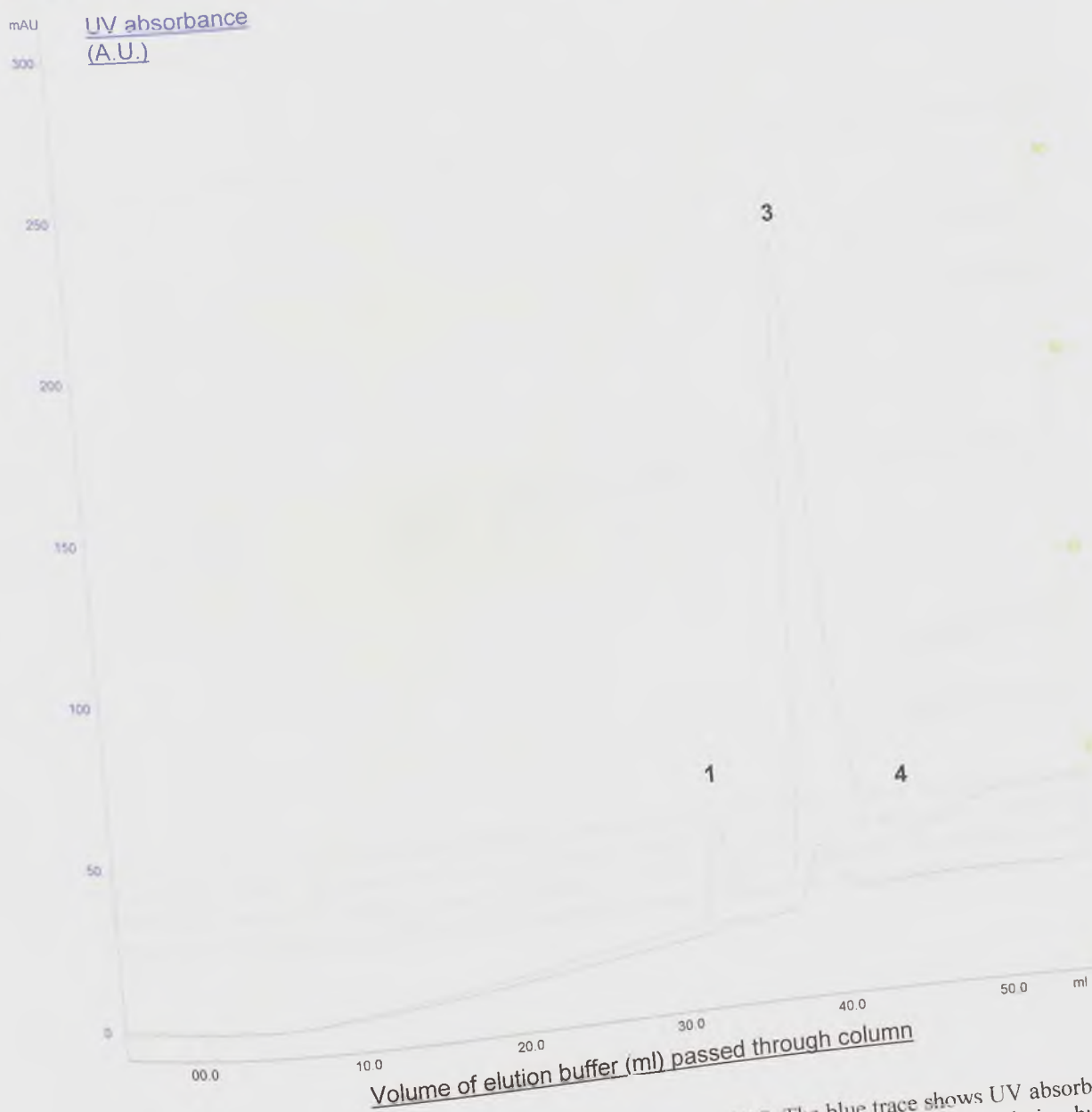
Figure 4.5. Elution profile of deglycosylated *Limulus* SAP from a HIC. The blue trace shows UV absorbance at 280nm, magenta shows UV absorbance at 215nm, and green trace shows concentration of elution buffer. The major peaks are labelled 1, 3 and 4 according to figure 4.3.

4.3.3 Isolation of *Limulus* SAP using a 30mM PE gradient

Limulus plasma was applied to a PE-agarose column and an elution of 10mM PC was performed which eluted CRP. A 0-100% 30mM PE gradient was then applied to the column to attempt to separate SAP. The elution trace (figure 4.7) showed only one peak with no apparent peak separation. SDS-PAGE gel analysis of the fractions collected from the PE gradient, also showed no apparent separation of SAP subunits as all PE gradient

lanes showed the presence of three bands as seen in figure 4.2, with equal staining in all

lanes (figure 4.6).

Figure 4.6. Silver stained SDS-PAGE of *Limulus* SAP isolated using a 0-30mM PE gradient after a 10mM PC elution.

Lane 1: Fraction 3 of SAP eluted using 30mM PE gradient. Lane 2: Fraction 5 of SAP eluted using 30mM PE gradient. Lane 3: Fraction 8 of SAP eluted using 30mM PE gradient. Lane 4: Fraction 10 of SAP eluted using 30mM PE gradient. Lane 5: Fraction 12 of SAP eluted using 30mM PE gradient. Lane 6: Remaining SAP eluted using EDTA after PE gradient fractions. Lane 7: Molecular weight markers.
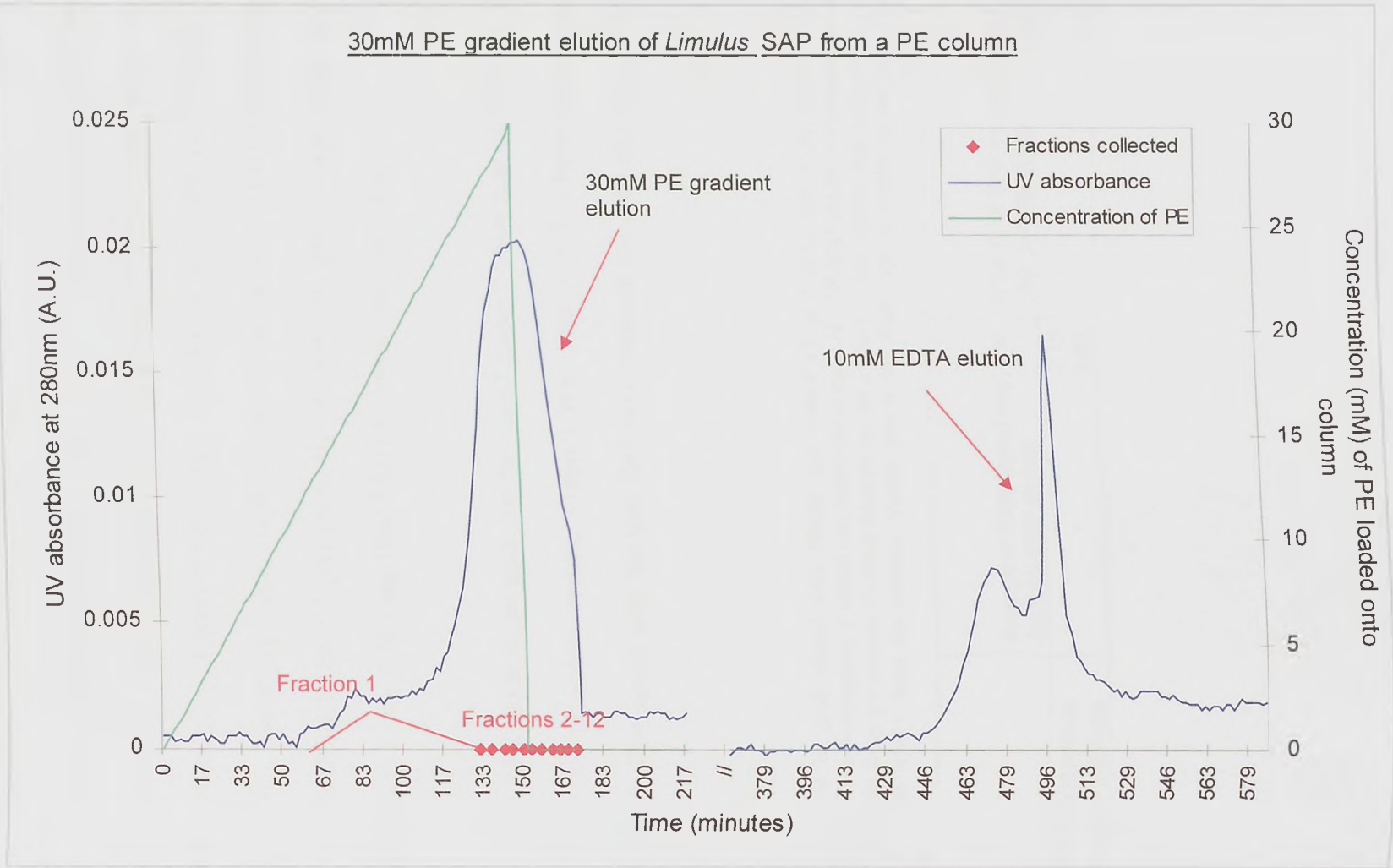
Figure 4.7. Elution profile of 30mM PE gradient of *Limulus* SAP following a 10mM PC elution from a PE column, followed by a 10mM EDTA elution. // indicates a break in time between elutions. Start of points of collection of fractions 1-12 are marked.

4.3.4 Isolation of *Limulus* SAP using ribose-5-phosphate

Isolated *Limulus* SAP was re-applied to a PE-agarose column and eluted with 10mM and 30mM ribose-5-phosphate. Remaining SAP on the column was eluted with 10mM EDTA. An elution trace showed a very small extended peak for the 10mM ribose-5-phosphate elution, but a large sharp peak for the following 10mM EDTA elution (see figure 4.8). The 30mM ribose-5-phosphate elution however, gave a large peak, and a smaller sharp peak for the following 10mM EDTA elution (see figure 4.9).

The presence of a much larger peak for the 30mM ribose-5-phosphate elution than for the 10mM ribose-5-phosphate elution, suggests that more SAP has been eluted using 30mM of ligand. This agrees with the total SAP concentrations for each elution (see table 4.1).

Table 4.1. Amount of SAP eluted with 10mM and 30mM ribose-5-phosphate. Column one shows concentration of ribose-5-phosphate. Column two shows the amount of SAP (mg) eluted using ligand per mg of SAP eluted totally from the column. Column three shows the amount of SAP eluted (mg) using EDTA per mg of SAP eluted totally from the column. Values shown are from typical results of multiple (2-3) experiments.

	Amount of SAP eluted (mg) per mg total SAP eluted.	Amount of SAP eluted by EDTA (mg) per mg of total SAP eluted
10mM ribose-5-phosphate	0.15	0.845
30mM ribose-5-phosphate	0.92	0.079
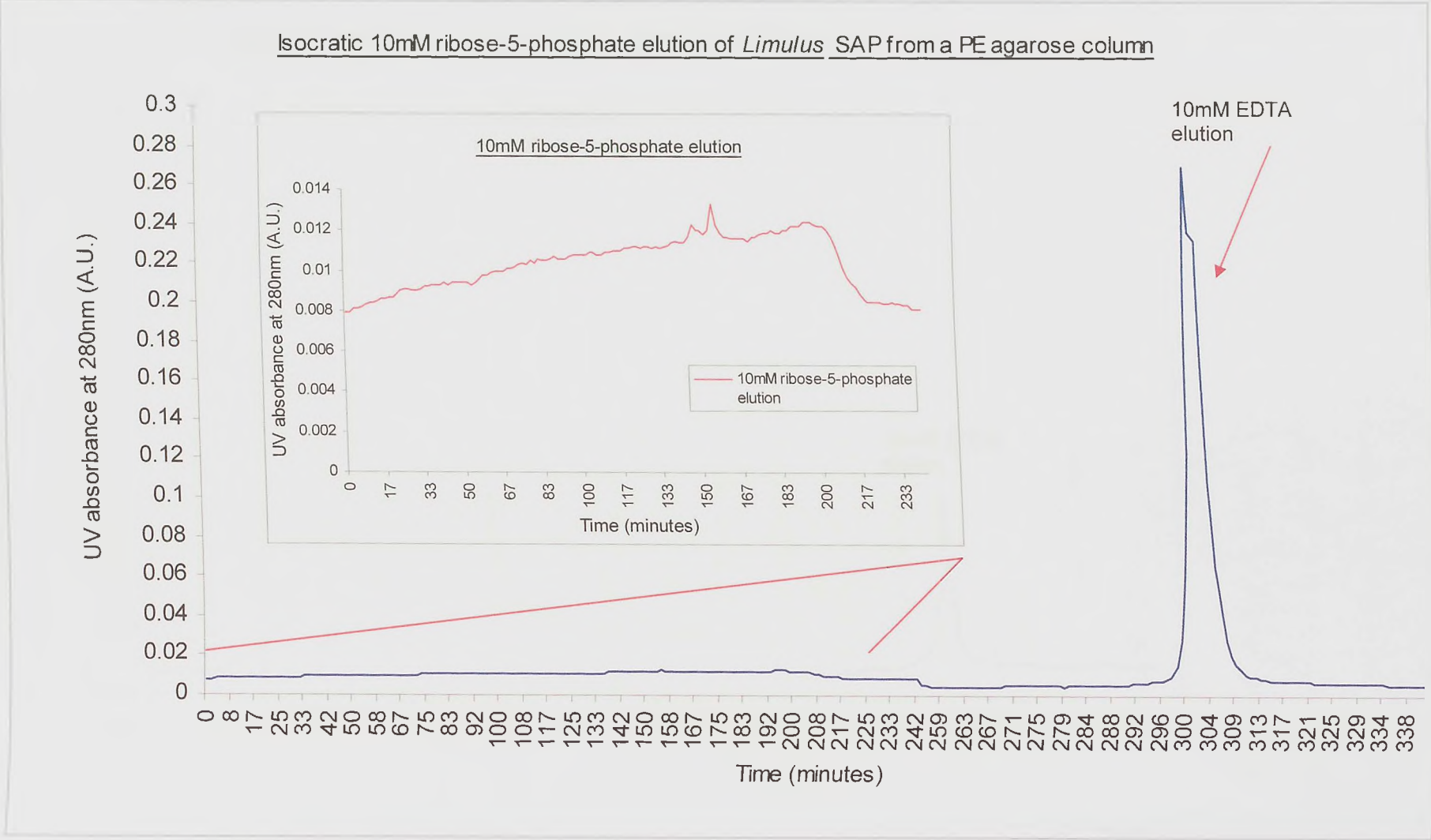
Figure 4.8. Elution profile of *Limulus* SAP eluted using 10mM ribose-5-phosphate followed by a 10mM EDTA elution.
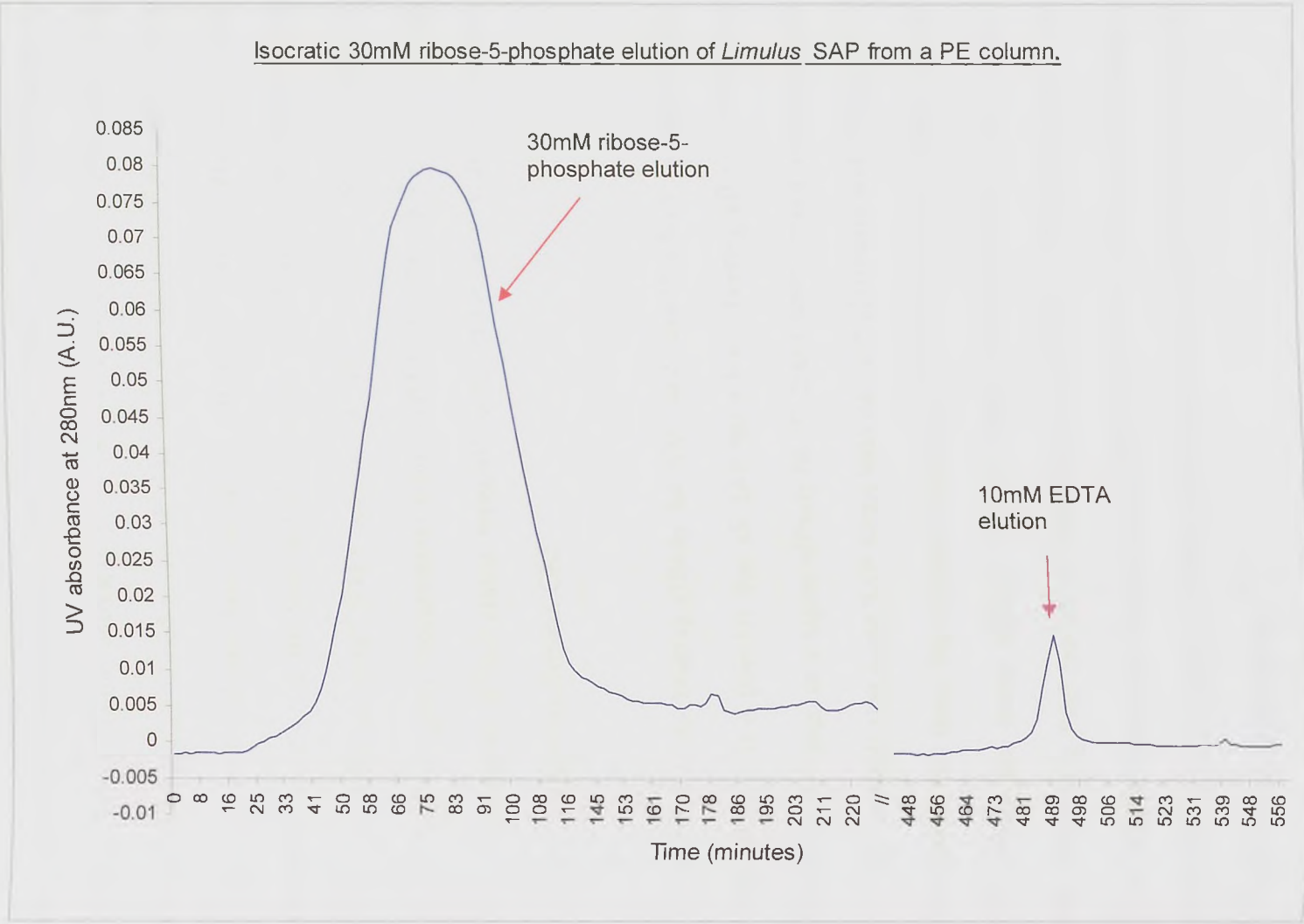
Figure 4.9. Elution profile of *Limulus* SAP eluted using sequential elutions of 30mM ribose-5-phosphate and 10mM EDTA. // indicates a break in time between elution.

The small amount of SAP eluted using 10mM ribose-5-phosphate corresponds to the very small peak shown on the elution trace, whilst the larger amount of SAP eluted using the following EDTA elution corresponds to the much larger EDTA peak. The 30mM ribose-5-phosphate elution eluted nearly all SAP from the column which agreed with the much larger peak on the elution trace, whilst the subsequent EDTA elution eluted only a small amount of protein (i.e. the SAP not eluted using 30mM ribose-5-phosphate) which corresponded with the very small EDTA peak on the elution trace. It is surprising that such a small difference in ligand concentration from 10mM to 30mM could give such a large difference in the amount of SAP eluted, especially considering the very slight effect it has on the homologous protein *Limulus* CRP.

SAP could have a low binding affinity for AMP and require a higher concentration to elute nearly all of it. It is possible that all SAP has a lower binding affinity for ribose-5-phosphate than CRP does, or a higher affinity for PE than ribose-5-phosphate compared to CRP. It is also possible that some SAP forms bind ribose-5-phosphate with higher affinity than other forms, and these high affinity binding forms are eluted with 10mM ribose-5-phosphate, whilst the lower affinity binding forms are eluted with 30mM ribose-5-phosphate. Remaining SAP that is not eluted at 30mM ribose-5-phosphate may have an even lower binding affinity for ribose-5-phosphate or none at all. However, this fraction is only small, and the most obvious difference is between the amounts of SAP eluted using 10 and 30mM ribose-5-phosphate.

If the two elution profiles of *Limulus* CRP and SAP eluted with 30mM ribose-5-phosphate are shown on the same graph (see figure 4.10), it appears that *Limulus* SAP has a lower affinity for 30mM ribose-5-phosphate than immobilised PE compared to *Limulus* CRP. As both profiles start when elution buffer containing 30mM ribose-5-phosphate is added to the PE column, the *Limulus* CRP elution profile shows a quicker response than

Limulus SAP indicating a higher affinity for ribose-5-phosphate than *Limulus* SAP. This is also shown by the much sharper peak for *Limulus* CRP than *Limulus* SAP, indicating that it takes a much longer time for *Limulus* SAP to elute from the column under the same conditions as *Limulus* CRP. This also occurs at a lower protein concentration of 2mg of SAP compared to 6mg of CRP, and it is possible that the *Limulus* SAP elution profile would show an extended peak if 6mg were placed on the PE column rather than 2mg. The *Limulus* CRP elution had been extended due to a reduced flow rate (around 60 minutes on figure 4.10) which showed an extended second peak after the initial first peak. Should this have not occurred the differences in elution profiles may be more evident and more comparable. Therefore, if an elution of 30mM ribose-5-phosphate was performed on *Limulus* CRP and SAP together it may be able to highlight the different levels of affinity to this ligand that both proteins have. However, the affinity of these proteins to ribose-5-phosphate in elution buffer is also governed by their affinity for immobilised PE on the column, and therefore not many conclusions could be drawn from this.
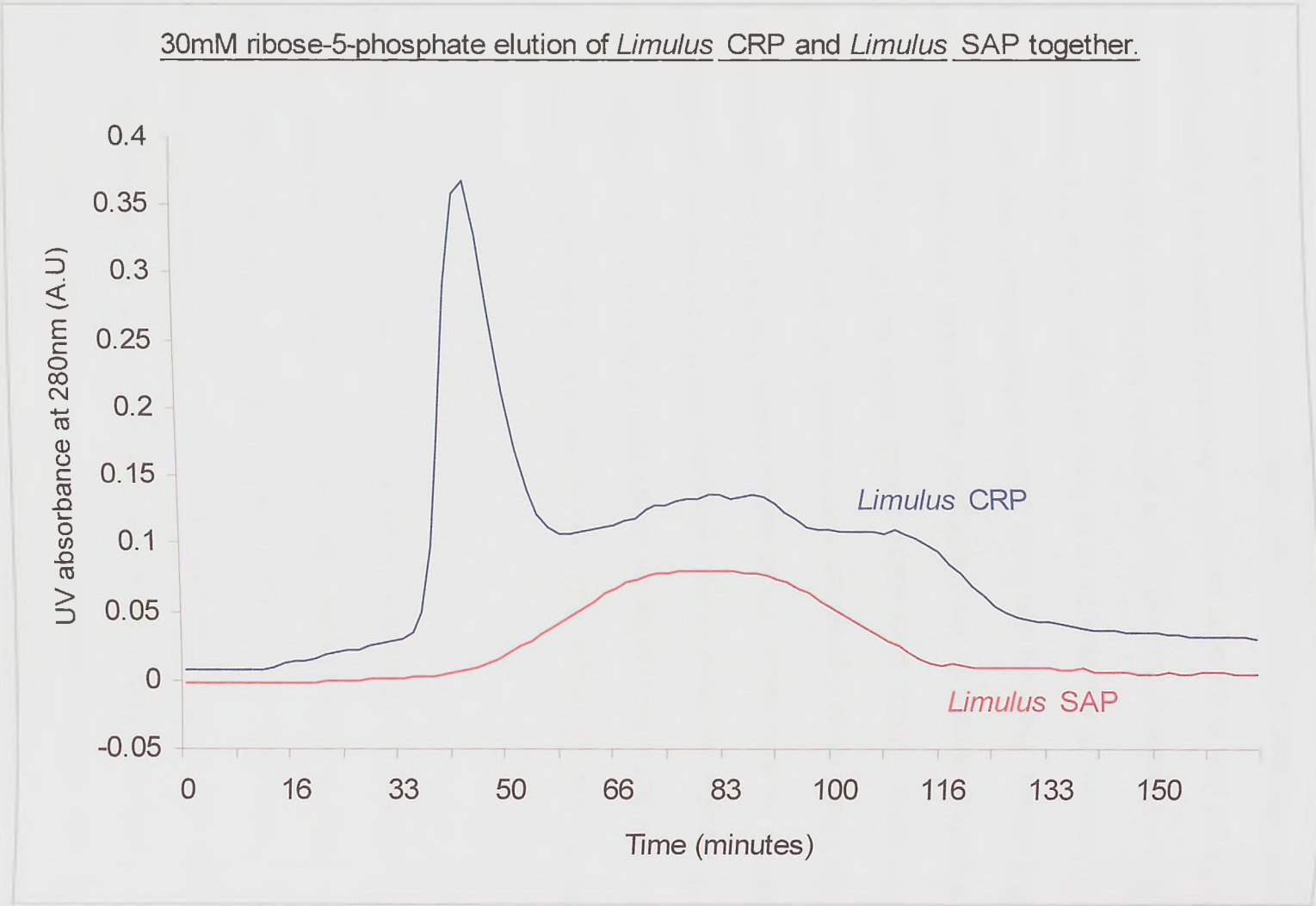
Figure 4.10 Elution profiles of *Limulus* CRP and SAP eluted using 30mM ribose-5-phosphate.

Increasing the length of time the 10mM ribose-5-phosphate elution ran over could elute more SAP, which would indicate all forms of SAP had low binding affinity for SAP. As this was not performed, SDS-PAGE analysis was performed on the fractions of SAP eluted using 10mM and 30mM ribose-5-phosphate, and the subsequent EDTA elutions, to determine whether there was a change in subunit composition. This may identify whether there is a change in SAP molecular aggregate forms that bind at 10 and 30mM ribose-5-phosphate assuming that changes in subunit size reflect changes in molecular aggregated SAP forms.

SDS-PAGE analysis (figure 4.11) shows that SAP eluted using 10mM ribose-5-phosphate has a darker lower molecular weight band of the three bands ~29kDa compared to the SAP eluted using 30mM ribose-5-phosphate which shows that the higher molecular weight band of the three bands ~29kDa stained darkest. The SAP eluted using the EDTA elution after 10mM ribose-5-phosphate elution showed the higher molecular weight band of the three bands ~29kDa more darkly stained, whilst the EDTA elution after 30mM ribose-5-phosphate showed the lower molecular weight bands of the three bands ~29kDa more darkly stained. There are also bands present at approximately 66kDa, 48.5kDa, and one present between 29 and 48.5kDa in all lanes on the gel. In lane 2 (SAP eluted using 10mM ribose-5-phosphate) the band present between 29kDa and 48.5kDa in size appears darker on the gel than in the other lanes. This is probably because any impurities present will be more concentrated due to the centrifugal concentration of the small amount of SAP protein eluted from the column using 10mM ribose-5-phosphate.

Therefore, there appears to be a reduction in the presence of the highest molecular weight band ~29kDa in size of SAP eluted by 10mM ribose-5-phosphate indicated by a reduction in staining of this band compared to the SAP eluted using 30mM ribose-5-phosphate and the EDTA elution following the 10mM ribose-5-phosphate. Some of this

175

band may represent a subunit which forms a molecular aggregate SAP form that has a

lower binding affinity for ribose-5-phosphate, and so is present in lower amounts in the

SAP eluted using 10mM ribose-5-phosphate elution but higher amounts in the subsequent

SAP eluted using EDTA and also the SAP eluted using 30mM ribose-5-phosphate. SAP

eluted using 30mM ribose-5-phosphate shows similar staining to whole SAP, as does the

SAP eluted using EDTA after the 10mM ribose-5-phosphate elution, which correlates with

the majority of SAP eluted being present in these fractions.

Figure 4.11. Silver stained SDS-PAGE gel of *Limulus* SAP eluted using 10mM and 30mM
ribose-5-phosphate from a PE agarose column.

Lane 1: Molecular weight marker. Lane 2: SAP eluted using 10mM ribose-5-phosphate. Lane 3: SAP
eluted using 30mM ribose-5-phosphate. Lane 4: SAP eluted using EDTA after 10mM ribose-5-
phosphate. Lane 5: SAP eluted using EDTA after 30mM ribose-5-phosphate. Lane 6: whole SAP.
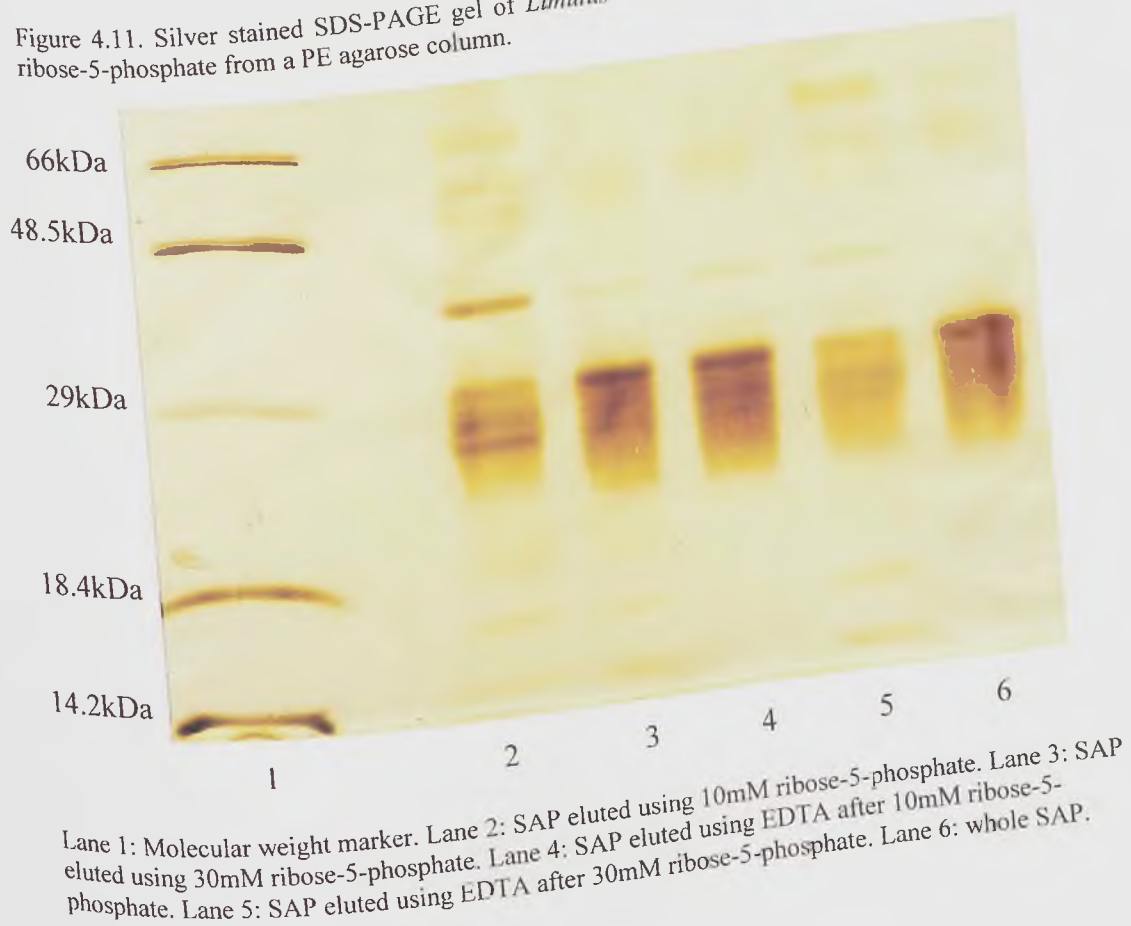
Surface hydrophobicity analysis of SAP eluted using 10mM ribose-5-phosphate using a

HIC (figure 4.12) showed the four peaks that were present in whole SAP (figure 4.3). Peak

1 looked comparatively larger than the other peaks compared to peak 1 of whole SAP.

Peaks 2 and 3 appeared as a single peak with 3 as a slight shoulder on peak 2, rather than

vice versa in whole SAP. Peak 4 was also present, but much smaller compared to that of

whole SAP. Therefore, it appears that ribose-5-phosphate at 10mM concentration eluted all

four molecular aggregate forms of SAP present in whole SAP represented by peaks on the elution trace, but the comparative sizes of the peaks are different suggesting a difference in the SAP present which may represent some weak separation of the four molecular aggregate forms of *Limulus* SAP as indicated by SDS-PAGE, although which bands correspond to which peaks on the FPLC elution is unknown.
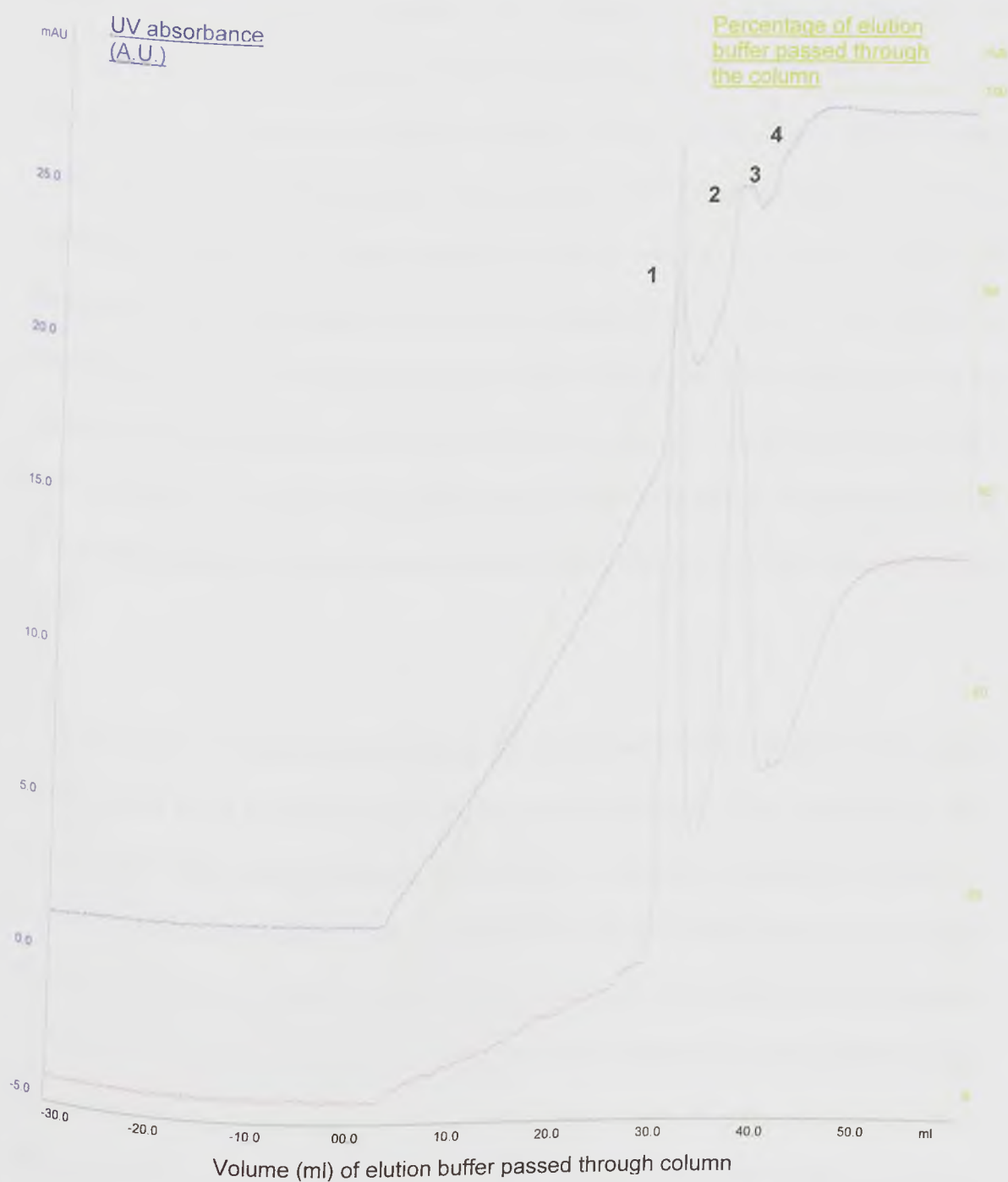
Figure 4.12. Elution profile of *Limulus* SAP previously eluted with 10mM ribose-5-phosphate, eluted from a HIC. The blue trace shows UV absorbance at 280nm, magenta trace shows UV absorbance at 215nm, and green trace shows concentration of elution buffer. The major peaks are labelled 1-4 according to peaks identified in figure 4.3.

The SAP eluted using a 10mM EDTA elution following the 10mM ribose-5-phosphate elution was also applied to a HIC by FPLC and separated using surface hydrophobicity. The UV elution trace (figure 4.13) showed the four peaks present in whole SAP. Peak 1 appears comparatively larger to the other peaks than it does in whole SAP similar to SAP eluted using 10mM ribose-5-phosphate. This suggests that protein that corresponds to this peak contains both SAP that binds to ribose-5-phosphate and that which does not. It could also possibly be contamination which is present in whole SAP too. Peak 3 appears slightly more evidently as a shoulder to peak 2 than it does in the SAP eluted using 10mM ribose-5-phosphate, and there is a slight shoulder on the ascending limb of peak 2 which is not present in the SAP eluted using 10mM ribose-5-phosphate or whole SAP. This suggests an additional group of SAP which does not bind 10mM ribose-5-phosphate and was not apparent on the elution trace for whole SAP. Peak 4 appears comparatively larger than it does on the trace for the SAP eluted using 10mM ribose-5-phosphate, suggesting that more SAP corresponding to this peak does not bind 10mM ribose-5-phosphate than that which does.

Overall, peaks 2-4 on the FPLC trace for SAP eluted using 10mM EDTA, appear comparatively larger to the corresponding peaks on the previous FPLC trace for the SAP eluted using 10mM ribose-5-phosphate, and peak 1 appears comparatively smaller. A change like this was expected if the SAP eluted with 10mM ribose-5-phosphate contained one or more different molecular aggregate forms of SAP as compared to the remaining EDTA eluted fraction. However, considering the small amount of protein actually eluted using 10mM ribose-5-phosphate, it is possible that there is contamination by trace impurities which would be amplified during concentration of the SAP fraction, and these cause the difference in FPLC elution trace profile.

Chapter 4. Isolation and characterisation of *Limulus* SAP
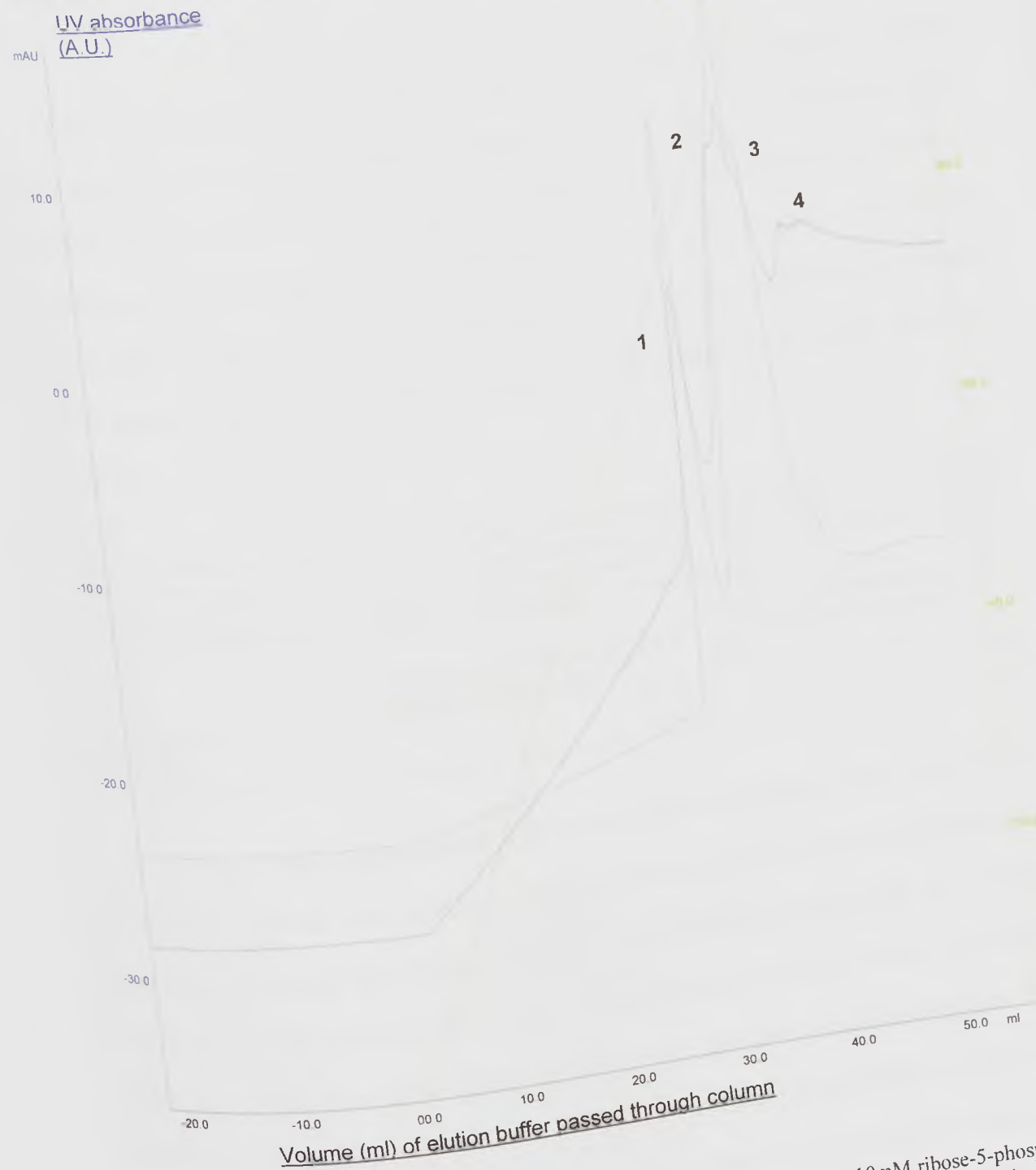
Figure 4.13. Elution profile of *Limulus* SAP previously eluted with EDTA after a 10mM ribose-5-phosphate elution, from a HIC. The blue trace shows UV absorbance at 280nm, magenta trace shows UV absorbance at 215nm, and green trace shows concentration of elution buffer. The major peaks are labelled 1-4 according to those identified in figure 4.3.

4.3.5 Isolation of *Limulus* SAP using AMP

SAP isolated from *Limulus* plasma by 30mM PE was re-applied to a PE-agarose column and eluted with 10mM and 30mM AMP. Remaining SAP on the column was eluted with 10mM EDTA. The elution trace for both 10mM and 30mM AMP elutions (not shown) showed an AMP dependent rise in UV absorbance similar to that of *Limulus* CRP eluted using AMP (see section 3.3.5). The 10mM EDTA peak was much larger after the 10mM AMP elution than the 30mM AMP elution which corresponded to the figures shown in table 4.2.

Table 4.2. Amount of SAP eluted with 10mM and 30mM AMP. Column one shows the concentration of AMP. Column two shows the amount of SAP (mg) eluted using ligand per mg of SAP eluted in total from the column. Column three shows the amount of SAP (mg) eluted by EDTA per mg SAP eluted in total from the column. The values shown are typical results of multiple (2-3) experiments.

	Amount of SAP eluted (mg) per mg SAP eluted in total.	Amount of SAP eluted by EDTA (mg) per mg of SAP eluted in total
10mM AMP	0.10	0.90
30mM AMP	0.96	0.04

10mM AMP elutes approximately 10% of the total SAP eluted, whilst 30mM AMP elutes nearly all SAP. This is similar to the SAP elution using 10mM and 30mM ribose-5-phosphate. It appears that slightly less SAP was eluted with 10mM AMP than 10mM ribose-5-phosphate used previously, but that slightly more SAP was eluted with 30mM AMP than 30mM ribose-5-phosphate used previously. The same principles apply as with SAP eluted using ribose-5-phosphate: SAP could have low binding affinity for AMP and require a higher concentration to elute nearly all of it within the time span of the elution; or there could be both low and high binding affinity forms of SAP for AMP, of which some forms elute at 10mM and the majority elute at 30mM. To determine whether this was true, SDS-PAGE was performed to identify changes in subunit sizes, and FPLC was performed to identify changes in molecular aggregate protein surface hydrophobicity.

SDS-PAGE analysis of SAP (figure 4.14) eluted using 10mM AMP showed three bands ~29kDa of which the higher molecular weight band was the lightest stained and the lower two molecular weight bands were darker stained although this was difficult to observe on the gel. The lower two molecular weight bands appear to be the same size of the lowest two molecular weight bands of whole SAP. SAP eluted using 30mM AMP showed three bands ~29kDa in size of which the higher molecular weight band was darker stained than the lower molecular weight bands. This staining is similar for the SAP eluted using 10mM EDTA after the 10mM AMP elution. SAP eluted using 10mM EDTA after the 30mM AMP elution shows similar staining to the SAP eluted using 10mM AMP where there are three bands with more equal staining.

Figure 4.14 Silver stained SDS-PAGE of *Limulus* SAP eluted with 10mM and 30mM AMP.
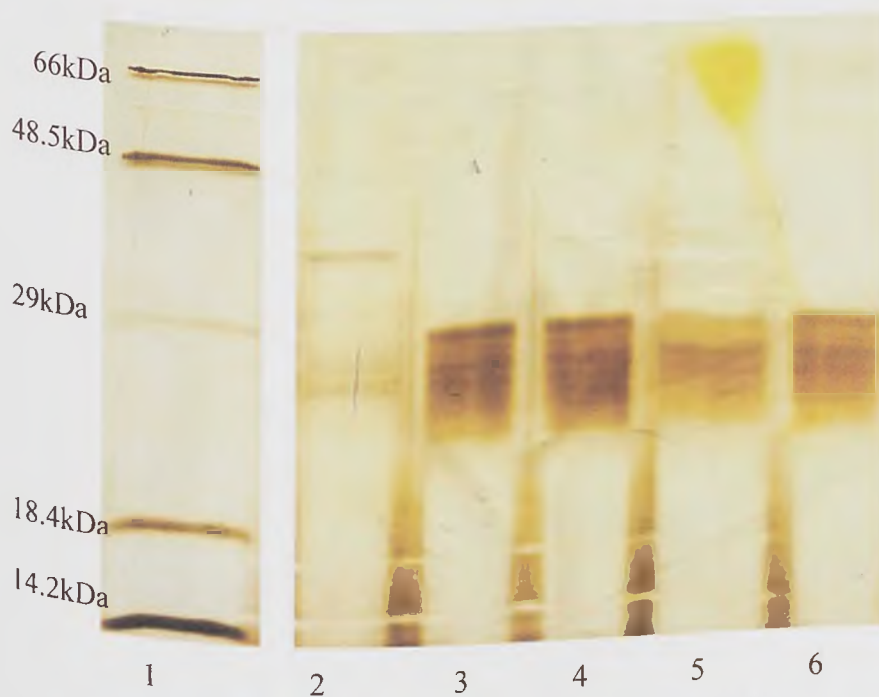
Lane 1: Molecular weight markers. Lane 2: SAP eluted using 10mM AMP. Lane 3: SAP eluted using 30mM AMP. Lane 4: SAP eluted using 10mM EDTA after 10mM AMP elution. Lane 5: SAP eluted using 10mM EDTA after 30mM AMP elution. Lane 6: whole SAP.

Overall, the fractions which correspond to the elutions that contained the most SAP had darker staining of, and therefore higher proportion of, SAP composing the highest molecular weight band of the three bands ~29kDa, whilst those that correspond to elution

that contained the least SAP had the darkest staining and therefore higher proportion of SAP composing the lowest molecular weight bands of the three bands ~29kDa. There are also bands present at approximately 66kDa, 45.8kDa, and one between 29 and 45.8kDa in all lanes. This is similar to the band pattern shown by SAP eluted using 10mM and 30mM ribose-5-phosphate and indicates a reduction in protein corresponding to the higher molecular weight band ~29kDa in SAP eluted using 10mM AMP compared to that eluted using 30mM AMP.

SAP that was isolated using 10mM AMP was analysed using surface hydrophobicity interactions. FPLC analysis of SAP eluted using 10mM AMP by application to a HIC, showed the presence of three peaks (figure 4.15). The first peak appears comparatively larger to the other peaks compared to that of the first peak of whole SAP (figure 4.3). Peak 2 is much smaller than peak 1 and corresponds to peak 2 or 3 of whole SAP with a shoulder on the descending limb. The third peak is the smallest of the three and corresponds to peak 4 of whole SAP.
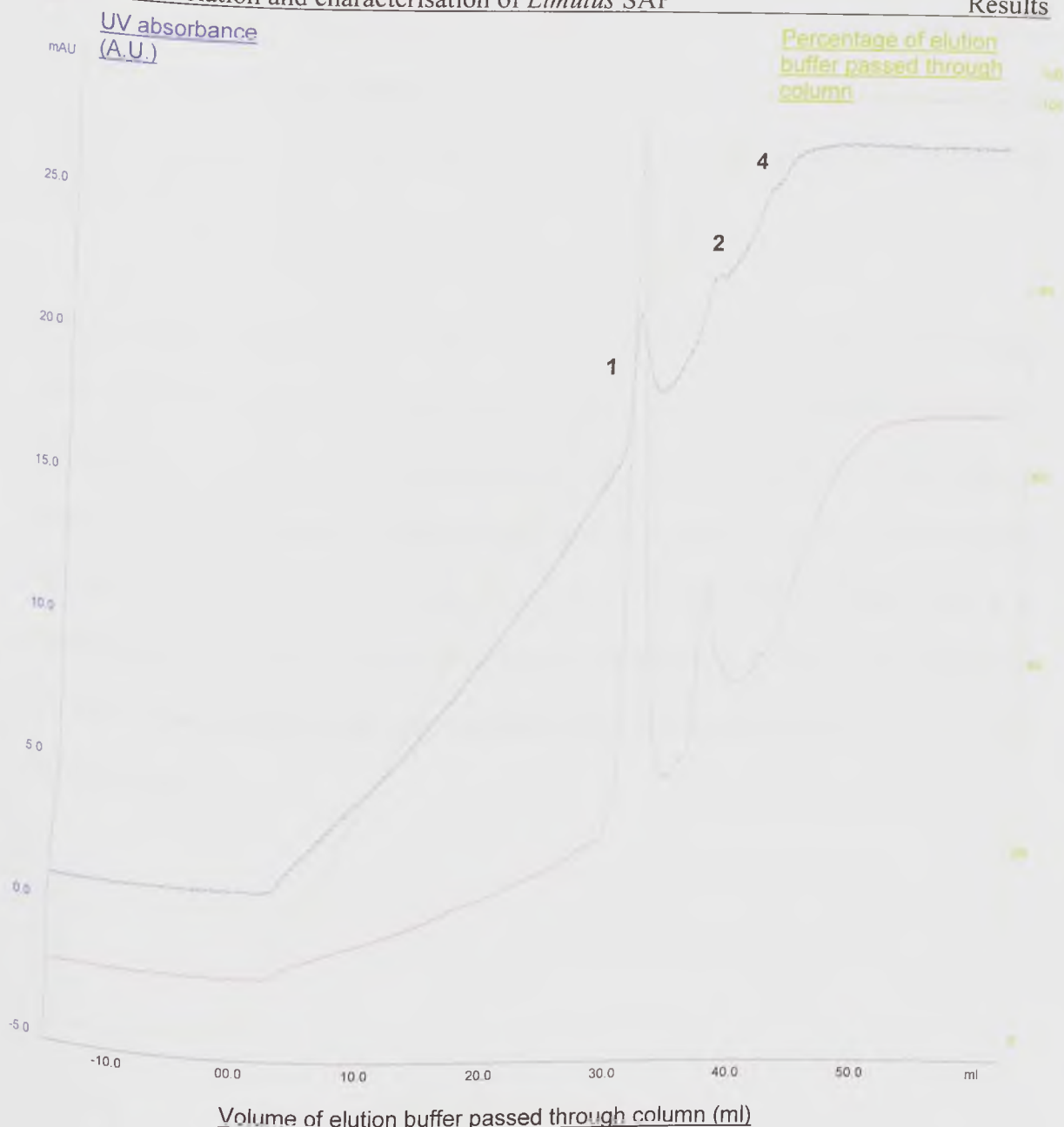
Figure 4.15. Elution profile of *Limulus* SAP previously eluted using 10mM AMP, eluted from a HIC. The blue trace shows UV absorbance at 280nm, magenta trace shows UV absorbance at 215nm, and green trace shows concentration of elution buffer on column. The major peaks are labelled 1, 2 and 4 according to those identified in figure 4.3.

Analysis of SAP eluted using 10mM EDTA after a 10mM AMP elution on the FPLC shows four peaks (figure 4.16). Peak 1 is approximately in the same proportion as it is in the SAP eluted using 10mM EDTA after the 10mM ribose-5-phosphate elution (figure 4.13). This suggests either that SAP corresponding to this peak is eluted more so with 10mM ribose-5-phosphate and AMP than that eluted after with 10mM EDTA. There is also a possibility that it is contamination present throughout all fractions, and not showing specificity to any ligand therefore appearing in all fractions. It appears larger in the traces

for 10mM AMP and ribose-5-phosphate perhaps because there was little SAP in these fractions so any protein contamination was amplified more so during SAP concentration than in the 10mM EDTA elution which contained a larger amount of SAP and so was concentrated less.

Peaks that correspond to peak 2 and 3 in whole SAP are present also in SAP eluted using 10mM EDTA, but close together as dual peaks not one as a shoulder of the other. This indicates that some SAP corresponding to peak 2 or 3 is eluted with 10mM AMP as shown in figure 4.15 as peak 2, whilst the remainder is eluted with EDTA. It also indicates that some SAP corresponding to peak 2 or 3 does not bind AMP at 10mM. Peak 4 is present in both SAP eluted using 10mM AMP and the following EDTA elution suggesting that SAP corresponding to this peak has both AMP binding and non-binding molecular aggregate forms.

UV absorbance (A.U.)

Percentage of elution buffer passed through column

Volume of elution buffer passed through column (ml)

Figure 4.16. Elution profile of *Limulus* SAP previously eluted with EDTA after a 10mM AMP elution, eluted using HIC. The blue trace shows UV absorbance at 280nm, magenta trace shows UV absorbance at 215nm, and green trace shows concentration of elution buffer. The major peaks are labelled 1-4 as identified on figure 4.3.

Overall, on the FPLC trace for SAP eluted using 10mM AMP, peaks 2-4 were comparatively much smaller than those on the FPLC trace for SAP eluted using EDTA after the 10mM AMP elution. Peak 1 appears comparatively much smaller on the FPLC trace for SAP eluted using EDTA after 10mM AMP than it does on the FPLC trace for

185

SAP eluted using 10mM AMP. This was to be expected if there was more proportionally

of one SAP molecular aggregate form than another, in the fraction eluted with 10mM AMP

compared to that eluted after with EDTA. However, considering the small amount of

protein actually eluted using 10mM AMP, it is possible that there is contamination by trace

impurities which would be amplified during concentration of the SAP fraction, and these

cause the difference in the FPLC elution trace profile.

4.3.6 Isolation of *Limulus* SAP using sodium chloride

Limulus SAP isolated from *Limulus* plasma using 30mM PE was re-applied to a PE-

agarose column and eluted using 10mM and 30mM sodium chloride. This was performed

to see if the elution of SAP by 10mM and 30mM ribose-5-phosphate and AMP was due to

ligand interaction and not an increase in salt concentration associated with the ligand. The

elution profile of 10mM NaCl elution (figure 4.17) and the elution profile of 30mM NaCl

elution (figure 4.18) shows no increase in UV absorbance (i.e. no peak) corresponding to

either salt concentration, but a sharp peak for the 10mM EDTA elution following them.

The amount of SAP eluted is shown in table 4.3 from single experiments.

Table 4.3. Amount of SAP eluted using 10mM and 30mM NaCl elutions. Column one shows the concentration of sodium chloride. Column two shows the proportion of SAP (mg) eluted using ligand per mg of SAP eluted totally. Column three shows the proportion of SAP (mg) eluted by EDTA per mg of SAP eluted totally.

	Amount of SAP eluted (mg) per mg total SAP eluted.	Amount of SAP eluted by EDTA (mg) per mg of total SAP eluted
10mM NaCl	0.007	0.993
30mM NaCl	0.08	0.912

Less SAP is eluted using 10mM NaCl than 30mM NaCl, but the amount eluted is still

much less than that by ribose-5-phosphate and AMP at the same concentrations. Therefore,

an increase in salt concentration has a slight effect on SAP elution which will be minimal

in terms of total SAP eluted using ribose-5-phosphate and AMP.
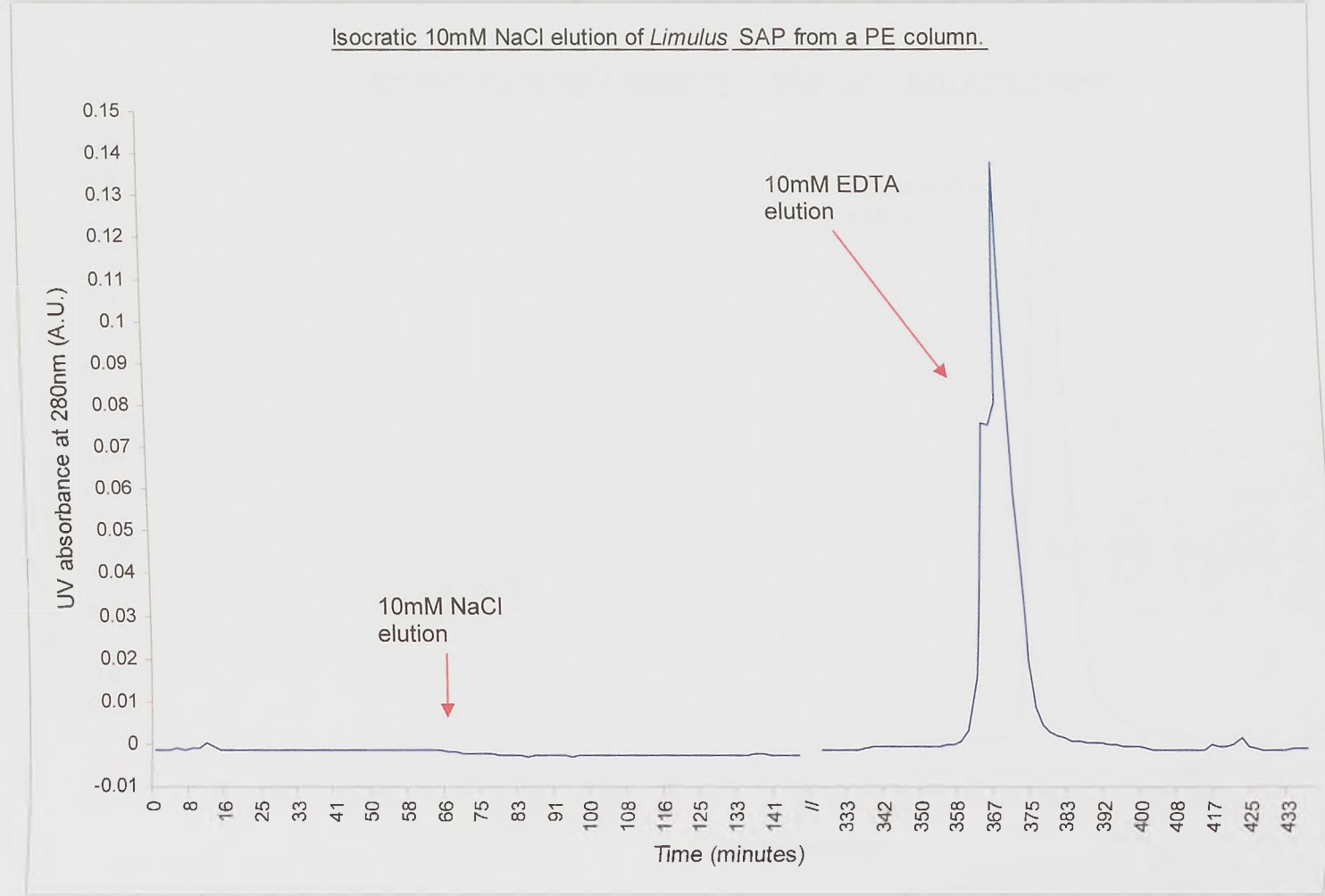
Figure 4.17. Elution profile of *Limulus* SAP eluted sequentially using 10mM NaCl then 10mM EDTA from a PE-agarose column. // indicates a break in time between elutions.
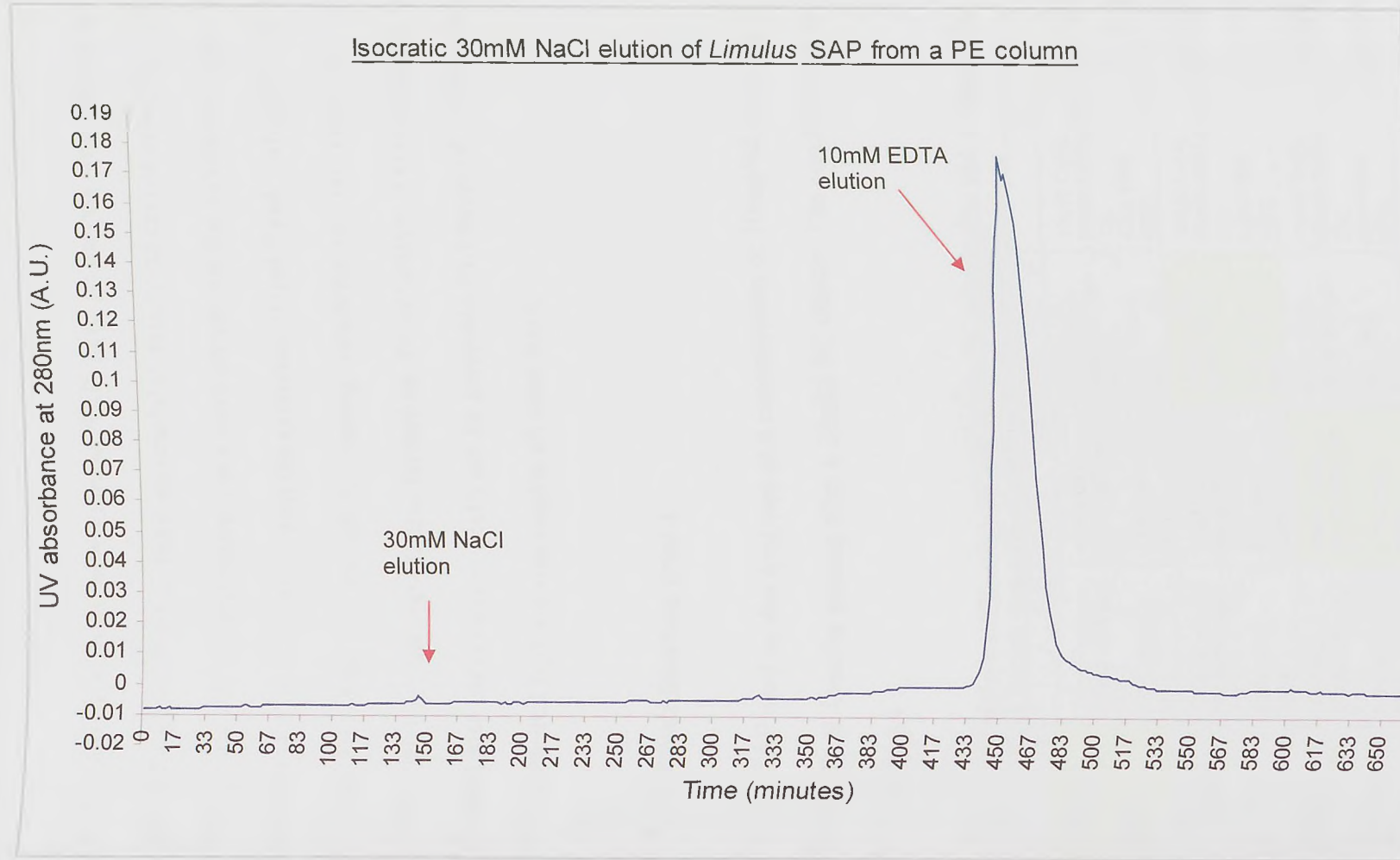
Figure 4.18. Elution profile of *Limulus* SAP eluted using 30mM NaCl followed by an elution of 10mM EDTA.

4.3.7 Results of crystallisation trials using *Limulus* SAP

Several crystallisation trials with *Limulus* SAP were performed with various buffers, precipitants, and ligands that had been successful in producing crystal growth in previous crystallisation trials. All crystallisation trials were set around the conditions: XpH 50mM MES, X% PEG 6000, 10mM CaCl$_2$, with the exception of trial 2 and 3 where the MES was replaced with Tris because of MES causing condensation on coverslips during crystallisation trial 1. Condensation also formed on the coverslips of crystallisation trial 2 and 3 indicating that condensation could not be prevented by change of buffer, and so as MES was the preferred buffer it was used in all other trays.

Limulus SAP crystallisation trial 1

The *Limulus* SAP used in this trial was at a concentration of 10mg/ml, obtained from a 30mM PE elution of *Limulus* plasma after a 10mM PC elution. The conditions of the tray are shown in table 4.4

Table 4.4. Table showing the conditions in each well of *Limulus* crystallisation trial 1. Highlighted areas show the wells from which crystals were taken to test for diffraction at SRS Daresbury.

	1	2	3	4	5	6
A	10mM CaCl$_2$, 50mM MES pH 7.0. 4.5% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.0. 5.0% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.0. 5.5% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.0. 6.0% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.0. 6.5% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.0. 7.0% PEG 6000.
B	10mM CaCl$_2$, 50mM MES pH 7.2. 4.5% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.2. 5.0% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.2. 5.5% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.2. 6.0% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.2. 6.5% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.2. 7.0% PEG 6000.
C	10mM CaCl$_2$, 50mM MES pH 7.5. 4.5% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.5. 5.0% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.5. 5.5% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.5. 6.0% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.5. 6.5% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.5. 7.0% PEG 6000.
D	10mM CaCl$_2$, 50mM MES pH 7.8. 4.5% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.8. 5.0% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.8. 5.5% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.8. 6.0% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.8. 6.5% PEG 6000.	10mM CaCl$_2$, 50mM MES pH 7.8. 7.0% PEG 6000.

Most wells showed growth of many tiny irregular shaped crystals some of which formed clumps. B2 and D3 showed growth of some small regular sharp sided crystals,

some in clumps and some single. C3 and B4 also showed growth of some regular shaped crystals and also showed growth of tiny bi-pyramids. C4 showed growth of some regular shaped crystals larger than the small irregular ones seen in most of the wells. Both B5 and C5 showed growth of small crystals regular in shape. A6 showed growth of some larger bi-pyramids than B4. Crystals from wells A1 (figure 4.19), B4 (figure 4.20), and C3 (figure 4.21) were used to test for diffraction. Crystals from wells B4 and C3 were used to test for diffraction at Daresbury Synchrotron Radiation Source station 14.1 in November 2005. Crystals from wells C3 and B4 were cryoprotected using cryobuffers containing MPD at concentrations 5%, 10%, 15% and 20%. Cryoprotection was achieved by adding 2μl of cryobuffer with 5% MPD into the liquid containing the crystals on the microbridge. This was repeated for 10%, 15% and 20% (x2) MPD cryobuffers at intervals of 10 minutes. 10μl of 20% MPD was then used to replace (exchanged with) 10μl of solution taken from the microbridge. The crystal from B4 disintegrated after addition of cryobuffer that contained 15% MPD. The crystal from well C3 was placed straight into the beam after exchange with 20% MPD cryobuffer and a snapshot of the crystal was taken at which showed diffraction to 20Å.
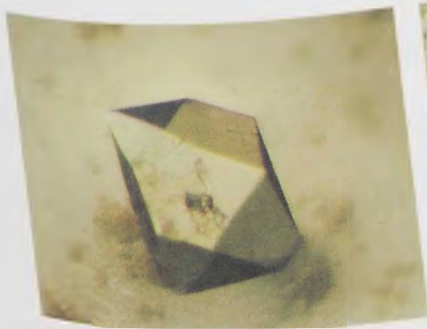
Figure 4.19. Crystal in well A1 of *Limulus* SAP crystallisation trial 1.

Figure 4.20. Crystal in well B4 of *Limulus* SAP crystallisation trial 1.

Figure 4.21. Crystals in well C3 of *Limulus* SAP crystallisation trial .

The crystal from well A1 was used to test for diffraction at Daresbury Synchrotron Radiation Source station 14.1 in July 2007. It was cryoprotected using cryobuffers containing PEG 400 at 5, 10, 20, and 30%, and 30mM AMP. 30mM AMP was included in the cryoprotectant as a ligand to soak into the crystals so that some diffraction occurring

would be that of SAP bound to PE. Cryoprotection was performed as mentioned previously for crystals in wells C3 and B4. After 15% PEG 400 cryobuffer was added the crystal started to disintegrate, so 4µl of 20% PEG 400 cryobuffer was added and after 4 minutes 10µl of 20% PEG 400 cryobuffer was exchanged with 10µl of microbridge solution. The crystal was frozen 5 minutes after the last cryobuffer was added, but the diffraction pattern showed the presence of ice rings suggesting that the crystal was not completely protected, and only showed diffraction to 9Å.

Eventually, all the wells in the tray dried up due to condensation forming on the coverslips which had to be replaced, and the wells mother liquor replaced by fresh mother liquor. Crystals that were used to obtain diffraction data were taken from the tray before all the wells dried up. There was no change in crystal growth in most of the wells when new mother liquor was added, except in wells A3, D2, C3, B4, D6 and A2. These showed crystal growth that was not present beforehand.

Limulus SAP crystallisation trial 2

Limulus SAP crystallisation trial 2 was the same as the first trial only with Tris buffer instead of MES, as the coverslips in the first tray had condensation on them and it was thought to be buffer related. *Limulus* SAP used in this trial was the same as that used in trial 1. The conditions of the tray are shown in table 4.5. The change of buffer from MES to Tris did not affect the presence of condensation on the coverslips and so eventually like *Limulus* crystallisation trial 1, all the wells in this tray had dried up and were replaced with 1ml of fresh mother liquor. This did not appear to encourage new crystal growth or affect existing crystals. Most wells showed precipitate formation or the growth of many tiny crystals. Flat crystals with regular shape were observed in well D2 (see figure 4.22). Clumps of small irregular crystals were observed in wells D1, D3, and C4, whilst growth of larger crystals was observed in C3 and C5. C3 showed growth of rod shaped crystals,

whilst C5 showed growth of bi-pyramid like crystals (see figure 4.23). No crystals in this

tray were used to obtain diffraction data.

Table 4.5. Table showing the conditions of each well in *Limulus* SAP crystallisation trial 2.

	1	2	3	4	5	6
A	10mM CaCl$_2$. 50mM Tris pH 7.0. 4.5% PEG 6000	10mM CaCl$_2$. 50mM Tris pH 7.0. 5.0% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.0. 5.5% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.0. 6.0% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.0. 6.5% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.0. 7.0% PEG 6000.
B	10mM CaCl$_2$. 50mM Tris pH 7.2. 4.5% PEG 6000	10mM CaCl$_2$. 50mM Tris pH 7.2. 5.0% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.2. 5.5% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.2. 6.0% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.2. 6.5% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.2. 7.0% PEG 6000.
C	10mM CaCl$_2$. 50mM Tris pH 7.5. 4.5% PEG 6000	10mM CaCl$_2$. 50mM Tris pH 7.5. 5.0% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.5. 5.5% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.5. 6.0% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.5. 6.5% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.5. 7.0% PEG 6000.
D	10mM CaCl$_2$. 50mM Tris pH 7.8. 4.5% PEG 6000	10mM CaCl$_2$. 50mM Tris pH 7.8. 5.0% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.8. 5.5% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.8. 6.0% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.8. 6.5% PEG 6000.	10mM CaCl$_2$. 50mM Tris pH 7.8. 7.0% PEG 6000.
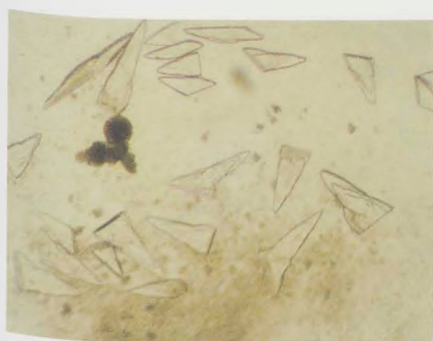
Figure 4.22. Crystals from well D2 of *Limulus* SAP crystallisation trial 2.

Figure 4.23. Crystals from well C5 of *Limulus* SAP crystallisation trial 2.

Limulus SAP crystallisation trial 3

This trial was an exact replica of *Limulus* SAP crystallisation trial 2 but using SAP at a concentration of 8mg/ml not 10mg/ml, as the concentration 8mg/ml had been used in previous crystal trials with *Limulus* SAP. Most wells showed formation of precipitate and no crystal growth was observed. However, wells C1 and D1 showed growth of small regular shaped crystals, and well D2 had a small bi-pyramid covered in precipitate. As with the first two trays, the mother liquor had to be replaced due to evaporation and condensation on the coverslips. Replacement of mother liquor only encouraged new crystal growth in well C1 which was that of the small regular shaped crystals mentioned. No crystals grown in this tray were used to obtain diffraction data.

Limulus SAP crystallisation trials 4-7

The following trials were all set around the conditions in trial 1 that showed promising crystal growth, wells A1, B2, and C5.

Limulus SAP crystallisation trial 4

This trial used *Limulus* SAP at a concentration of 12.8mg/ml, isolated from PEG cut *Limulus* plasma using 30mM PE after a 10mM PC elution. The change of concentration from 8mg/ml to 12mg/ml was performed to see the effect of increasing the protein concentration to higher than 10mg/ml that had been used in the trails 1 and 2. This trial was an extension of the first SAP trial with the same conditions but of percentage PEG 6000 from 5.0% to 7.5% rather than from 4.5% to 7%. The conditions are shown in table 4.6.

Table 4.6. Table showing the conditions of each well in *Limulus* SAP crystallisation trial 4.

	1	2	3	4	5	6
A	10mM $CaCl_2$. 50mM MES pH 7.0. 5.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 5.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 6.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 6.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 7.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 7.5% PEG 6000.
B	10mM $CaCl_2$. 50mM MES pH 7.2. 5.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.2. 5.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.2. 6.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.2. 6.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.2. 7.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.2. 7.5% PEG 6000.
C	10mM $CaCl_2$. 50mM MES pH 7.5. 5.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 5.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 6.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 6.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 7.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 7.5% PEG 6000.
D	10mM $CaCl_2$. 50mM MES pH 7.8. 5.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.8. 5.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.8. 6.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.8. 6.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.8. 7.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.8. 7.5% PEG 6000.
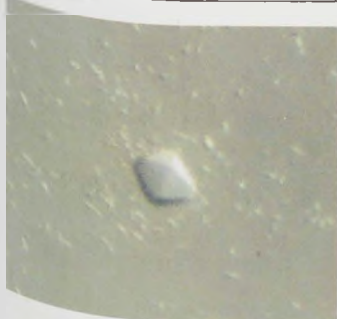
Figure 4.24 Crystal from well B4 of *Limulus* SAP crystallisation trial 4.

Most wells of this tray showed formation of precipitate and no crystal growth. Wells B1, D1, D2, A2, and A3 showed growth of tiny irregular shaped crystals similar to trays 1-2. Well B4 showed growth of clusters of flat crystals with a few small irregular shaped crystals and tiny bi-pyramids (see

figure 4.24) similar to wells D2 and C5 from trial 2. No crystals grown in this trial were used to test for diffraction.

Limulus SAP crystallisation trial 5

This trial used *Limulus* SAP at a concentration of approximately 7.8mg/ml, isolated using a 30mM PE elution from *Limulus* plasma after a 10mM PC elution. The concentration was reduced from that of SAP used in trial 4 (12mg/ml) to 7.8mg/ml as previous crystal trials were performed with *Limulus* SAP at a concentration of 8mg/ml, and crystal trial 4 showed no promising growth of crystals. The trial concentrated around trial 1 conditions in well A1 as it had good crystal growth. The pH of MES was 6.8, 7.0, 7.2, and 7.4 instead of 7.0 to 7.8 making the pH range smaller so exact differences between small pH changes could be seen, and the percentage of PEG 6000 varied from 4% to 5%. The conditions of the tray are shown in table 4.7.

Table 4.7. Table showing the conditions of the wells in *Limulus* SAP crystallisation trial 5.

	1	2	3	4	5
A	10mM CaCl$_2$. 50mM MES pH 6.8. 4.0% PEG 6000	10mM CaCl$_2$. 50mM MES pH 6.8. 4.2% PEG 6000	10mM CaCl$_2$. 50mM MES pH 6.8. 4.5% PEG 6000	10mM CaCl$_2$. 50mM MES pH 6.8. 4.7% PEG 6000	10mM CaCl$_2$. 50mM MES pH 6.8. 5.0% PEG 6000
B	10mM CaCl$_2$. 50mM MES pH 7.0. 4.0% PEG 6000	10mM CaCl$_2$. 50mM MES pH 7.0. 4.2% PEG 6000	10mM CaCl$_2$. 50mM MES pH 7.0. 4.5% PEG 6000	10mM CaCl$_2$. 50mM MES pH 7.0. 4.7% PEG 6000	10mM CaCl$_2$. 50mM MES pH 7.0. 5.0% PEG 6000
C	10mM CaCl$_2$. 50mM MES pH 7.2. 4.0% PEG 6000	10mM CaCl$_2$. 50mM MES pH 7.2. 4.2% PEG 6000	10mM CaCl$_2$. 50mM MES pH 7.2. 4.5% PEG 6000	10mM CaCl$_2$. 50mM MES pH 7.2. 4.7% PEG 6000	10mM CaCl$_2$. 50mM MES pH 7.2. 5.0% PEG 6000
D	10mM CaCl$_2$. 50mM MES pH 7.4. 4.0% PEG 6000	10mM CaCl$_2$. 50mM MES pH 7.4. 4.2% PEG 6000	10mM CaCl$_2$. 50mM MES pH 7.4. 4.5% PEG 6000	10mM CaCl$_2$. 50mM MES pH 7.4. 4.7% PEG 6000	10mM CaCl$_2$. 50mM MES pH 7.4. 5.0% PEG 6000

Most wells showed precipitate formation and growth of small and regular/irregularly shaped crystals (see figures 4.25 and 4.26). Well D1 had some fairly large crystals (figure 4.27), one of which was flat and regularly shaped, whilst other crystals in the well were irregular in shape. No crystals grown in this tray were used to test for diffraction.
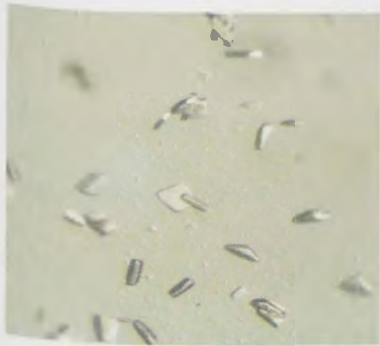
Figure 4.25 Crystals grown in well A2 of *Limulus* SAP crystallisation trial 5.
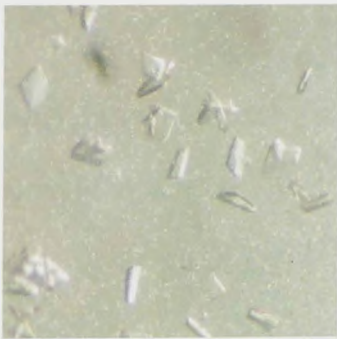
Figure 4.26. Crystals grown in well A3 of *Limulus* SAP crystallisation trial 5.
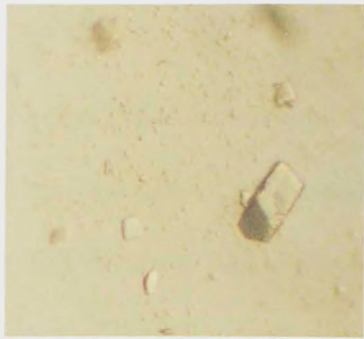
Figure 4.27. Crystals grown in well D1 of *Limulus* SAP crystallisation trial 5

Limulus SAP crystallisation trial 6

This trial was concentrating on the conditions around well B2 of trial 1 which showed some promising crystal growth of regular crystals. The conditions were 10mM $CaCl_2$, 50mM MES pH 7.2 to 7.5, and PEG 6000 4.5-5.5%. The *Limulus* SAP used in this trial was the same as that used in trial 5. The conditions of the tray are shown in table 4.8.

Table 4.8. Table showing the conditions in each well of *Limulus* SAP crystallisation trial 6.

	1	2	3	4	5
A	10mM $CaCl_2$. 50mM MES pH 7.2. 4.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.2. 4.7% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.2. 5.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.2. 5.2% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.2. 5.5% PEG 6000
B	10mM $CaCl_2$. 50mM MES pH 7.3. 4.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.3. 4.7% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.3. 5.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.3. 5.2% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.3. 5.5% PEG 6000
C	10mM $CaCl_2$. 50mM MES pH 7.4. 4.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.4. 4.7% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.4. 5.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.4. 5.2% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.4. 5.5% PEG 6000
D	10mM $CaCl_2$. 50mM MES pH 7.5. 4.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 4.7% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 5.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 5.2% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 5.5% PEG 6000

All wells contained precipitate and many tiny irregular shaped crystals. No crystals were large enough to test for diffraction.

Limulus SAP crystallisation trial 7

This trial was an extension of trial 1, but concentrating on the conditions around well C5 which showed crystal growth in crystallisation trials 1 and 2. *Limulus* SAP that was used in this trial was the same as that used in trials 5 and 6. The conditions were 10mM $CaCl_2$, 50mM MES pH 7.5-7.8, and PEG 6000 6% to 7%. The conditions of the tray are shown in table 4.9.

Table 4.9. Table showing the conditions of each well in *Limulus* SAP crystallisation trial 7. Highlighted areas show the wells from which crystals were taken to use for testing at SRS Daresbury.

	1	2	3	4	5
A	10mM $CaCl_2$. 50mM MES pH 7.5. 6.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 6.2% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 6.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 6.7% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.5. 7.0% PEG 6000.
B	10mM $CaCl_2$ 50mM MES pH 7.6. 6.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.6. 6.2% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.6. 6.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.6. 6.7% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.6. 7.0% PEG 6000.
C	10mM $CaCl_2$. 50mM MES pH 7.7. 6.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.7. 6.2% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.7. 6.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.7. 6.7% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.7. 7.0% PEG 6000.
D	10mM $CaCl_2$. 50mM MES pH 7.8. 6.0% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.8. 6.2% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.8. 6.5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.8. 6.7% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.8. 7.0% PEG 6000.

The wells which had small regular and irregular shaped crystals in were: A1, A2, A3 (figure 4.28), B3, A4, B4, A5 and B5, which showed that perhaps the conditions best for crystal growth were at pH 7.5 but that the percentage of PEG 6000 did not make any difference to crystal growth. Well D2 (figure 4.29) had growth of long rod shaped crystals. Other wells showed precipitate and no crystal growth. A long rod shaped crystal from D2 was used to obtain diffraction data at Daresbury Synchrotron Radiation Source station 14.1 in March 2007. The crystal was cryoprotected using 5, 10%, 15% and 20% MPD as mentioned previously for *Limulus* SAP crystallisation trial 1. The crystals used for diffraction were approximately 200μm long and had diffraction in some parts to 3.7Å (see figure 4.30).
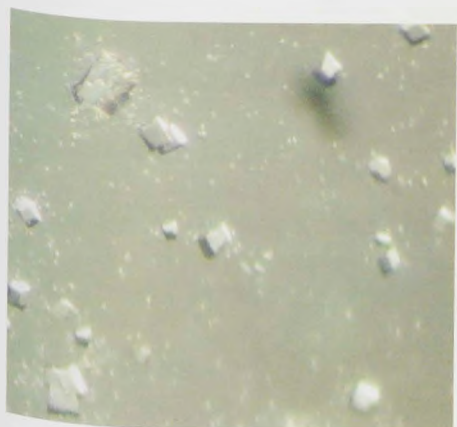
Figure 4.28. Crystals from well A3 from *Limulus* SAP crystallisation trial 7

Figure 4.29 Crystals from well D2 of *Limulus* SAP crystallisation trial 7.
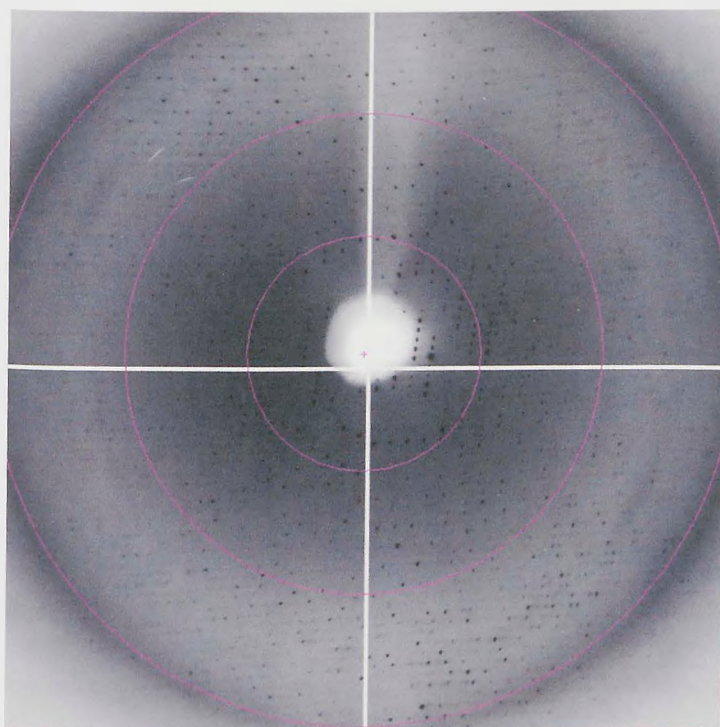
Figure 4.30 Zoomed in part of the diffraction pattern from a crystal in well D2 *Limulus* SAP crystal trial 7. The coloured rings identify where the diffraction is at 11.2, 5.6, and 3.7Å respectively from inner to outer circle.

***Limulus* SAP crystallisation trials 8-11**

Four further crystallisation trials (8-11) were performed using *Limulus* SAP with conditions that varied around the conditions on trial 1. The pH of MES varied around 7.0, the percentage of PEG 6000 varied around 7%, and the concentration of $CaCl_2$ varied around 10mM. All trays used *Limulus* SAP isolated from *Limulus* plasma using 30mM PE after a 10mM PC elution, at a concentration of 8mg/ml. The conditions of each tray are shown in tables 4.10 to 4.13.

***Limulus* SAP crystallisation trial 8**

Table 4.10. Table showing conditions in each well of *Limulus* SAP crystallisation trial 8.

	1	2	3	4	5	6
A	6mM $CaCl_2$. 50mM MES pH 6.5. 3% PEG 6000.	6mM $CaCl_2$. 50mM MES pH 6.5. 4% PEG 6000.	6mM $CaCl_2$. 50mM MES pH 6.5. 5% PEG 6000.	6mM $CaCl_2$. 50mM MES pH 6.5. 6% PEG 6000.	6mM $CaCl_2$. 50mM MES pH 6.5. 7% PEG 6000.	6mM $CaCl_2$. 50mM MES pH 6.5. 8% PEG 6000.
B	10mM $CaCl_2$. 50mM MES pH 7.0. 3% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 4% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 5% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 6% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 7% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 8% PEG 6000.
C	12mM $CaCl_2$. 50mM MES pH 7.5. 3% PEG 6000.	12mM $CaCl_2$. 50mM MES pH 7.5. 4% PEG 6000.	12mM $CaCl_2$. 50mM MES pH 7.5. 5% PEG 6000.	12mM $CaCl_2$. 50mM MES pH 7.5. 6% PEG 6000.	12mM $CaCl_2$. 50mM MES pH 7.5. 7% PEG 6000.	12mM $CaCl_2$. 50mM MES pH 7.5. 8% PEG 6000.

Most wells contained precipitate or tiny irregular shaped crystals. Wells B1, C3, C4 and B6 showed growth of small thin rod shaped crystals (see figure 4.31). Crystals in well B6 looked similar to those in well D2 of *Limulus* SAP crystallisation trial 7 but much smaller. Well C6 showed crystal growth in small spiky clumps (see figure 4.32). Crystals in well B6 were too small, and those in well C6 were very irregular and therefore neither were suitable for use in diffraction testing.
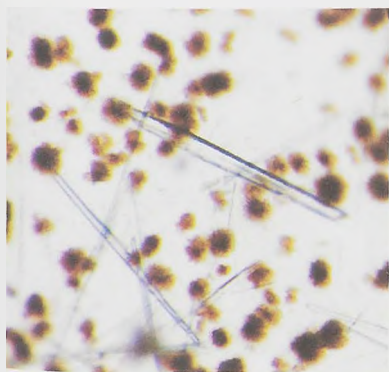
Figure 4.31. Crystals grown in well B6 of *Limulus* SAP crystallisation trial 8.
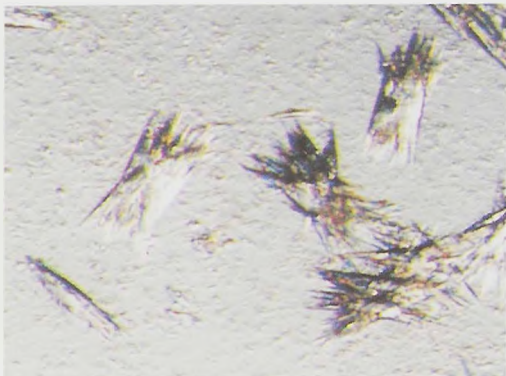
Figure 4.32. Crystals grown in well C6 of *Limulus* SAP crystallisation trial 8.

Limulus SAP crystallisation trial 9

Table 4.11. Table showing conditions in each well of *Limulus* SAP crystallisation trial 9

	1	2	3	4	5	6
A	6mM CaCl$_2$. 50mM MES pH 6.5. 9% PEG 6000.	6mM CaCl$_2$. 50mM MES pH 6.5. 10% PEG 6000.	6mM CaCl$_2$ 50mM MES pH 6.5. 11% PEG 6000.	6mM CaCl$_2$. 50mM MES pH 6.5. 12% PEG 6000.	6mM CaCl$_2$. 50mM MES pH 6.5. 13% PEG 6000.	6mM CaCl$_2$. 50mM MES pH 6.5. 14% PEG 6000.
B	10mM CaCl$_2$. 50mM MES pH 7.0. 9% PEG 6000.	10mM CaCl$_2$. 50mM MES pH 7.0. 10% PEG 6000.	10mM CaCl$_2$. 50mM MES pH 7.0. 11% PEG 6000.	10mM CaCl$_2$. 50mM MES pH 7.0. 12% PEG 6000.	10mM CaCl$_2$. 50mM MES pH 7.0. 13% PEG 6000.	10mM CaCl$_2$. 50mM MES pH 7.0. 14% PEG 6000.
C	12mM CaCl$_2$. 50mM MES pH 7.5. 9% PEG 6000.	12mM CaCl$_2$. 50mM MES pH 7.5. 10% PEG 6000.	12mM CaCl$_2$. 50mM MES pH 7.5. 11% PEG 6000.	12mM CaCl$_2$. 50mM MES pH 7.5. 12% PEG 6000.	12mM CaCl$_2$. 50mM MES pH 7.5. 13% PEG 6000.	12mM CaCl$_2$. 50mM MES pH 7.5. 14% PEG 6000.

Wells A1-A3 (figure 4.33) showed crystal growth similar to well B6 of *Limulus* SAP crystallisation trial 8. Well B3 also showed the growth of long rods shaped crystals (figure 4.34), as did well B4 although there was also the presence of smaller irregular crystals (figure 4.35). Wells C1 and C2 (figure 4.36) showed crystals of spiky irregular shape similar to those in well C6 of *Limulus* SAP crystallisation trial 8. Crystals grown in wells C4 and C6 were more regular in shape (figure 4.37).
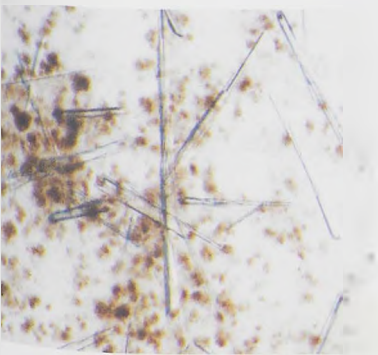
Figure 4.33. Crystals from well A1 of *Limulus* SAP crystallisation trial 9.

Figure 4.34. Crystals from well B3 of *Limulus* SAP crystallisation trial 9
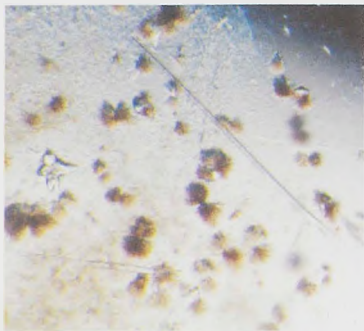
Figure 4.35. Crystals from well B4 of *Limulus* SAP crystallisation trial 9.
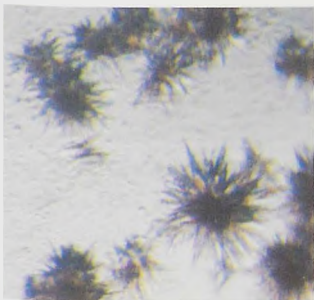
Figure 4.36. Crystals from well C2 of *Limulus* SAP crystallisation trial 9.
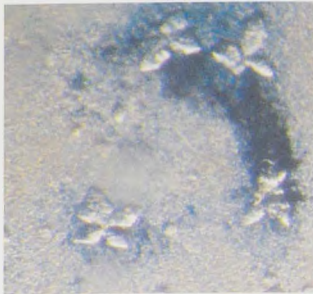
Figure 4.37. Crystals from well C6 of *Limulus* SAP crystallisation trial 9.

Limulus SAP crystallisation trial 10

Table 4.12. Table showing conditions in each well of *Limulus* SAP crystallisation trial 10.

	1	2	3	4	5	6
A	8mM CaCl$_2$. 50mM MES pH 6.5. 3% PEG 6000.	8mM CaCl$_2$. 50mM MES pH 6.5. 4% PEG 6000.	8mM CaCl$_2$. 50mM MES pH 6.5. 5% PEG 6000.	8mM CaCl$_2$. 50mM MES pH 6.5. 6% PEG 6000.	8mM CaCl$_2$. 50mM MES pH 6.5. 7% PEG 6000.	8mM CaCl$_2$. 50mM MES pH 6.5. 8% PEG 6000.
B	10mM CaCl$_2$. 50mM MES pH 7.0. 3% PEG 6000.	10mM CaCl$_2$. 50mM MES pH 7.0. 4% PEG 6000.	10mM CaCl$_2$. 50mM MES pH 7.0. 5% PEG 6000.	10mM CaCl$_2$. 50mM MES pH 7.0. 6% PEG 6000.	10mM CaCl$_2$. 50mM MES pH 7.0. 7% PEG 6000.	10mM CaCl$_2$ 50mM MES pH 7.0. 8% PEG 6000.
C	14mM CaCl$_2$. 50mM MES pH 7.5. 3% PEG 6000.	14mM CaCl$_2$. 50mM MES pH 7.5. 4% PEG 6000.	14mM CaCl$_2$. 50mM MES pH 7.5. 5% PEG 6000.	14mM CaCl$_2$. 50mM MES pH 7.5. 6% PEG 6000.	14mM CaCl$_2$. 50mM MES pH 7.5. 7% PEG 6000.	14mM CaCl$_2$. 50mM MES pH 7.5. 8% PEG 6000.

Most wells showed small or large rod shaped crystals. Wells A1, C3, B5 and A6 showed growth of long rod shaped crystals (see figure 4.38 and 4.39) similar to those found in wells D2 from crystal trial 7. Wells B1, A2, B2, A3, B3, A4, B4, and C4 showed growth of

small rod shaped crystals and small regular shaped crystals (see figures 4.40 and 4.41).

Wells C4, A5, B6 and C6 showed growth of spiky crystal clumps similar to those of well

C2 crystal trial 9. Larger more regular shaped crystal growth was observed in wells A1 and

C1. No crystals were used in diffraction testing.
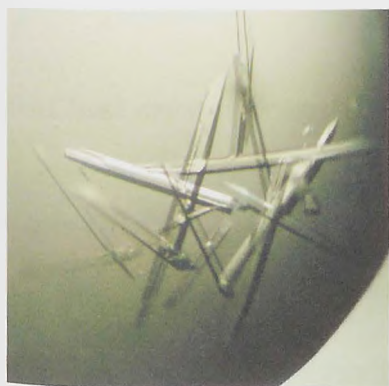
Figure 4.38 Crystals from well C3
of *Limulus* SAP crystallisation
trial 10.
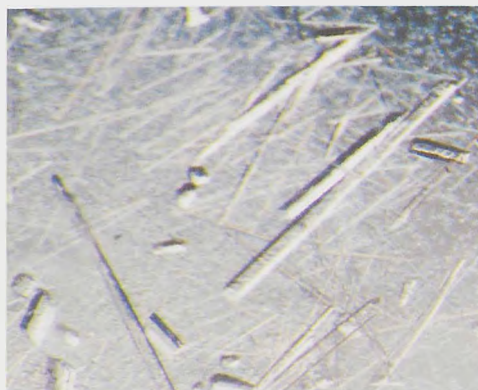
Figure 4.39 Crystals from well A1 of
Limulus SAP crystallisation trial 10.

Figure 4.40 Crystals from well B1 of
Limulus SAP crystallisation trial 10.

Figure 4.41. Crystals from well B2 of
Limulus SAP crystallisation trial 10.

Limulus SAP crystallisation trial 11

Table 4.13. Table showing conditions in each well of *Limulus* SAP crystallisation trial 11.

	1	2	3	4	5	6
A	8mM $CaCl_2$. 50mM MES pH 6.5. 9% PEG 6000.	8mM $CaCl_2$. 50mM MES pH 6.5. 10% PEG 6000.	8mM $CaCl_2$ 50mM MES pH 6.5. 11% PEG 6000.	8mM $CaCl_2$ 50mM MES pH 6.5. 12% PEG 6000.	8mM $CaCl_2$ 50mM MES pH 6.5. 13% PEG 6000.	8mM $CaCl_2$. 50mM MES pH 6.5. 14% PEG 6000.
B	10mM $CaCl_2$. 50mM MES pH 7.0. 9% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 10% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 11% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 12% PEG 6000.	10mM $CaCl_2$ 50mM MES pH 7.0. 13% PEG 6000.	10mM $CaCl_2$. 50mM MES pH 7.0. 14% PEG 6000.
C	14mM $CaCl_2$. 50mM MES pH 7.5. 9% PEG 6000.	14mM $CaCl_2$. 50mM MES pH 7.5. 10% PEG 6000.	14mM $CaCl_2$ 50mM MES pH 7.5. 11% PEG 6000.	14mM $CaCl_2$. 50mM MES pH 7.5. 12% PEG 6000.	14mM $CaCl_2$ 50mM MES pH 7.5. 13% PEG 6000.	14mM $CaCl_2$. 50mM MES pH 7.5. 14% PEG 6000.

Most wells (B1, A2, B2, C2, A3, B3, C3, A4, B4, and A5) showed growth of small rod shaped crystals as seen previously in trials 8-10. Some wells also showed growth of clusters of spiky and irregular shaped crystals. No crystals were suitable for use in testing for diffraction data.

Areas of best crystal growth

The areas that showed crystal growth dependent on the concentration of PEG 6000 and calcium chloride, and the pH of MES at 50mM are shown in tables 4.14 and 4.15. The information was collated together using data from all crystal trays, and areas of conditions that were not tried are shown. It seems that the best concentration of PEG 6000 for growing spiky crystals was at 6-10% and to grow bi-pyramids was a 5%-7% PEG concentration. Both rod shaped crystals and regular shaped crystals seemed to grow through the full range of PEG 6000 concentrations, although rod shaped crystals seemed to be more abundant with PEG 6000 concentrations higher than 7.5, and regular crystals seemed to be more abundant at PEG 6000 concentrations 6.5 and lower.

Most rods shaped crystals grew at 50mM MES pH 6.5 to 7.0, although some did grow at higher pH values of 7.5 and 7.8. Spiky crystals also grew at both pH 7.0 and 7.5 but not beyond these values, although an increase in the number of crystallisation trails around all pH values may have shown otherwise. Regular shaped crystals however, grew at pH values 6.5 to 7.8 and showed no apparent preference of pH. There appears to be much crystal growth at pH 7 and 7.5 which may be due to many repetitions of crystal trials in this area.

Concentration of PEG 6000

pH of 50mM MES	3.0	4.0	4.2	4.5	4.7	5.0	5.2	5.5	6.0	6.2	6.5	6.6	6.7	7.0	7.5	8.0	9.0	10.0	11.0	12.0	13.0	14.0
6.5	□■	□				□		□								□	□	□	□	□	□	
6.8																						
7	□	□		■		□			□		✧			□◇		□✧	□	□	□	□		
7.2						■			■◇		■◇											
7.3																						
7.4		■																				
7.5	■					□		■	□✧■		■					✧	✧	□✧	□	■	■	■
7.6																						
7.7																						
7.8						■◇		■		□												

Crystals used for diffraction
Rod shaped crystals □
Spiky crystals ✧
Regular shaped crystals ■
Bi-pyramids ◇
Small irregular and regular shaped crystals
Precipitate
Conditions not tried

Table 4.14. Table showing the best areas of crystal growth for *Limulus* SAP crystallisation trials 1-11 dependent on the concentration of PEG 6000 and the pH of MES at 50mM concentration.

pH of 50mM MES	Concentration of calcium chloride (mM).				
	6	8	10	12	14
6.5	□	□✧■			
6.8					
7			□✧■		
7.2			■◇		
7.3					
7.4			■		
7.5			□✧■◇	□✧■	□✧■
7.6					
7.7					
7.8			□■◇		

Conditions not tried

Crystals used for diffraction

Rod shaped crystals □

Spiky crystals ✧

Regular shaped crystals ■

Bi-pyramids ◇

Precipitate

Table 4.15. Table showing the best areas of crystal growth for Limulus SAP crystallisation trials 1-11 dependent on the concentration of calcium chloride and the pH of 50mM MES within the wells.

It appears that the best concentration of calcium chloride for crystal growth is at 10mM although much crystal growth is observed at both higher and lower concentrations. From observation of table 4.15, it appears that good conditions for crystal growth are 50mM MES pH 7.0, 7.5 and 7.8, and at 8mM, 10mM, 12mM and 14mM calcium chloride. The most variety in crystal shape grew in conditions of 50mM MES pH 7.5 and 10mM calcium chloride. Coupled with the results from table 4.14, it can be suggested that the most varied crystal growth appears in conditions 50mM MES pH 7.5, 10mM calcium chloride, and 6% PEG 6000. These results may differ if more crystallisation conditions were tried.

4.4 Discussion

4.4.1. Whole *Limulus* SAP

Limulus SAP is present at approximately 0.3mg/ml of haemolymph in these experiments which is slightly higher than the 0.12-0.17mg/ml range identified by Shrive *et al.*, (1999), and may be due to fluctuations in SAP plasma levels from pooled haemolymph of horseshoe crabs taken at different periods of time. The presence of different sized subunits of SAP was indicated by SDS-PAGE which showed three bands approximately 29kDa in size. Deglycosylation of SAP did not reduce the number of bands on SDS-PAGE, but did appear to reduce the size of the top two bands indicating that these subunits were glycosylated. Of the three SAP bands ~29kDa in size, deglycosylation appeared to increase the proportion of the middle band compared to non-deglycosylated SAP as indicated by darker staining of this band. This would suggest there was more of this SAP subunit after deglycosylation, than the other two subunits. The subunits may form different molecular aggregate SAP forms each composed of a single subunit type, or may form molecular aggregate SAP forms composed of different subunit types. As the structure shows a single subunit type, it is likely that the subunits only come together with like subunits. There is also the possibility that the different sizes of subunits on SDS-PAGE after deglycosylation are due to other modifications or contamination from other proteins such as CRP, which may still be present in trace amounts after elution using 10mM PC.

Denaturation of *Limulus* SAP followed by deglycosylation resulted in smaller molecular weight bands on SDS-PAGE than native SAP and native deglycosylated SAP. There appeared to be a reduction in the number of bands in denatured and deglycosylated SAP from three in non-denatured and non-deglycosylated to two, which may have been due to the darker staining of two bands and relatively lighter staining of another. An overall reduction in the size of denatured and deglycoyslated SAP was probably due to both a

removal of glycosylated groups and an alteration of shape by denaturation allowing easier movement of subunits through the gel.

Analysis of SAP by surface hydrophobicity indicated the presence of four peaks on the FPLC elution trace, of which peaks 2 and 4 were shoulders on the third and largest peak. Therefore, there is a possibility of at least four SAP molecular aggregate forms which differ in surface hydrophobicity. Analysis of deglycosylated SAP by surface hydrophobicity, indicated the presence of only three (1, 3 and 4) peaks on the FPLC elution trace compared four peaks in non-deglycosylated SAP. This suggested a reduction by deglycosylation in the number of molecular aggregate forms of SAP that differ in surface hydrophobicity. Peak 2 is likely to correspond to glycosylated SAP that has a similar surface hydrophobicity to SAP composing either peaks 1, 3, or 4 after removal of the glycosylated group. It is possible that as peak 3 is the largest of the peaks identified in the elution traces for surface hydrophobicity, the protein composing this peak corresponds to the darkest stained band for deglycosylated SAP on SDS-PAGE.

The presence of SAP molecular aggregate forms that vary by glycosylation and surface hydrophobicity in the haemolymph of *Limulus polyphemus*, suggest a family of SAP proteins which may have different binding roles within the horseshoe crab but similar overall function. There is also a possibility that specific molecular aggregate forms are expressed in individual horseshoe crabs for a specific purpose, although this cannot be verified in this study as *Limulus* plasma used was pooled from many individuals. This is now being followed up.

4.4.2 *Limulus* SAP PE binding

Limulus SAP molecular aggregate and subunit forms were not separated using a 30mM

PE gradient as indicated by the elution trace and SDS-PAGE analysis. This suggests that SAP molecular aggregate forms have a similar binding affinity for PE with each other, or at least not dissimilar enough to be separated using a 30mM PE gradient. A possible way to achieve separation may have been to use a lower concentration gradient e.g. 10mM, over an extended period of time. However, lower concentrations and prolonged gradient elutions using ligands for CRP did not appear to provide separation of CRP molecular aggregate forms (see chapter 3) and this may also be the case for *Limulus* SAP.

4.4.3 *Limulus* SAP elution using AMP and ribose-5-phosphate

Limulus SAP was eluted from a PE column using both AMP and ribose-5-phosphate, suggesting that these ligands either bind to SAP at the PE binding site or interfere with it in some way such as binding near it so that SAP dissociates from PE. At 10mM concentration, AMP eluted approximately 10% of SAP from the column compared to EDTA which eluted the remaining 90%. At 30mM AMP concentration, the percentage of SAP eluted using AMP was approximately 96% compared to 4% eluted afterwards with EDTA. Similar values were obtained with 10mM and 30mM ribose-5-phosphate where the percentage of SAP removed was approximately 15% and 92% respectively. Such a small change in elution buffer concentration leading to such a large difference in the amount of SAP eluted between 10mM and 30mM ribose-5-phosphate and AMP was not expected, especially considering the same change in concentration does not greatly alter the amount of CRP eluted using these ligands at 10mM and 30mM. There are two possible suggestions for these results, either a) the binding affinity of ~10-15% of SAP for ribose-5-phosphate and AMP is high so that it is eluted at 10mM concentration whilst the remaining SAP has a lower binding affinity and so is eluted at higher ligand concentration (30mM), or b) all SAP binds at equal affinity for AMP and ribose-5-phosphate but 30mM ligand concentration is required to elute the majority of SAP within the time given during elution.

Should the latter be true, an extended elution time using 10mM AMP or ribose-5-phosphate should elute a higher proportion of SAP.

SDS-PAGE analysis of SAP eluted at 10mM and 30mM ribose-5-phosphate showed a difference in subunit composition, with the SAP eluted using 10mM ribose-5-phosphate showing darker stained lower molecular weight bands of the three bands ~29kDa in size compared to SAP eluted using 30mM ribose-5-phosphate which showed more equal band staining but with the higher molecular weight band of the three bands ~29kDa stained darkest similar to whole SAP. A similar SDS-PAGE result was seen for SAP eluted using 10mM and 30mM AMP. The SAP eluted at 30mM ligand concentration had similar staining on SDS-PAGE to whole SAP which was to be expected as the majority of SAP was eluted at 30mM ligand concentration. SDS-PAGE results support the proposal that a) ~10-15% of SAP has higher binding affinity for ribose-5-phosphate and AMP at 10mM than the remaining SAP because of larger presence of one subunit form of SAP over the others at 10mM ligand concentration. This indicates more of one molecular aggregate form of SAP that may be composed entirely or mostly of this subunit, which therefore may have higher affinity for ribose-5-phosphate and AMP.

Surface hydrophobicity analysis of SAP eluted using 10mM ribose-5-phosphate showed the presence of the four main peaks on the elution trace also present in whole SAP. Peak one appeared much larger than peaks 2 and 3 on the elution trace observed with SAP that had been eluted using 10mM ribose-5-phosphate, whilst on the elution trace observed with whole SAP, peak one was much smaller than peaks 2 and 3. An increase in comparative peak 1 size may correspond to the darker staining of the lower molecular weight SAP band seen in SDS-PAGE and indicate increase in concentration of this SAP molecular aggregate form, should the whole aggregate be composed of like subunits. Surface hydrophobicity analysis of the SAP eluted using EDTA following 10mM ribose-5-phosphate elution, showed that the peak sizes on the elution trace are much more similar to those found in

whole SAP, and all four peaks were present. Therefore, 10mM ribose-5-phosphate did not elute SAP of one molecular aggregate form in particular that varied in surface hydrophobicity but did show some partial separation.

Surface hydrophobicity analysis of SAP eluted using 10mM AMP showed the presence of three of the four peaks identified on the elution trace of whole SAP. This indicated that 10mM AMP was eluting three particular molecular aggregate forms of SAP, which shows some specificity of SAP for AMP. As with SAP that was eluted using 10mM ribose-5-phosphate, the proportional size of peak 1 on the elution trace is larger than the other peaks suggesting a higher proportion of SAP corresponding to this peak than in whole SAP. This may be related to the darker staining of the lower molecular weight band of the three bands ~29kDa in SDS-PAGE of SAP eluted using 10mM AMP. Surface hydrophobicity analysis of the SAP eluted using EDTA after the 10mM AMP elution, showed the presence of the four main peaks on the elution trace as identified in whole SAP. All peaks are in similar proportions to those found in the elution trace for whole SAP. Therefore, 10mM AMP did not elute any one molecular aggregate form of SAP in particular that varied in surface hydrophobicity but did show some partial separation. The discrimination between peaks 2 and 3 on the elution traces for surface hydrophobicity analysis of SAP eluted using 10mM ribose-5-phosphate and AMP was difficult due to the closeness of the peaks. It could not be determined for sure which peaks they corresponded to on the elution trace for whole SAP.

Therefore, there appeared to be some partial separation of *Limulus* SAP molecular aggregate forms binding to 10mM ribose-5-phosphate and AMP as indicated by a slightly different band staining on SDS-PAGE, and changes in surface hydrophobicity by FPLC. However, as 10mM ribose-5-phosphate and AMP eluted only a small amount of SAP which had to be concentrated to apply to the HIC, there is a possibility that trace impurities such as haemocyanin and other plasma proteins appear as a peak(s) on the FPLC trace. In

all four traces shown for SAP isolated using 10mM ribose-5-phosphate and AMP, and the subsequent EDTA elutions, peak one appears comparatively larger than it does on the FPLC trace for whole SAP, and appears at its largest on the FPLC trace for the SAP eluted using 10mM AMP and ribose-5-phosphate. This would suggest that, as the SAP eluted using ribose-5-phosphate and AMP are concentrated, the comparative size of protein corresponding to peak 1 against peaks 2-4 gets larger indicating that it corresponds to trace impurity becoming concentrated. As the surface hydrophobicity analysis was only performed once for each sample, it is possible that multiple analyses may show the absence of the smaller peaks or shoulders which also may indicate trace impurities. Work is underway to determine which protein forms each of the peaks present on the FPLC traces.

Salt elutions of SAP at 10mM and 30mM confirm that SAP elution is due to ligand-protein interaction rather than a change in ionic strength of buffer. This confirms a specificity of SAP for ribose-5-phosphate and AMP which interferes with SAP-PE binding in order for it to be eluted from a PE-agarose column. It infers that ribose-5-phosphate and AMP are interfering with the PE binding site on SAP either in an indirect way by binding near the site, or by directly competing with PE. SAP affinity for ribose-5-phosphate and AMP appears similar to that for PE, as 30mM concentration is required to elute similar amounts, i.e. 30mM of all ligands is needed to elute almost all SAP.

4.4.4 Crystallisation trials using *Limulus* SAP

Almost all of the crystallisation trials were successful in producing crystal growth as expected from previous trials, although none tried were successful in producing crystals that gave high quality diffraction data. This may be in part to conditions within the trays, but may also be due to a combination of SAP isoforms within the SAP mixture as indicated by SDS-PAGE, glycosylation, and FPLC analysis of surface hydrophobicity. These may affect SAP-SAP interactions during crystal formation thereby creating a non-uniform

crystal lattice which produces poor diffraction. One possibility to attempt to create more uniform crystals would be to isolate the molecular aggregate isoform groups identified by FPLC and use them for crystallisation trials. Deglycosylation could also be performed to reduce the effect of oligosaccharides on crystal packing and lattice formation.

There is also a possibility that SAP solution used in the crystal trials has some contamination from other *Limulus* plasma proteins. High molecular weight bands are present on SDS-PAGE at approximately 66kDa, 48kDa, and 35kDa which may represent non-SAP proteins or SAP dimers. CRP may be a trace contaminant, as well as α_2-macroglobulin (~200kDa) and haemocyanin (~66kDa) or other components of the haemolymph. Contamination of pure SAP with non-SAP proteins would affect crystal growth. The trace contaminants could also be present on the FPLC analysis although which peaks they represent is unknown. Further experiments are being undertaken to determine this.

The best conditions for crystal growth appear to be 10mM calcium chloride, 6% PEG 6000 and 50mM MES pH 7.5. However, crystal growth is observed at higher and lower concentrations of calcium chloride and PEG, and at higher and lower pH values of 50mM MES. A more detailed look at the effect of calcium concentration and pH on crystal growth could be achieved by conditions that have much higher and lower concentrations of calcium, and also varying the pH of 50mM MES. Increasing the number of trials performed using the higher and lower concentrations of calcium chloride could increase the number of crystals seen at these concentrations. This may also be true for the conditions tried only a few times for varying the pH of 50mM MES and the concentration of PEG 6000. Therefore, it is difficult to come to a conclusion about the conditions best for crystal growth using these trials, and that more trials would need to be performed centred on those areas which show more crystal growth than others: 10mM calcium chloride, 6% PEG 6000, 6.5% PEG 6000, and 50mM MES pH 7.0 and 7.5.

4.4.5 Overall conclusions

Limulus SAP contains at least three subunit forms of different molecular weight as identified by SDS-PAGE. Two of the subunit forms appear to be glycosylated as deglycosylation reduces their molecular weight as determined by SDS-PAGE, although the third subunit form could be glycosylated but this is not apparent on the gel. At least four molecular aggregate forms of SAP vary in surface hydrophobicity of which one molecular aggregate form appears to be due to glycosylation.

Human SAP shows binding to 1' and 6' phosphorylated sugars, and 3' sulphated sugars (Loveless *et al.*, 1992) and also dAMP (Hohenester *et al.*, 1997), so if *Limulus* SAP shows similar binding properties to human SAP it would also be expected to show binding to ribose-5-phosphate and AMP which have a 5' phosphorylated group, although the position of the phosphate group is different. Ligands ribose-5-phosphate and AMP elute SAP from a PE column suggesting binding at the same binding site or near to it interfering with SAP-PE interaction. The binding affinity of SAP for ribose-5-phosphate and AMP shows that some SAP binds at 10mM ligand concentration, which may be represented by the first peak on the surface hydrophobicity analysis traces with smaller peaks representing other SAP molecular aggregate forms that did not bind ribose-5-phosphate at 10mM ligand concentration as avidly. However, extended elutions at 10mM ligand concentrations were not performed so the possibility of equal or similar binding affinity of all SAP molecular aggregate forms for ribose-5-phosphate or AMP cannot be ruled out. There is also the possibility of trace contaminants being present which may cause differences between elution traces for surface hydrophobicity analysis.

The proportion of SAP eluted using 30mM AMP is slightly higher than the proportion of SAP eluted using 30mM ribose-5-phosphate suggesting a slightly higher affinity of SAP for AMP. However, overall the proportion of SAP eluted using ribose-5-phosphate and AMP at either 10mM or 30mM is not too dissimilar probably due to a similarity in structure and hence binding between ribose-5-phosphate and AMP.

Chapter 5. Isolation of the genes for pentraxins CRP and SAP from *Mustelus canis*.

5.1 Introduction

The smooth dogfish, *Mustelus canis*, is a vertebrate of the class Chondrichthyes and the subclass Elasmobranchii. It is found on the Northwestern Atlantic coast ranging from Massachusetts to Florida, and Southern Brazil to Northern Argentina. *M. canis* are viviparous and can have litters of between 3-18 pups (Conrath and Musick 2002). Maturity is reached at 2-3 years for males and 3-4 years for females (TeWinkel 1950; Conrath and Musick 2002; Conrath *et al.*, 2002). Male *M. canis* can live 10 years, whilst female *M. canis* can live for up to 16 years (Conrath *et al.*, 2002). *M. canis* have flat, blunt teeth that are used to grind their prey which are mainly Crustacea such as *Cancer irroratus* (Atlantic Rock Crab), *Limulus polyphemus* (Atlantic horseshoe crab), *Ovalipes ocellatus* (Lady crab) and *Callinectes sapidus* (Blue crab), but also include Mollusca such as razor clams, Osteichthyes (Bony fish), and Polchaete (Gelsleichter *et al.*, 1999).

M. canis have both innate immune type proteins such as CRP and SAP (Robey *et al.*, 1983), and also adaptive immune type proteins such as immunoglobulins (Marchalonis and Edelman 1965, 1966). Both CRP and SAP like proteins were found in *M. canis* using their calcium dependant ligand binding affinities for phosphocholine and sepharose respectively (Robey *et al.*, 1983), which is similar to the binding affinities of *L. polyphemus* CRP (Robey and Liu 1983) and SAP (Tharia *et al.*, 2002). The concentration of dogfish CRP was found to be around 400μg/ml of whole serum, and SAP was 100μg/ml. These concentrations are significantly higher than those found in human serum (CRP (<50μg/l) and SAP (30-50mg/l)), but the SAP concentration is similar to the high level of SAP (120μg-170μg/ml) like protein in the serum of *L. polyphemus* (Shrive *et al.*, 1999). This indicates that although *M. canis* are more evolved than *L. polyphemus* and also have

immunoglobulins, they still require CRP and SAP like proteins in high concentrations to function as a major part of their immune system.

Approximate molecular weights of dogfish CRP and SAP protomers were found to be 26kDa and 27.5kDa respectively, similar to the molecular weights of the CRP and SAP protomers in humans (CRP: 23kDa, SAP: 25.5kDa) and *L. polyphemus* (CRP: ~25kDa, SAP: 23.8kDa) (Marchalonis and Edelman 1968; Robey *et al.*, 1983; Tennent *et al.*, 1993; Emsley *et al.*, 1994; Tharia *et al.*, 2002). Crystals of dogfish CRP have been grown and subjected to X-ray diffraction analysis which indicated that they were arranged in paired hexamers, perhaps partly held together by disulphide bonds (Robey *et al.*, 1983; Samudzi *et al.*, 1993). Both reports of Robey and Liu (1983) and Samudzi *et al.*, (1993) indicate that dogfish CRP has some heterogeneity although whether this is to the extent of that shown for CRP in horseshoe crabs is unknown. N-terminal analysis of doublet bands of dogfish CRP on SDS-PAGE gave similar but not identical amino acids sequences (Samudzi *et al.*, 1993), although this was not apparent in the study by Robey *et al.*, (1983) who sequenced the N-terminal amino acids of dogfish CRP and SAP shown in table 5.1.

Table 5.1. N-terminal sequences of *M. canis* CRP and SAP found by Robey *et al.*, (1983).

Name of sequence	Sequence
M. canis CRP N-terminus from Robey *et al.*, (1983)	Ser Pro Val Ala Ala Ser Tyr Arg Ala Thr Ala Gly Leu Ala Gly Lys Ala Leu Asp Phe
M. canis SAP N-terminus from Robey *et al.*, (1983)	Gly Phe Pro Gly Lys Ser Lys Iso Phe

As *M. canis* belongs to the elasmobranchs which are one of the earliest evolved vertebrate classes and possess both CRP and SAP like protein and immunoglobulins, they provide a good model to study the evolution of the innate and adaptive immune system from the earlier evolved invertebrates like *L. polyphemus* that possess only innate immune like proteins to the later evolved vertebrates such as humans that possess both innate and adaptive immune proteins. The crystal structure of dogfish CRP and SAP has been

determined (A. K. Shrive, T. Greenhough unpublished results) but the full amino acid sequences are not known. Determining the amino acid sequence of dogfish CRP and SAP would help to elucidate areas of sequence homology with other CRP and SAP which may in turn indicate areas of similar structure and possibly function. This would enable us to look at evolution of the pentraxins from *L. polyphemus* to *M. canis* to humans.

It was decided to attempt to determine the gene sequences of dogfish CRP and SAP by using previously successful methods to find the sequence of *Limulus* SAP (Tharia *et al.*, 2002). The amino acid sequence of peptides derived from cyanogen bromide and trypsin cleavage of dogfish CRP in addition to the identification of some amino acids from the electron density maps of dogfish CRP and SAP, and also peptides from dogfish SAP (Robey *et al.*, 1983) were used to design degenerate oligonucleotide primers. The primers were used to amplify regions of genomic dogfish DNA encoding genes for dogfish CRP and SAP. However, the sequences retrieved using PCR did not show sequence identity to pentraxin-like proteins and so amplification of dogfish DNA encoding CRP and SAP was unsuccessful. If it had worked the sequenced DNA would have been used to design more specific primers to amplify cDNA encoding dogfish CRP and SAP from reverse transcription of mRNA. The amplified cDNA product would be used in Rapid Amplification of 5' cDNA ends (RACE) to produce double-stranded cDNA. The double stranded cDNA would then be amplified using two sets of primers until there was enough cDNA to sequence (Sambrook and Russell 2001).

5.2 Materials and Methods

5.2.1 Collection of *M. canis* liver and blood

Blood and liver samples were obtained from a single *M. canis* from the Marine Resources Centre at the Marine Biological Laboratory, Woods Hole, Massachusetts. During bleeding, the dogfish was restrained and blood taken from the posterior end

between the cloacal opening and the tail. A 20ml syringe with a 22 gauge needle was inserted exactly on the midline of the dogfish body to bleed from the caudal blood vessels. 20-50mls of blood was taken and added to heparinised elasmobranch ringers solution (0.28M NaCl, 0.54M Urea, 0.01M KCl, 0.0045M NaHCO$_3$, 0.007% NaH$_2$PO$_4$.H$_2$O, 0.01M CaCl$_2$.2H$_2$O, 100 units/ml heparin, and 1g/litre dextrose) in a 1:1 volume to volume ratio. The elasmobranch ringers solution was used to dilute the blood without the blood cells lysing and releasing their contents, and the heparin reduced the clotting. The blood and ringers solution was centrifuged to pellet the blood cells, and the supernatant plasma removed. Sodium azide at 0.02% was added to the plasma to prevent fungal or bacterial growth, and then stored at 4°C.

After bleeding, the dogfish was sacrificed and sections of its liver lobes were removed and cut up into slices. These slices were immediately put onto dry ice in order to freeze them, and subsequently kept at -80°C.

Collection of *M. canis* blood and liver was performed in collaboration with Prof P. B. Armstrong.

5.2.2 Isolation of pentraxins from *M. canis* plasma

CRP and SAP were isolated from *M. canis* plasma using sequential elutions of phosphocholine (PC) and phosphoethanolamine (PE) during affinity chromatography similar to the *Limulus* pentraxins (see section 2.1). Filtered plasma was first diluted with calcium wash buffer (50mM Tris, 10mM CaCl$_2$, 150mM NaCl) to reduce the salt concentration. It was then applied to a 25ml PE-agarose affinity column equilibrated with calcium wash buffer (see section 2.1.3). CRP was eluted with 20mM PC in calcium wash buffer, and SAP was subsequently eluted with 50mM PE in calcium wash buffer. The flow rates were 0.5ml/min in accordance with maximum flow rate through the column, and all

buffers were at pH 7.4 and had been filtered and de-gassed. The column was cleaned of remaining proteins by regeneration with EDTA and high salt buffers (see section 2.1.3). CRP and SAP were dialysed with EDTA for 24 hours, followed by calcium wash buffer for 24 hours to remove the ligand (i.e. PC or PE) and replace lost calcium ions. CRP and SAP were then concentrated down to around 10mg/ml and stored at 4°C (see section 2.1.5 and 2.1.6)

Isolation of dogfish pentraxins was performed by I. Burns and A. Roberts.

5.2.3 Cyanogen bromide cleavage of *M.canis* CRP

In order to obtain peptides of dogfish CRP to use for primer design, CRP was fragmented using cyanogen bromide which cleaves at the carboxy side of methionine residues in the protein sequence. As methionine residues are relatively rare in a protein sequence compared to the other amino acids, it is a good place to cleave to get peptides of a sufficient size to sequence.

Before the cyanogen bromide digest, the protein was subjected to reduction and sulphydryl modification to prevent cross-linking of cysteine containing peptides by blocking free sulphydryl groups. After being kept at -80°C for 20 minutes, around 2mg of dogfish CRP was freeze dried using a speed vacuum for approximately two hours. This was done to remove buffer and other components from the protein as the volume was too small to dialyse. The solid was re-dissolved in 6M guanidine hydrochloride in 0.6M Tris pH 8.6 and 15µl of 0.1M β-mercaptoethanol. Nitrogen gas was bubbled through for 3 minutes and then the solution incubated at room temperature in the dark for 3 hours. 150µl of iodoacetic acid (268mg/ml) in 0.1M NaOH was then added, and nitrogen bubbled through for 3 minutes and incubated for 3 hours in the dark again. The solution was then dialysed against 5mM NH_4HCO_3 for 24 hours at 4°C in the dark, and stored at -20°C.

For cyanogen bromide digestion, the sulphydryl modified dogfish CRP was re-dissolved in 250μl of 70% formic acid. One crystal of cyanogen bromide was added and nitrogen bubbled through the solution for three minutes before incubating for 24 hours at room temperature. Cleavage was stopped with 10 volumes of deionised water, and the cleaved peptides stored at -20°C.

Cyanogen bromide cleavage solution was freeze-dried using a speed-vacuum and re-dissolved in 400μl of 1M Tris pH 8.0 and the peptides analysed using SDS-PAGE and Western Blotting. The gel used was 17.5% resolving gel and 4% stacking gel (see section 2.1.7) run in a dual buffer system (upper reservoir 0.1M Tris, 0.1M Tricene, 0.1% SDS, 0.1mM sodium thioglycolate to protect against N-terminal blocking; lower reservoir 0.2M Tris pH 8.9). The gel was blotted onto polyvinylidene difluoride (PVDF) membrane using buffer containing 35% methanol in 10mM N-Cyclohexyl-3-aminopropanesulfonic acid (CAPS) buffer at pH 11.0 to enhance binding onto membrane, and 3.2mM DL-Dithiothreitol (reducing agent). The PVDF membrane with peptides bound to it was stained with Coomassie Brilliant Blue R-250 stain (see chapter 2.1), rinsed with water to remove buffer components, and then air-dried in order for selected bands of size 12, 9, and 4kDa to be cut out and sent for amino acid sequencing to Alta Bioscience, University of Birmingham.

The cyanogen bromide digest and subsequent SDS-PAGE and Western blotting were performed by I. Burns and A. Roberts.

5.2.4 Trypsin cleavage of *M. canis* CRP

In a second attempt to obtain a peptide sequence to design primers for isolation of the dogfish CRP gene, a trypsin digest was used. Trypsin hydrolyses a protein at the carboxyl side of arginine, lysine, and 5'-aminoethylcysteine residues (Aitkens *et al.*, 1989).

Sometimes anomalous cleavage may occur during prolonged digestion, which is mainly attributed to the presence of chymotrypsin which cleaves on the carboxy side of hydrophobic residues especially phenylalanine, tryptophan, tyrosine and leucine (Aitkens *et al.*, 1989).

Trypsin (proteomics grade) was obtained from Sigma as a lyophilized powder which had the chymotryptic activity removed. Trypsin was reconstituted in 1mM HCl to give a concentration of 1mg/ml at pH 3.0. The dogfish CRP was diluted from 8mg/ml to 5mg/ml with 100mM Tris pH 9.0, which gave an overall pH of 8.5 that was required for the trypsin digest. Trials of the digestion were performed using trypsin (at 1mg/ml) added to the dogfish CRP at varying proportions from 1:100 ratio to 1:20 ratio, and digested for periods of time between 30 minutes to 6 hours. The products of the digests were run on gels as mentioned previously for the cyanogen bromide digest and suitable bands of the PVDF blot at sizes of approximately 19, 18 and 4kDa were cut out and sent for amino acid sequencing at Alta Bioscience, University of Birmingham.

The trypsin digest was performed by I. Burns.

5.2.5 Purification of DNA from *M. canis* liver

Genomic DNA from the dogfish liver was isolated using the Wizard SV Genomic DNA Purification System (Promega). The purification was performed as follows: 20mg of dogfish liver was incubated with 'Digestion Solution Mastermix' (200µl of Nuclei Lysis Solution (Promega), 50µl 0.5M EDTA pH 8.0, 20µl 20mg/ml Proteinase K, and 5µl 4mg/ml RNase solution (Promega)) at 55°C for approximately 18 hours. The solution was centrifuged to pellet undigested liver, and the supernatant transferred to a new 1.5ml micro centrifuge tube. 250µl of Wizard SV Lysis buffer was added to the solution and mixed. The lysate was added to an assembled minicolumn (Promega) and centrifuged at 13000

rpm for 3 minutes in a microcentrifuge to bind the DNA to the column. After discarding the flow-through, 650μl of Wizard SV wash solution was added to the column and centrifuged at 13000 rpm for 1 minute. Column flow through was discarded. This was repeated another 3 times discarding the flow through each time. After the final wash, the minicolumn assembly was centrifuged for 2 minutes at 13000 rpm to dry the binding matrix and then the column transferred to a sterile 1.5ml micro centrifuge tube. 250μl of nuclease-free water preheated to 65°C, was added to the column. After 2 minutes incubation at room temperature, the minicolumn assembly was centrifuged at 13000 rpm for 1 minute to elute the DNA from the column. Another 250μl of nuclease free water was added and incubated for 2 minutes. The column assembly was then centrifuged again for 2 minutes at 13000 rpm to elute any remaining DNA from the column, which was then stored at -20°C. An aliquot of DNA was diluted in TE buffer (10mM Tris pH 7.5, 1mM EDTA) and the DNA concentration determined for double stranded DNA at 260nm and 280nm.

5.2.6 Restriction digests

Restriction digests were performed on dogfish DNA with *Bam*H1 and *Sal*1 to check that DNA was not damaged. Approximately 1μg of DNA was used for each digest, which was a total of 30μl with DNA, enzyme (10 units of *Bam*H1 or *Sal*1), enzyme buffer (1x), and nuclease free water. Controls for the digests were set up with plasmids for *Bam*H1 and *Sal*1 digests, and also digests with no enzymes to check that the non-enzyme components were restriction enzyme free. Each digest was left at 37°C for 2 hours.

Restriction digests were also performed on plasmid DNA to check whether PCR product had been incorporated into the plasmid. Approximately 1μg of DNA was used for each digest, which was a total of 30μl with DNA, 10 units of *Eco*R1, enzyme buffer (1x), and nuclease free water. Controls were set up with no enzyme components. The digest was left at 37°C for 2 hours.

5.2.7 Primers for *M. canis* CRP and SAP

Primers were ordered from MWG. Primers were sent lyophylised, they were first dissolved in nuclease free water to give a concentration of 100pmol/µl. They were then diluted with nuclease free water in small aliquots of 200µl to obtain a primer concentration of 20pmol/µl.

5.2.8 PCR of *M. canis* DNA

PCR was set up using the primers designed for the dogfish CRP and SAP. Each PCR mixture was composed of Mastermix (Promega) (50 units/ml of *Taq* DNA polymerase, 400µM dATP, 400µM dGTP, 400µM dCTP, 400µM dTTP, 3mM MgCl$_2$); 50pmoles of each primer; and 200ng of DNA, made up to 25µl with nuclease free water. There were several different PCR programs used on both the general PCR machine and the gradient PCR machine. The programs used were based upon a general PCR program but with different annealing temperatures. The general program was:

- 95°C for 5 minutes - denaturation
- 35 cycles of 95°C for 1 minute, X°C for 30 seconds, 72°C for 1 minute - annealing
- 72°C for 10 minutes - extension
- Hold at 4°C.

Variations of the general program were used on the general PCR machine which involved annealing temperatures (X) of 45°C, 46°C, 50°C, 55°C, and 60°C. Variations of the general program were performed using the gradient PCR machine with gradients of 47°C-57°C, 45°C-55°C, 44°C-54°C, 50°C-65°C, and 55°C-65°C annealing temperature. PCR products were kept at -20°C until use for gel electrophoresis.

5.2.9 Gel electrophoresis of PCR product

To identify the sizes of DNA fragments from PCR and those from restriction digests and amplified using PCR, gel electrophoresis was used. The gels used for separating out DNA fragments were made of agarose: 0.8% for restriction digests and 1-1.5% for PCR products. Agarose gels were made using an agarose powder dissolved in Tris Acetic Acid (TAE) buffer (40mM Tris Acetate, 2mM EDTA), which was left to set in a gel casting unit. DNA samples were mixed with loading dye that contained sucrose to help the samples sink to the bottom of the well and coloured dye to visualise migration through the gel. Samples and size ladders (see figure 5.1) were loaded into the wells and gels run at 80-100V until the DNA had migrated a suitable distance as judged by dye in the loading buffer. Gels were stained using ethidium bromide at 0.5µg/ml for approximately 10 minutes, then rinsed with large volume water washes for approximately 5 minutes. The DNA bands were visualised on a UV transilluminator.
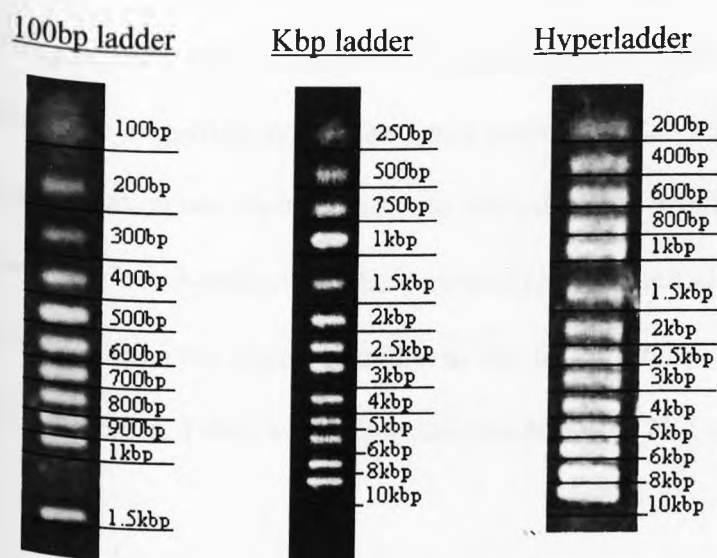
Figure 5.1. Diagram of the three ladders used 100bp, kbp, and hyperladder, and sizes of their bands on a gel.

5.2.10 Topo cloning

To Topo clone, firstly the PCR product was inserted into the Topo vector plasmid pCR2.1-Topo (Invitrogen) (see figure 5.2). This was done by using 4µl of PCR product

added to 1µl of salt solution (1.2M NaCl, 0.06M $MgCl_2$) and 1µl of Topo vector which was then incubated at room temperature for 30 minutes. PCR product is inserted into the vector inside the gene for β-Galactosidase enzyme which hydrolyses X-gal from colourless to blue. Therefore if the gene is disrupted by the insertion of PCR product, the resulting protein from the gene will be non-functional and therefore not convert X-gal from colourless to blue and colonies remain colourless or "white".

2µl was taken and added to either 100µl of *E. coli* cells strain:α-select, genotype: F-*deoR endA*1 *recA*1 *relA*1 *gyrA*96 *hsd*R17(r_k ⁻, m_k ⁺) *supE*44 *thi*-1 *phoA* Δ(*lacZYA* *argF*)U169 Φ80*lacZ*ΔM15 λ⁻ (Bioline); or "1 shot" of *E. coli* strain: TOP10, genotype: *F-mcrA ?(mrr-hsdRMS-mcrBC) ?80lacZ?M15 ?lacX74 recA1 araD139 ?(araleu)* 7697 *galU galK rpsL (StrR) endA1 nupG* (Invitrogen), then mixed and incubated on ice for 30 minutes. Cells were then heat-shocked for around 45 seconds at 42°C, and put on ice again for approximately two minutes. 900µl of SOC (2% Tryptone, 0.5% Yeast Extract, 0.4% Glucose, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$ and 10mM $MgSO_4$) medium was then added, and cells incubated for an hour at 37°C to allow one replication so they could express the ampicillin resistance genes present in the Topo vector. 100µl of each Topo cloning reaction was then spread onto individual LB Broth agar plates coated in X-Gal (5-bromo-4-chloro-3-indolyl-β-galactopyranoside), and left at 37°C overnight. White colonies were taken from the plate and added to sterile 15ml tubes that contained 5ml of LB Broth with Ampicillin. Tubes were incubated overnight at 37°C, and the agar plates kept at 4°C.

M13 Reverse Primer
lacZα ATG
Hind III   Kpn I   Sac I BamH I   Spe I

```
CAG GAA ACA GCT ATG AC  ATG ATT ACG CCA AGC TTG GTA CCG AGC TCG GAT CCA CTA
GTC CTT TGT CGA TAC TG  TAC TAA TGC GGT TCG AAC CAT GGC TCG AGC CTA GGT GAT
```

BstX I EcoR I   EcoR I

```
GTA ACG GCC GCC AGT GTG CTG GAA TTC GCC CTT  PCR Product  AAG GGC GAA TTC TGC
CAT TGC CGG CGG TCA CAC GAC CTT AAG CGG GAA              TTC CCG CTT AAG ACG
```

EcoR V   BstX I   Not I   Xho I   Nsi I Xba I   Apa I

```
AGA TAT CCA TCA CAC TGG CGG CCG CTC GAG CAT GCA TCT AGA GGG CCC AAT TCG CCC TAT
TCT ATA GGT AGT GTG ACC GCC GGC GAG CTC GTA CGT AGA TCT CCC GGG TTA AGC GGG ATA
```

T7 Promoter   M13 Forward (-20) Primer

```
AGT GAG TCG TAT TAC AAT TCA CTG GCC GTC GTT TTA CAA CGT CGT GAC TGG GAA AAC
TCA CTC AGC ATA ATG TTA AGT GAC CGG CAG CAA AAT GTT GCA GCA CTG ACC CTT TTG
```

+1 Plac lacZα

f1 ori

pUC ori

pCR®2.1-TOPO®
3.9 kb

Kanamycin

Ampicillin

Figure 5.2. Diagram of Topo vector pCR 2.1 adapted from www.invitrogen.com

5.2.11 Isolation of DNA from *E. coli*

To isolate the DNA from the *E. coli* colonies grown in the Ampicillin containing LB Broth, a Wizard Plus SV Minipreps DNA Purification System (Promega) and Perfectprep Plasmid Minikit (Eppendorf) were used.

For the Wizard Plus SV Minipreps DNA Purification System (Promega) the following protocol was followed. 1ml of the cell culture was transferred to a micro centrifuge tube and the cells pelleted using centrifugation. The supernatant was removed, and 250µl of cell resuspension solution (Promega) was added and the cells resuspended. 250µl of cell lysis solution (Promega) was then added and mixed with the resuspended cells by inversion of the tubes. Tubes were incubated for 5 minutes at room temperature until the solution had

cleared. Then, 10µl of Alkaline Protease (Promega) was added and left for 5 minutes. 350µl of neutralization solution (Promega) was then added and mixed by inversion of the tubes. The solution was then centrifuged for 10 minutes at 13000 rpm and the supernatant which contained the DNA, was added to a minicolumn assembly (Promega). The minicolumn assembly was centrifuged at 13000 rpm for 1 minute to allow the DNA to bind, and the flow through discarded. The minicolumn was then washed by adding 750µl of column wash solution (Promega) and centrifugation for 1 minute at 13000 rpm. The flow through was discarded and the minicolumn washed again with 250µl of column wash solution by centrifugation at 13000 rpm for 2 minutes. The flow through was discarded again and the column then transferred to a sterile 1.5ml mini centrifuge tube. To remove DNA from the column, 100µl of nuclease free water was added to the minicolumn and spun at 13000 rpm for 1 minute. The DNA in the water was kept at -20°C. Insertion of a PCR fragment was confirmed by restriction enzyme digest of the plasmid.

For the Perfectprep Plasmid Minikit (Eppendorf) DNA isolation, the following protocol was used. 1.5ml of cell culture was transferred to a mini centrifuge tube and pelleted by centrifugation at 13000rpm for 20 seconds. The supernatant was removed and the bacteria were resuspended using 100µl of solution 1 (buffered RNase A solution - Eppendorf) and vortexing. The bacteria were then lysed by adding 100µl of solution 2 (alkaline lysis solution – Eppendorf) to the resuspension and mixing by inversion. Bacterial lysate was neutralized by adding 100µl of solution 3 (neutralisation solution – Eppendorf) and immediately mixing by inversion. Centrifugation of this neutralized lysate at 13000rpm for 30 seconds pelleted the bacteria cell walls and proteins, then the supernatant could be removed and transferred to a spin column in a collection tube (Eppendorf). 450µl of DNA binding matrix (Eppendorf) was added to the supernatant in the spin column to bind plasmid DNA. Plasmid DNA was then purified by centrifuging the spin column/collection tube assembly for 30 seconds at 13000rpm. The filtrate was discarded and the spin column placed back into the collection tube, where plasmid DNA was further purified by adding

400µl of diluted purification solution (buffered salt solution – Eppendorf in a 1:1 ratio with 95% ethanol) and centrifuging for 60 seconds at 13000rpm. Residual diluted purification solution was removed by centrifuging the spin column/collection tube assembly again at 13000rpm for 60 seconds. The spin column was transferred to a fresh collection tube, and 50µl of water at 65°C was added to the DNA binding matrix in the spin column. The spin column/collection tube assembly was then centrifuged at 13000rpm for 60 seconds, and the filtrate that contained the plasmid DNA was stored at -20°C. Insertion of a PCR fragment into the plasmid DNA was confirmed by restriction digest.

5.2.12 Preparation of DNA for sequencing

Plasmid DNA was precipitated from the Wizard Plus SV Minipreps DNA purification System sample by adding 3M sodium acetate at a tenth of the volume of DNA solution, and then ethanol at twice the volume of the DNA solution. This was kept overnight at -20°C and then spun at 13000 rpm for 15 minutes to pellet the DNA. The supernatant was decanted, 100µl of 75% ethanol was added, and the solution spun again at 13000 rpm for 10 minutes. The supernatant was decanted again and the remaining ethanol left to evaporate, leaving the pellet of DNA which was then sent for sequencing to MWG using their comfort read service.

Plasmid DNA from the Perfectprep Plasmid Minikit DNA isolation sample was also sent for sequencing, but was sent just in water as the DNA concentration was high enough for sequencing at MWG to be performed.

The sequencing primers M13 reverse and T7 which flank the PCR product insertion site were used during sequencing at MWG.

5.2.13 Sequence alignment

Cyanogen bromide and trypsin peptides of dogfish CRP, and the sequences of dogfish CRP and SAP (Robey *et al.,* 1983) were aligned with other pentraxin sequences: human CRP and SAP, carp pentraxin, salmon pentraxin, trout pentraxin, rat CRP and SAP, mouse SAP, hamster SAP, guinea pig CRP and SAP, rabbit CRP, pig CRP and SAP, cattle SAP, *Limulus* CRP 1.1, 1.4, and 3.3, *Limulus* SAP, and *Tachypleus* CRP 1, 8 and 16. Alignment was performed using Clustal.

5.2.14 Sequence search

Basic Local Alignment Search Tool (BLAST) was performed using EBI BLAST (www.ebi.ac.uk) and NCBI BLAST (www.ncbi.nlm.nih.gov) for the nucleotide, protein and translated nucleotide databases. In EBI, the sequence (either protein or nucleotide) was loaded into the query box and the parameters set for nucleotide or protein. The sensitivity was tried at all levels and the resulting aligned sequences were sorted according to p-value. In NCBI, the sequence was loaded into the query box, the database chosen for the relevant sequence (e.g. non-redundant protein sequences for a protein sequence) and the program selected for the relevant sequence (e.g. blastp for protein alignment with proteins). When analysing the sequence obtained from PCR, the nucleotide to nucleotide, protein to translated nucleotide, and protein to protein BLAST search programs were used in NCBI.

5.3 Results

5.3.1 *M. canis* and *S. acanthias* primers

N-terminal sequencing of dogfish CRP peptides from cyanogen bromide cleavage (see section 5.2.3) gave two fragments:

Fragment 1: Proline, Alanine, Alanine, Serine, Serine

Fragment 2: Methionine, Valine, Glycine, Glutamic acid, Aspartic acid, Glutamine, Asparagine, Serine, Leucine, Glycine, Glycine, Serine, Phenylalanine, Glutamic acid, Histidine, Serine.

These fragments and known fragments of dogfish CRP and SAP (Robey *et al.*, 1983) were aligned with known sequences of mammalian, fish and invertebrate pentraxins (see figure 5.3, and section 5.2.13). The first fragment (fragment 1) of cyanogen bromide cleavage peptide aligned against the previously found dogfish CRP N-terminus fragment (Robey *et al.*, 1983) which was near the N-terminus of human CRP and so was called N-terminus sequence, whilst the second fragment (fragment 2) from the cyanogen bromide cleavage aligned with the conserved area of ligand and calcium binding site in all pentraxins. Previous work by A. K. Shrive determined the N-terminus of dogfish CRP using the same methods as described for this research (unpublished results), and this sequence is shown on figure 5.3 also. Using the cyanogen bromide cleaved peptide sequence and the conservation of residues within the areas to which they were aligned, primers were designed. Nucleotide sequences for pentraxin like proteins from *Squalus acanthias* (spiny dogfish) (Parton *et al.*, 2007) were found in the EBI database and primers were designed from these also. The *S. acanthias* primers were designed from the pentraxin sequence near the N-terminus and around 150 base pairs into the sequence similar to those designed from the *M. canis* cyanogen bromide peptide fragments. *S. acanthias* and *M. canis* are not too dissimilar in features and lifecycle, and are both elasmobranchs, so it was thought that the primers designed from the *S. acanthias* pentraxin sequence might share enough homology to anneal to the *M. canis* pentraxin (i.e. CRP or SAP) DNA also.

```
Limulus_CRP1.1          -----------LEEGEITSKVKFPPSSSPSFPRLVMVGTLP  VLGQDQDSVGGKFDATQSLEGELSELNLWNTVLNHEQIKYLSKC
Limulus_CRP3.3          -----------LEEGEITSKVKFPPSSSPSFPRLVMVGTLP  VLGQDQDSVGGKFDATQSLEGELSELNLWNTVLNHEQIKYLSKC
Limulus_CRP1.4          -----------LEEGEITSKVKFPPSSSPSFPRLVMVGTLP  VLGQEQDSVGGEYDAEQSLEGELSELNLWNTVLNHEQIKHLSKC
Tachypleus_1_tCRP2      -------------------KVKFPPSSSPSFPRLVMVGTLP  VLGQDQDTVGGGFDAKQSLEGELSELNLWDTVLNHEQIKHLSKC
Tachypleus_8_tCRP1      -------------------KVKFPPSSSPSFPRLVMVGTLP  VIGQDQDSVGGGFDEKQSLVGELSELNLWDMVLNHEQIKHLSKC
Tachypleus_16_tCRP3     -------------------KVKFPPSTASSFPRLVMDGTLP  VLGQDQDTVGGGFVASESMEGELSELNMWNSVLNSNQILHLSNC
Limulus_SAP            ------------AVDIRDVKISFPGTQNPKFPHLRFMQTLP  VIGQEQDKIGGGFEEQESWSGELSDLQVWDEALTTHQVSTVASC
guinea_pig_crp        -------------GTDMSKKTFVFPKETDNSYVSLKAQLKKP  ILGQDQDSFGGSFDANQSFVGDIGDVNMWDFVLSPKEIDMVYSG
human_crp             -------------QTDMSRKAFVFPKESDTSYVSLKAPLTKP  ILGQEQDSFGGNFEGSQSLVGDIGNVNMWDFVLSPDEINTIYLG
rabbit_crp            --------------AGMHKKAFVFPKESDNSYVSLNAQLKKP  ILGQDQDSFGGSFEKQQSLVGDIGNVNMWDYALSPEEINTIYAG
rat_crp               -------------HEDMSKQAFVFPGVSATAYVSLEAESKKP  ILGQEQDSYGGGFDANQSLVGDIGDVNMWDFVLSPEQINAVYVG
golden_hamster_crp    -------------QKDMSKTAFVFPKESANSYVSLEAQSKKT  ILGQEQDSYGGGFDANQSLVGDIGDVNMWDIVLSPEQINTVCVG
pig_crp               -------------QTDMIGKVFVFPKESENSYVSLTARLTKP  ILGQEQDAFAGGFEKNQCLVGDIGDVNMWDYVLSPEEINTVYAG
golden_hamster_sap    -------------QTDLTGKVFVFPRQSETDYVKLIPRLDKP  ILGQEQDNYGGGFDNYQSFVGEIGDLNMWDSVLTPEEIKSVYQG
armenian_hamster_sap  -------------QTDLTGKVFVFPRESESDYVKLIPRLEKP  ILGQEQDNYGGGFDKSQSFVGEMGDLNMWDSVLTPEEIKSVYEG
rat_sap               -------------QTDLNQKVFVFPRESETDYVKLIPWLEKP  VLGQEQDTYGGGFDKTQSFVGEIADLYMWDSVLTPENIHSVDRG
mouse_sap             -------------QTDLKRKVFVFPRESETDHVKLIPHLEKP  VLGQEQDNYGGGFQRSQSFVGEFSDLYMWDYVLTPQDILFVYRD
pig_sap               -------------HRDLTGKVFVFPRESATDHVKLITKLEKP  VLGQEQDSYGGGFDKTQSFVGEIGDLYMWGSVLSPNEIRLVYQG
cattle_sap            -------------QTDLRGKVFVFPRESSTDHVTLITKLEKP  VLGQEQDSYGGGFDKNQSFMGEIGDLYMWDSVLSPEEILLVYQG
human_sap             -------------HTDLSGKVFVFPRESVTDHVNLITPLEKP  VLGQEQDSYGGKFDRSQSFVGEIGDLYMWDSVLPPENILSAYQG
guinea_pig_sap        -------------QTDLDKKVFVFPRESSSDHVNLITKLETP  ILGQEQDSYGGKFDRGQSFLGEIGDVYMWDSVLSPDDVQAVYYG
salmon_ptxn           -----------EHQDLSGKVFVIPMATSTSHVKLHANVSEP  MLVQEQDSYGGGFDVSQSFVGEVTDVHFWDSVISPCEIQLYMQL
trout_ptxn            -----------DLQDLSGKVFVIPMTTSTSHVKLHANVSKP  MLVQEQDSYGGGFDSSQSFVGEVTDVHFWDSVISPCEIQMYMEL
carp_pentraxin        LLS---LTAAATEVGLVGKVLLFPTKTNTSFVALTPEKPLS  LLGQDPDNYLGAFEVEQSFVGEITDVHMWDHVLSGSQIMAVYSN
dogfish_sap_Robey et al  --------------GFPGKSLIF-----------------  -------------------------------------------
dogfish_crp_Robey et al  -SPV--AASYRATAGLAGKALDF-----------------  -------------------------------------------
dogfish_n-terminus_cbr   --P---AASS----------------------------   -------------------------------------------
dogfish_sequence_cbr     --------------------------------------   MVGEDQDSLGGSFEHS---------------------------
dogfish_n-terminus_prev  -SPP--AASSSATAGL----------------------   -------------------------------------------
human CRP numbering      10.........................47          152.......................................195
```

Figure 5.3. Alignment shows sequences 10-47 and 152-195 (human CRP numbering) where dogfish CRP and SAP fragments align to known pentraxins. The green background indicates identical residues, whilst the yellow background indicates similar residues. Fragment 1 from cyanogen bromide digest is highlighted in grey as dogfish_n-terminus_cbr, fragment 2 from cyanogen bromide digest is highlighted in grey as dogfish_sequence_cbr, and previous N-terminus sequence of dogfish CRP determined by A. K. Shrive (unpublished results) is highlighted in grey as dogfish_n_terminus_prev. Previously discovered dogfish CRP and SAP sequences by Robey *et al.*, (1983) are also highlighted in grey.

The sequences for designing primers were used in a nucleotide and protein BLAST search to identify similar sequences. *M. canis* and *S. acanthias* CRP N-terminus primer sequences did not match with any sequences on BLAST searches suggesting they could be used for a primer without much non-specific binding during PCR. However, primers corresponding to fragment 2 (C-terminus primer) from *M. canis* and *S. acanthias*, had sequence similarity with other proteins using BLAST search and therefore a chance of amplifying non-dogfish CRP encoding DNA during PCR. To reduce the possibility of non CRP-encoding dogfish DNA being amplified, the primers were designed to be less degenerate as possible. Back translation of the sequences was performed and areas of least degeneracy highlighted for primer design. The primers were also non-complimentary so that they did not anneal with each other. Final designed primers are shown in table 5.2. A proline was added to the end of the *M. canis* N-terminus sequence primer as it is conserved throughout the CRP and SAP sequences shown in figure 5.3. The third base coding for proline was not used for primer design due to degeneracy at this site.

Table 5.2. Primers used in preliminary PCR of dogfish DNA to amplify dogfish CRP gene. Primers are detailed using the code as shown.

Primer	Sequence
M. canis N-terminus sequence	Gly Lys Ala Leu Asp Phe Pro
M. canis. N-terminus primer	5'-GGNAARGCNYTNGAYTTYCC-3'
M. canis C-terminus sequence	Met Val Gly Glu Asp Gln Asp
M. canis. C-terminus primer	5'-TCYTGRTCYTCNCCNACCAT-3'
S. acanthias. N-terminus primer	5'-TTTCCAACCAAAACCACGGC-3'
S. acanthias. C-terminus primer	5'-ATCTGTAATCTCCCCAAC-3'
Primer base symbol	**Primer base**
A	Adenosine
T	Thymine
C	Cytosine
G	Guanine
N (as designated by MWG)	Adenosine/Thymine/Cytosine/Guanine
R (as designated by MWG)	Adenosine/Guanine
Y (as designated by MWG)	Cytosine/Thymine

5.3.2 PCR using *S. acanthias* and *M. canis* primers on DNA from liver stored at –80°C

DNA from two pieces of the dogfish liver that had been frozen at -80°C was extracted and a restriction digest with *Bam*H1 and *Sal*1 set up to check the integrity of the DNA. The restriction digest products were run on a gel which indicated that there was little DNA degradation.

PCR using this DNA was performed at 55°C annealing temperature using the two DNA samples from the dogfish liver with both the *M. canis* and *S. acanthias* primers. The PCR products shown on figure 5.4 are of size ~100bp for *M. canis* primer product, and ~200bp for *S. acanthias* primer product, with contamination in both blanks (lanes 3 and 6).
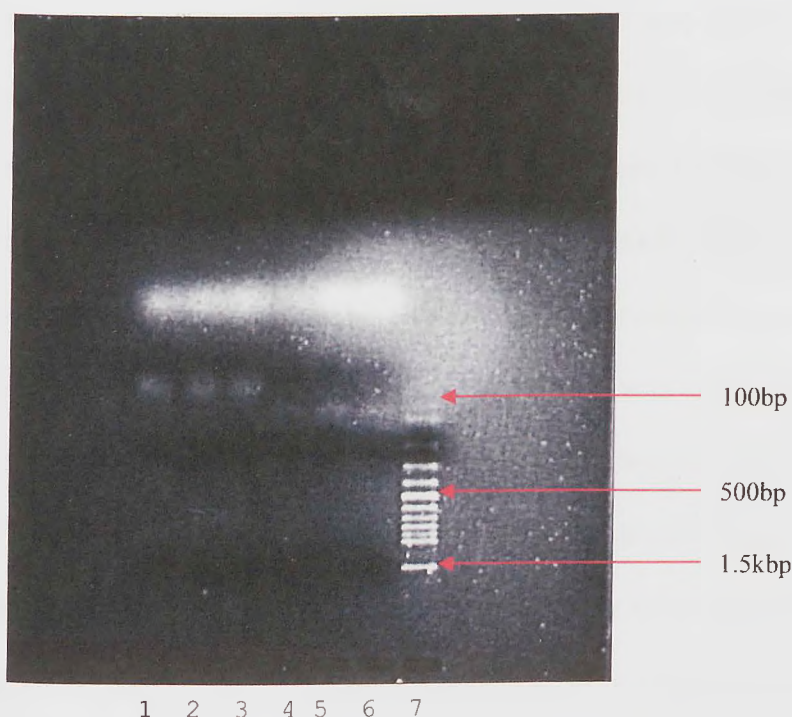
Figure 5.4. gel showing product of PCR at 50°C annealing temperature with *M. canis* and *S. acanthias* primers. Lane 1)DNA sample 1 *M. canis* 2)DNA sample 2 *M. canis* 3)blank *M. canis* 4)DNA sample 1 *S. acanthias* 5)DNA sample 2 *S. acanthias* 6)blank *S. acanthias* 7)100bp ladder

Amplified products of the PCR reaction with the dogfish primers were inserted into Topo vector pCR-2.1 which was then taken up by *E. coli* cells during Topo cloning (as described in section 5.2.10). White colonies were selected after transfection and growth. DNA isolated from the *E. coli* clones was subjected to a restriction digest with *Eco*R1 which cuts either side of the insert (see figure 5.2. and section 5.2.6) and so digestion with it should give the plasmid on its own linearised and the insert PCR product. Restriction

231

digest products were run on a gel which showed only single bands in each lane that did not correspond to the PCR fragment size seen previously (i.e. ~100bp). This meant that either there was no insert in the plasmid and the Topo cloning did not work, or that *Eco*R1 was not working. If *Eco*R1 did not work, there should be at least two bands (coiled and supercoiled plasmid) on the gel and not a single band, so it was likely that *Eco*R1 had worked fully and that the insert was missing so only the very small surrounding plasmid DNA between the *Eco*R1 sites was cut.

5.3.3 Dogfish CRP and SAP primers

Given the poor results from PCR using the previous primers, new primers were designed in an attempt to reduce the degeneracy for dogfish CRP and also for SAP. This time, several smaller pools of degenerate primers were designed, so multiple combinations of primers were tried. Primers for dogfish CRP were designed from the cyanogen bromide peptide fragments and Clustal alignment as shown previously. However, the N-terminus sequence was uncertain and had low homology with other pentraxins, so a region encoding the pentraxin motif that is present in all pentraxins was used. This sequence was approximately 100 residues from the N-terminus of human CRP (see figure 5.5). Sequences were then checked against the electron density map of dogfish CRP (unpublished work) to determine whether the amino acids coded for by the primer would fit into the electron density map. Those which did not were changed to an amino acid (and therefore codon in the primer) which was suitable for both the electron density map and the sequence alignment. Dogfish SAP primers were also designed from the existing dogfish SAP fragment sequence (Robey *et al.*, 1983) and the high area of sequence homology around the ligand and calcium binding sites in pentraxin sequences (see figure 5.5). The final designed primers and the amino acid sequence they coded for, were used in BLAST searches to identify any other potential gene sequences and hence proteins, that the primers may anneal to and amplify. There were no high similarity alignments against non-pentraxin proteins. The final designed primers are shown in table 5.3.

```
limulus_crp_1.1         ------------LEEGEITSKVKFPP    WHHVCHTWSSWEGEAT    LVVLGQDQDSVGGKFDATQSLEGELSELNLWNTVL
limulus_crp_1.1         ------------LEEGEITSKVKFPP    WHHVCHTWSSWEGEAT    LVVLGQDQDSVGGKFDATQSLEGELSELNLWNTVL
limulus_crp_3.3         ------------LEEGEITSKVKFPP    WHHVCHTWSSWEGEAT    LVVLGQDQDSVGGKFDATQSLEGELSELNLWNTVL
Tachypleus_1_tCRP2      --------------------KVKFPP    WHHACHTWSSWKGEAT    LVVLGQDQDTVGGGFDAKQSLEGELSELNLWDTVL
Tachypleus_8_tCRP1      --------------------KVKFPP    WHHVCHTWSSWEGEAT    LVVIGQDQDSVGGGFDEKQSLVGELSELNLWDMVL
Tachypleus_16_tCRP3     --------------------KVKFPP    WHHVCYLWSSWEGSSS    TVVLGQDQDTVGGGFVASESMEGELSELNMWNSVL
limulus_sap            ------------AVDIRDVKISFPG     WYHVCHVWSGVDGRMA    TVVIGQEQDKIGGGFEEQESWSGELSDLQVWDEAL
rabbit_crp            --------------AGMHKKAFVFPK     PTHLCASWESSTGIAE    SIILGQDQDSFGGSFEKQQSLVGDIGNVNMWDYAL
human_crp            ------------QTDMSRKAFVFPK      PVHICTSWESASGIVE    SIILGQEQDSFGGNFEGSQSLVGDIGNVNMWDFVL
guinea_pig_crp       ------------GTDMSKKTFVFPK      PVHICTSWESVSGIIE    MIILGQDQDSFGGSFDANQSFVGDIGDVNMWDFVL
rat_crp             ------------HEDMSKQAFVFPG       PTHICATWESATGIVE    SIILGQESYGGGFDANQSLVGDIGDVNMWDFVL
golden_hamster_crp   ------------QKDMSKTAFVFPK      PTHICASWESATGIAE    SIILGQEQDSYGGGFDANQSLVGDIGDVNMWDIVL
pig_crp             ------------QTDMIGKAFVFPK       PMHFCMTWESTSGITE    SIILGQEQDAFAGGFEKNQCLVGDIGDVNMWDYVL
rat_sap             ------------QTDLNQKVFVFPR       PIHFCTSWESSSGIAE    SIVLGQEQDTYGGGFDKTQSFVGEIADLYMWDSVL
mouse_sap           -------------QTDLKRKVFVFPR     PVHLCTTWESSSGIVE    SIVLGQEQDNYGGGFQRSQSFVGEFSDLYMWDYVL
pig_sap             ------------HRDLTGKVFVFPR       PVHICTSWESSTGIAE    RIVLGQEQDSYGGGFDKTQSFVGEIGDLYMWGSVL
cattle_sap          ------------QTDLRGKVFVFPR       PVHICTSWESSTGIAE    KIVLGQEQDSYGGGFDKNQSFMGEIGDLYMWDSVL
human_sap           ------------HTDLSGKVFVFPR       PVHICVSWESSSGIAE    KIVLGQEQDSYGGKFDRSQSFVGEIGDLYMWDSVL
guinea_pig_sap      ------------QTDLDKKVFVFPR       PVHICTSWESSSGIAE    KTILGQEQDSYGGKFDRGQSFLGEIGDVYMWDSVL
golden_hamster_sap  ------------QTDLTGKVFVFPR       PVHFCTSWESSSGIAE    SIILGQEQDNYGGGFDNYQSFVGEIGDLNMWDSVL
armenian_hamster_sap ------------QTDLTGKVFVFPR      PVHFCTSWESSSGIAD    SIILGQEQDNYGGGFDKSQSFVGEMGDLNMWDSVL
carp_ptxn           FFCLLLSLTAAATEVGLVGKVLLFPT      GTHLCLTWDSETGLSA    TVLLGQDPDNYLGAFEVEQSFVGEITDVHMWDHVL
salmon_ptxn         ------------EHQDLSGKVFVIPM      WISICWTWDSKSGLTQ    SIMLVQEQDSYGGGFDVSQSFVGEVTDVHFWDSVI
trout_ptxn          ------------DLQDLSGKVFVIPM      WISICWTWDSTTGLTQ    SIMLVQEQDSYGGGFDSSQSFVGEVTDVHFWDSVI
dogfish_crp_Robey,  -SPV---AASYRATAGLAGKALDF--      ----------------    ----------------------------------
dogfish_crp_n-term1 --------------------------      --HYCLTW--------    ----------------------------------
dogfish_crp_n-term2 --------------------------      --HPCLTW--------    ----------------------------------
dogfish_crp_c-term  --------------------------      ----------------    --MVGEDQD-------------------------
dogfish_sap_Robey,  ---------------GFPGKSLIF--      ----------------    ----------------------------------
dogfish_sap_n-term  ---------------GFPGKS-----      ----------------    ----------------------------------
dogfish_sap_c-term1 --------------------------      ----------------    ------------------------EMTDVH------
dogfish_sap_c-term2 --------------------------      ----------------    ------------------------EMTNVH------
human CRP numbering  ..............1.........13      93...........108    132...............................166
```

Figure 5.5. Alignment shows sequences 1-13, 93-108, and 132-166 (human CRP numbering) where dogfish CRP and SAP fragments and primers align to known pentraxins. The green background indicates identical residues, whilst the yellow background indicates similar residues. The fragments of CRP and SAP sequences from work by Robey *et al.,* (1983), and the primers designed as shown in table 5.3 are highlighted in grey.

Table 5.3. Primers designed for amplification of *M. canis* CRP and SAP genes. For single letter codes please see table 5.2. CRP primers were designed from sequence data in this study, whilst SAP primers were designed from sequence data from Robey, Tanaka, and Liu (1983).

Primer	Sequence						
M. canis CRP N-terminus sequence	His	Tyr/Phe	Cys	Leu	Thr	Trp	
M. canis CRP primer N-terminus 5'-3'	CA (T/C)	TA(T/C)/TT(C)	TG(T/C)	TT(A/G)/CT(N)	AC(N)	TGG	
Variant 1	CAY	TAY	TGY	TTR	ACN	TGG	
Variant 2	CAY	TAY	TGY	CTN	ACN	TGG	
Variant 3	CAY	TTY	TGY	TTR	ACN	TGG	
Variant 4	CAY	TTY	TGY	CTN	ACN	TGG	
M. canis CRP C-terminus sequence	Met	Val	Gly	Glu	Asp	Gln	Asp
M. canis CRP primer C-terminus 3'-5'	TAC	CA(N)	CC(N)	CT(T/C)	CT(A/G)	GT(T/C)	CT
M. canis SAP N-terminus sequence	Gly	Phe	Pro	Gly	Lys	Ser	
M.canis SAP primer N-terminus 5'-3'	GG(N)	TT(C/T)	CC(N)	GG(N)	AA(A/G)	TC/AG	
Variant 1	GGN	TTY	CCN	GGN	AAA	TC	
Variant 2	GGN	TTY	CCN	GGN	AAG	TC	
Variant 3	GGN	TTY	CCN	GGN	AAA	AG	
Variant 4	GGN	TTY	CCN	GGN	AAG	AG	
M. canis SAP C-terminus sequence	Glu	Met	Thr	Asp/Asn	Val	His	
M. canis SAP primer C-terminus 3'-5'	CT(T/C)	TAC	TG(N)	CT(A/G)/TT(A/G)	CA(N)	GT	
Variant 1	TTC	TAC	TGN	TTR	CAN	GT	
Variant 2	TTC	TAC	TGN	CTR	CAN	GT	
Variant 3	CTC	TAC	TGN	TTR	CAN	GT	
Variant 4	CTC	TAC	TGN	CTR	CAN	GT	

5.3.4 PCR using new dogfish CRP primers on DNA from liver stored at -80°C

PCR at 55°C annealing temperature was set up using all the CRP N-terminus primers (1-4) and the CRP C-terminus primer (1) as the melting temperatures of the primers were all around 50°C, so performing PCR with an annealing temperature of 55°C should get more specific annealing. Gel electrophoresis showed no PCR products for the reactions with each of the primers pairs, or the blanks. There should have been bands at around 100bp as this was the expected size of the fragment the CRP primers should have amplified. So, PCR was repeated but at a lower annealing temperature of 50°C giving products around 100bp in both the reaction samples and the blanks (see figure 5.6). This could have been the dogfish CRP fragment, contamination, or the primers annealing together as it was also present in the blanks (lanes 5 and 7).
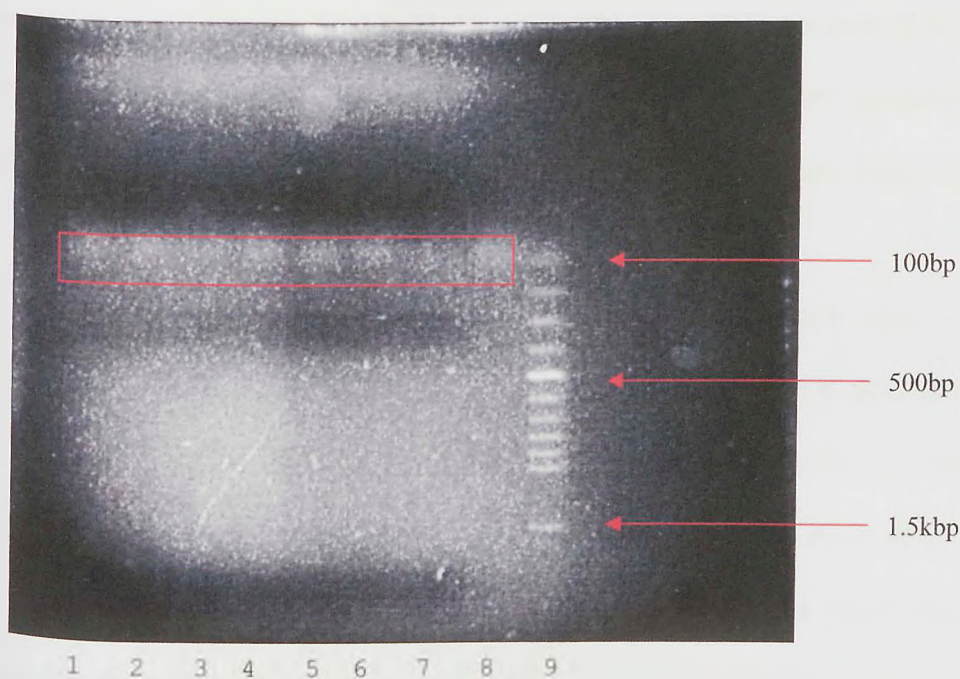
Figure 5.6. Gel showing products and blanks of PCR at annealing temperatures of 50°C with primers CRP N-terminus 1-4 and C-terminus 1. Lane 1) N-terminus 1 C-terminus 1 2)N-terminus 2 C-terminus 1 3)N-terminus 3 C-terminus 1 4)N-terminus 4 C-terminus 1 5)blank N-terminus 1 C-terminus 1 6)blank N-terminus 2 C-termins 1 7)blank N-terminus 3 C-terminus 1 8)blank N-terminus 4 C-terminus 1 9)100bp ladder. The box indicates where the 100bp band is.

In theory, only one set of primers from within the degenerate pool should anneal at the correct place and amplify part of the CRP gene. Therefore, only one band should be seen on the gel and not one for each reaction, and there should be no bands in the blanks.

However, as the primers are very similar, the low annealing temperature of the performed PCR may have encouraged annealing of all the primers even though they were not specific. This would explain the presence of bands in all the reaction lanes but not the blanks. There must still be contamination in the blank tubes whether it is dogfish DNA or DNA from another source. A gradient of annealing temperatures during PCR would show whether annealing of all primers is occurring at low temperatures, and as the temperature increases only the correct more specific primers should bind to the DNA and amplify the CRP gene.

Gradient PCR was performed for CRP N-terminus 1 C-terminus 1 and also N-terminus 2 C-terminus 1 at 47-57°C annealing temperatures. A gradient of 47-57°C was more suitable for the melting temperatures of the primers. It was unclear from the gel of PCR products for CRP N-terminus 1 C-terminus 1 whether there were any bands at the lower temperatures used in the gradient PCR. However, the same PCR gradient of 47-57°C using N-terminus 2 C-terminus 1 primers gave very faint bands at the lower temperatures around 400-500bp (lanes 1-5), and throughout all temperatures at ~100bp (see figure 5.7)

It was thought that the primers annealed to the dogfish DNA more so at lower temperatures probably due to their degenerate nature. The gradient PCR was repeated for CRP N-terminus 2 C-terminus 1 at annealing temperatures of 44-54°C as 54°C was the highest temperature that showed bands on figure 5.6. This gave three PCR products at lower temperatures (lanes 1-7) shown as triple bands on gel around 400-500bp and 700bp in size (see figure 5.8). There were also bands at <100bp for all temperatures indicating some contamination still.
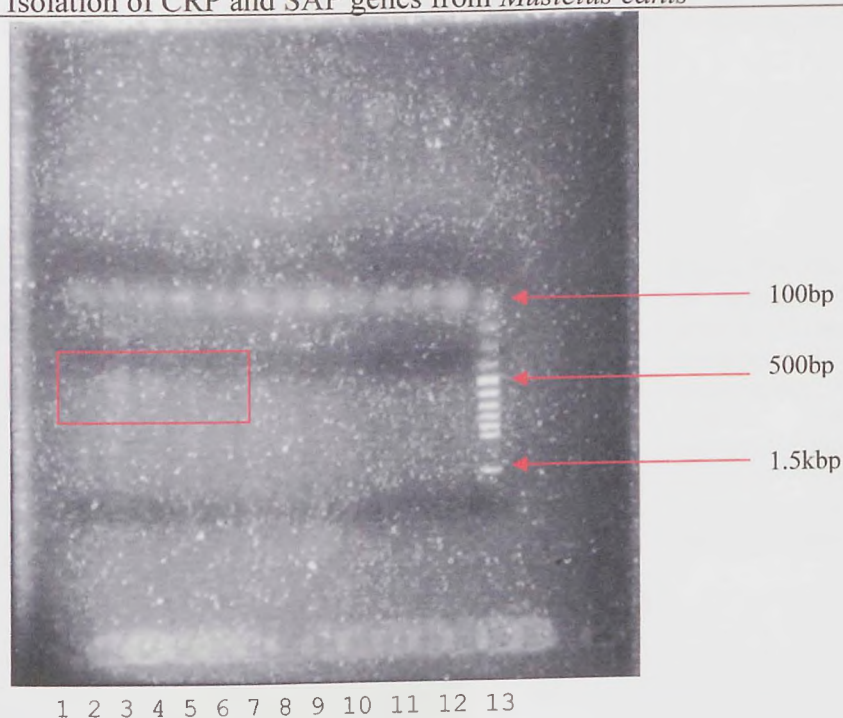
Figure 5.7. Gel showing products of gradient PCR with CRP primers N-terminus 2 C-terminus 1 at annealing temperatures 47-57°C. Lanes show product for each temperature covered by gradient PCR machine. Lane 1)47.0°C 2)47.3°C 3)47.9°C 4)48.8°C 5)49.9°C 6)51.2°C 7)52.5°C 8)53.9°C 9)55.4°C 10)56.3°C 11)56.8°C 12)57.1°C 13)100bp ladder
The red box indicates where the bands of 400-550bp are.
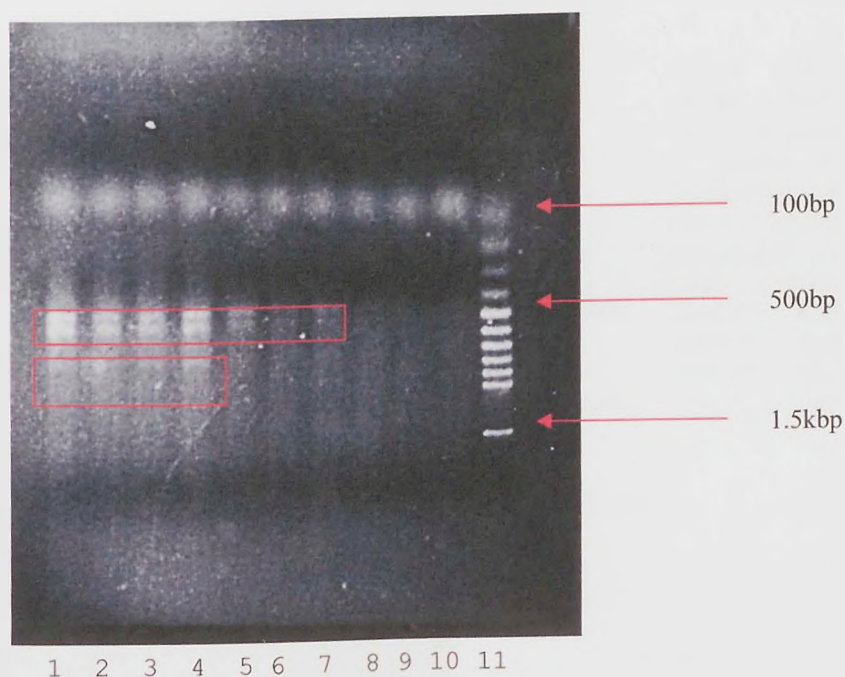
Figure 5.8. Gel showing products of gradient PCR with CRP primers N-terminus 2 C-terminus 1 at annealing temperatures 44-54°C. Lanes show product for each temperature on gradient PCR. Lane 1)44.9°C 2)45.8°C 3)46.9°C 4)48.2°C 5)49.5°C 6)50.9°C 7)52.4°C 8)53.3°C 9)53.8°C 10)54.1°C 11)100bp ladder
The red boxes indicate where the bands ~400-500bp and 700bp are.

The gradient PCR was repeated with CRP N-terminus 3 C-terminus 1 and N-terminus 4

C-terminus 1 primers at 44-54°C. Both PCR reactions showed the same triple bands on the

237

gel as the N-terminus 2 C- terminus 1 PCR products at lower annealing temperatures (see

figures 5.9 and 5.10).
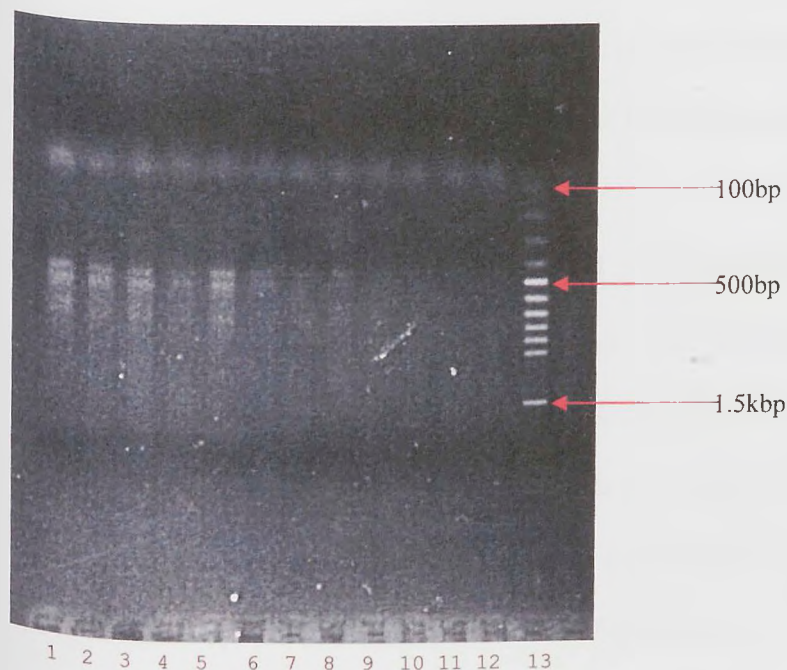
Figure 5.9. Gel showing products of gradient PCR with CRP primers N-terminus 3 and C-terminus 1 at annealing temperatures 44-54°C: Lane 1)44.0°C 2)44.3°C 3)44.9°C 4)45.8°C 5)46.9°C 6)48.2°C 7)49.5°C 8)50.9°C 9)52.4°C 10)53.3°C 11)53.8°C 12)54.1°C 13)100bp ladder
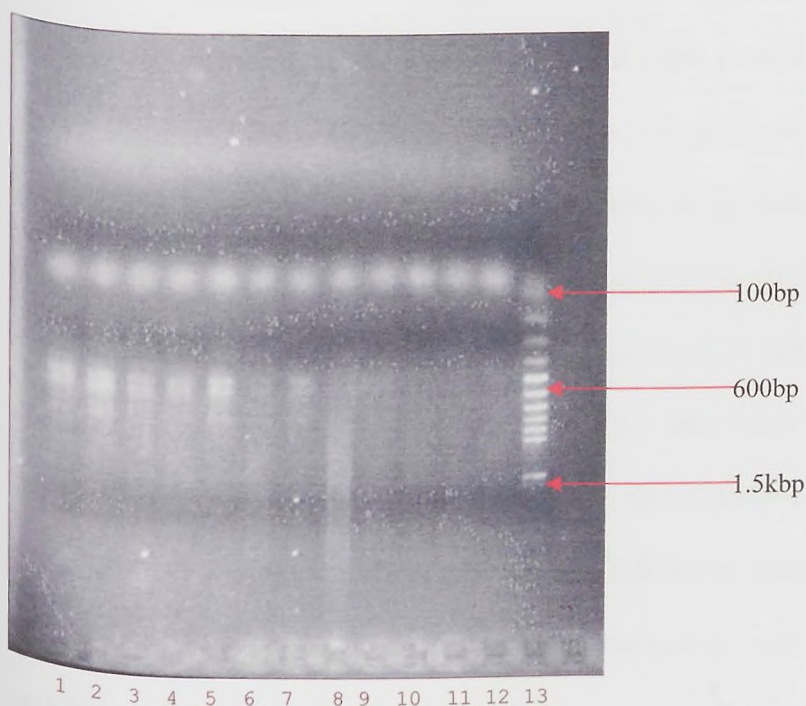
Figure 5.10. Gel showing products from gradient PCR with CRP primers N-terminus 4 C-terminus 1 at annealing temperature 44-54°C. Lanes show product for each temperature on gradient PCR. Lane 1)44.0°C 2)44.3°C 3)44.9°C 4)45.8°C 5)46.9°C 6)48.2°C 7)49.5°C 8)50.9°C 9)52.4°C 10)53.3°C 11)53.8°C 12)54.1°C 13)100bp ladder

This indicates that all three N-terminus primers 2-4 and the C-terminus primer anneal to the dogfish DNA at the lower temperatures of the gradient PCR. The presence of three bands indicates that the primers are annealing to three different areas of dogfish DNA and

therefore annealing is not specific to the CRP gene, if indeed they anneal to it at all. The largest band at ~700bp disappears at higher temperatures of the gradient PCR before the two smaller bands at ~400bp do, indicating that the two smaller bands have more specific primer annealing than the larger band, and one of these two bands may correspond to PCR product from the CRP gene although the size of the fragment is larger than expected.

It was noted that all the PCR reactions that produced fragments contained the CRP C-terminus primer and given that this is a pool of several different primers, it was suspected that within this pool there may be sequences acting as both 5' and 3' primers. PCR at an annealing temperature of 44-54°C was performed using just the CRP C-terminus primer which shows the same two larger bands of the triple bands seen from the PCR reactions for the N-terminus 2, 3, 4, and C-terminus 1 (see figure 5.11).

Therefore, initial general PCR gave 100bp DNA fragments that were the products of non-specific annealing as the bands on the gel were present in both reaction tubes and blanks. Gradient PCR with primers N-terminus 2, 3, and 4 with C-terminus 1 primer also produced the same band at 100bp and also three larger products but only two of these seemed to be specific as the third band at ~700bp was not amplified at higher temperatures. The two bands at ~400bp were also found to be non-CRP DNA products as they were the product of PCR using the C-terminus primer only. This meant that the C-terminus primer was annealing as both the 5' and 3' primer and therefore the amplified product was unlikely to be dogfish CRP DNA as amplification was not specific to sequence corresponding to dogfish CRP DNA. The non-specific annealing of the primers was probably due to both the low annealing temperatures of the gradient PCR and the degeneracy of the primers themselves.
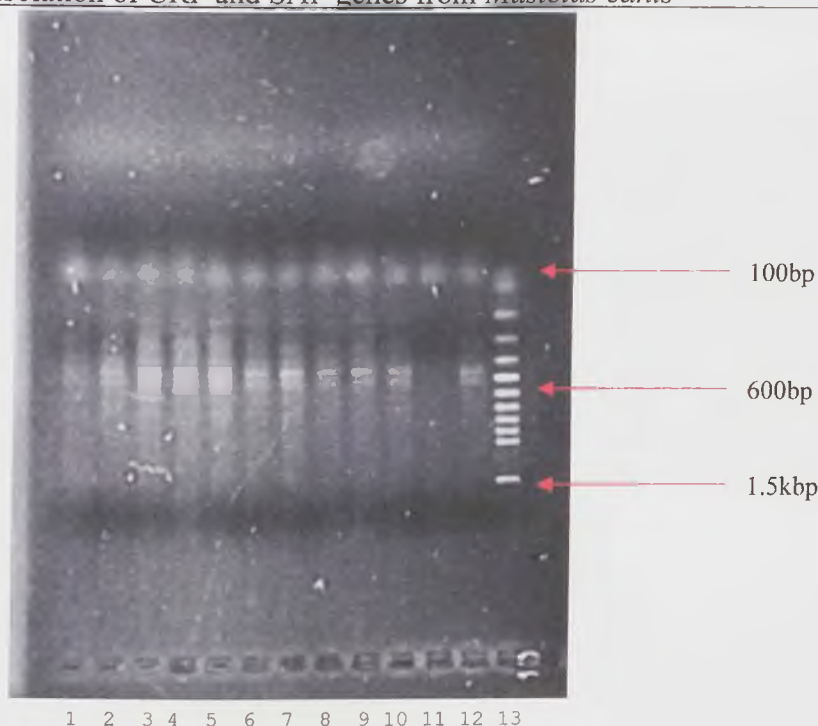
Figure 5.11. Gel showing products from gradient PCR with CRP primer C-terminus 1 at annealing temperature 44-54°C. Lanes show product for each temperature on gradient PCR. Lane 1)44.0°C 2)44.3°C 3)44.9°C 4)45.8°C 5)46.9°C 6)48.2°C 7)49.5°C 8)50.9°C 9)52.4°C 10)53.3°C 11)53.8°C 12)54.1°C 13)100bp ladder

5.3.5 PCR using dogfish SAP primers on DNA from liver stored at -80°C

Even though the gradient PCR using the CRP primers was not successful in terms of amplification of the dogfish CRP gene, it did show amplification of DNA. Therefore, gradient PCR was used to test the dogfish SAP primers. Multiple PCR were performed using primer combinations of N-terminus and C-terminus primers. However, only primer combinations of N-terminus (1, 3 and 4) C-terminus 4 appeared to produce PCR products.

Gradient PCR of annealing temperatures 45-55°C using SAP primers N-terminus 4 C-terminus 4 gave two products at lower annealing temperatures (lanes 1-6 see figure 5.12) with the smaller band around 300bp which was the size of the expected amplified dogfish SAP fragment, and the larger band ~700bp. As both bands disappeared at the higher temperatures they were unlikely to be the result of contamination, but it also indicated that primer annealing was not specific as there were two products.
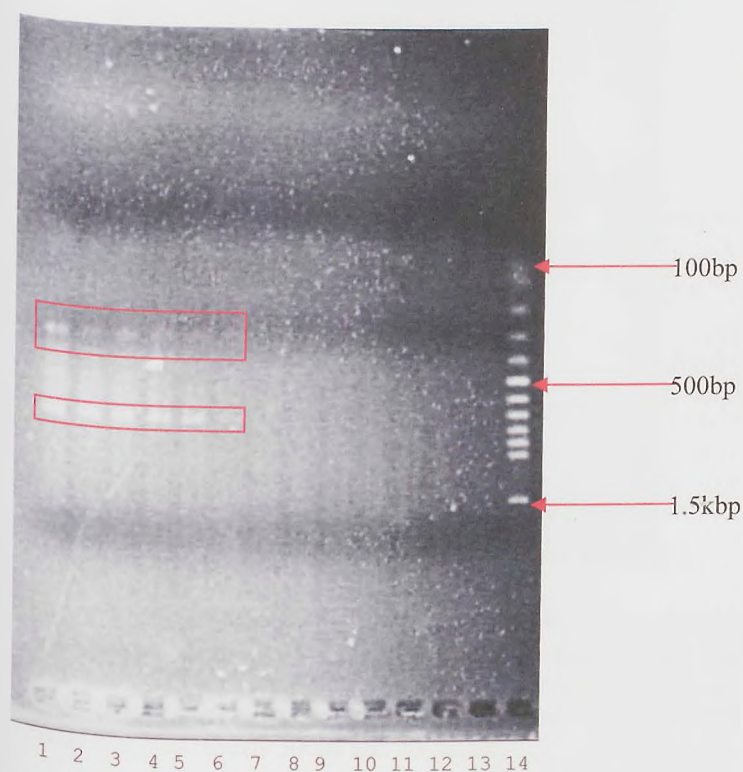
Figure 5.12. Gel showing products of gradient PCR with SAP primers N-terminus 4 and C-terminus 4 at annealing temperatures 45-55°C. Lanes show product for each temperature on gradient PCR. Lane 1)45.0°C 2)45.3°C 3)45.9°C 4)46.8°C 5)47.9°C 6)49.2°C 7)50.5°C 8)51.9°C 9)53.4°C 10)54.3°C 11)54.8°C 12)55.1°C 13)PCR reaction with no DNA (primers on own) 14)100bp ladder. Red boxes indicate where the bands of sizes ~300bp and 700bp are.

Gel electrophoresis of gradient PCR products using combinations of N-terminus (1-3) and the C-terminus 4 primers were performed to see whether they gave the same sized products as the N-terminus 4 C-terminus 4 primers. Gradient PCR at annealing temperature of 45-55°C for SAP primers N-terminus 1 C-terminus 4, and N-terminus 3 C-terminus 4, gave the same two bands that disappeared at higher temperatures (see figures 5.13 and 5.14), but that with primers N-terminus 2 C-terminus 4 showed no bands.

1 2 3 4 5 6   7 8 9 10 11 12 13 14
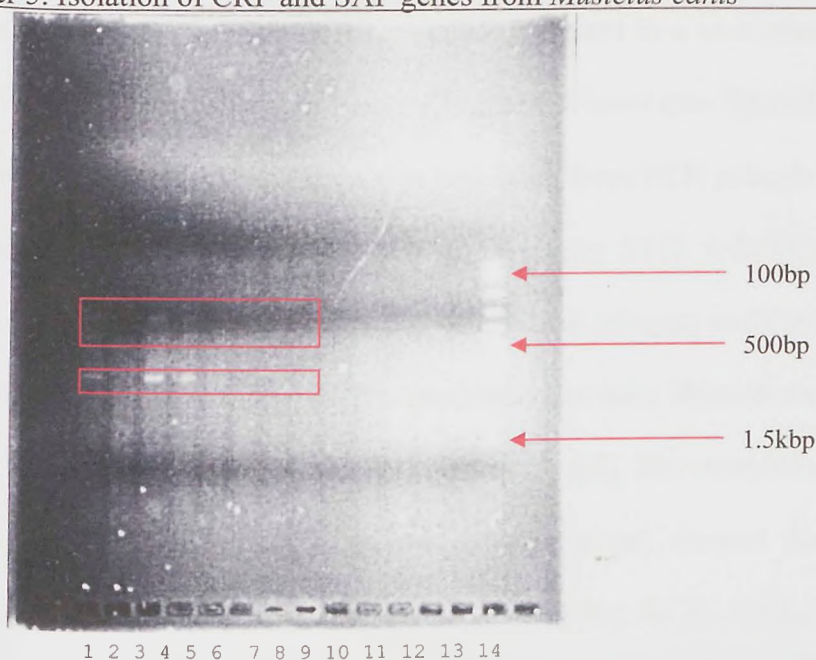
Figure 5.13. Gel showing products of gradient PCR with SAP primers N-terminus 1 and C-terminus 4 at annealing temperatures 45-55°C. Lanes show product for each temperature on gradient PCR. Lane 1)45.0°C 2)45.3°C 3)45.9°C 4)46.8°C 5)47.9°C 6)49.2°C 7)50.5°C 8)51.9°C 9)53.4°C 10)54.3°C 11)54.8°C 12)55.1°C 13)PCR reaction with no DNA (primers on own) 14)100bp ladder. Red boxes indicate where the bands at approximately 300bp and 700bp are.

1 2   3 4   5 6   7 8 9 10 11 12 13 14
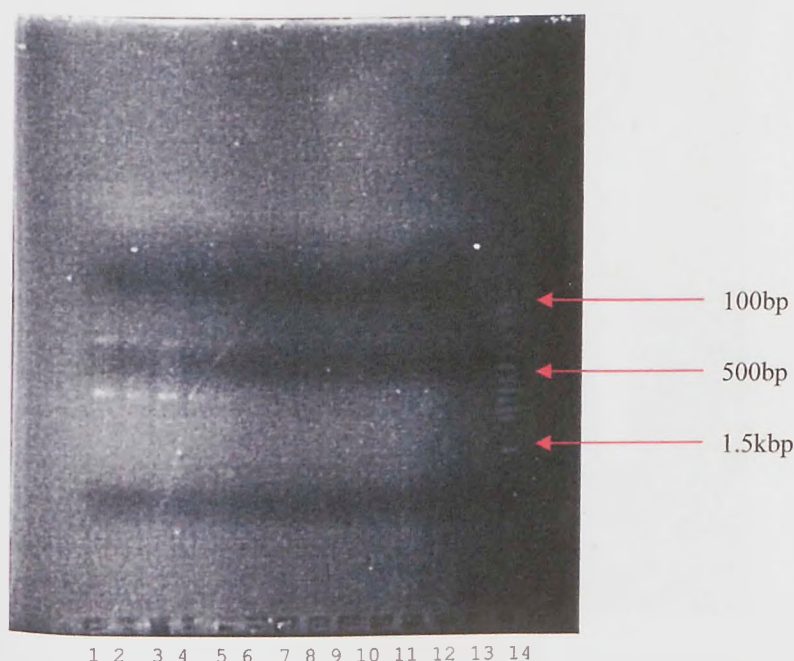
Figure 5.14. Gel showing products of gradient PCR with SAP primers N-terminus 3 and C-terminus 4 at annealing temperatures 45-55°C. Lanes show product for each temperature on gradient PCR. Lane 1)45.0°C 2)45.3°C 3)45.9°C 4)46.8°C 5)47.9°C 6)49.2°C 7)50.5°C 8)51.9°C 9)53.4°C 10)54.3°C 11)54.8°C 12)55.1°C 13)PCR reaction with no DNA (primers on own) 14)100bp ladder.

Topo cloning was used with PCR product from the SAP N-terminus 1 C-terminus 4, and N-terminus 4 C-terminus 4 reactions at the temperature which showed the brightest band which was around 46°C (lane 3 figure 5.13 and 5.14). After Topo cloning, six white colonies were taken from each plate (i.e. six for each PCR product) and grown in LB-

Broth. DNA was isolated from each sample and used in a restriction digest with *Eco*R1

which should cut the Topo vector either side of the insert (see figure 5.2). Insert fragments

were found in DNA from colonies that had insert from PCR using both the N-terminus 1

C-terminus 4, and N-terminus 4 C-terminus 4 (figure 5.15). A DNA preparation from the

colony 2 and 4 for the N-terminus 1 C-terminus 4 primers, and colonies 1 and 6 for N-

terminus 4 C-terminus 4 primers was performed as these showed some bands present on

the gel at ~300bp (these bands are very faint in the gel). The restriction digest was repeated

on the newly isolated DNA from the colonies which showed the same sized insert

fragments present on the gel previously suggesting that the *E. coli* had amplified the PCR

product thought to be the dogfish SAP DNA fragment.
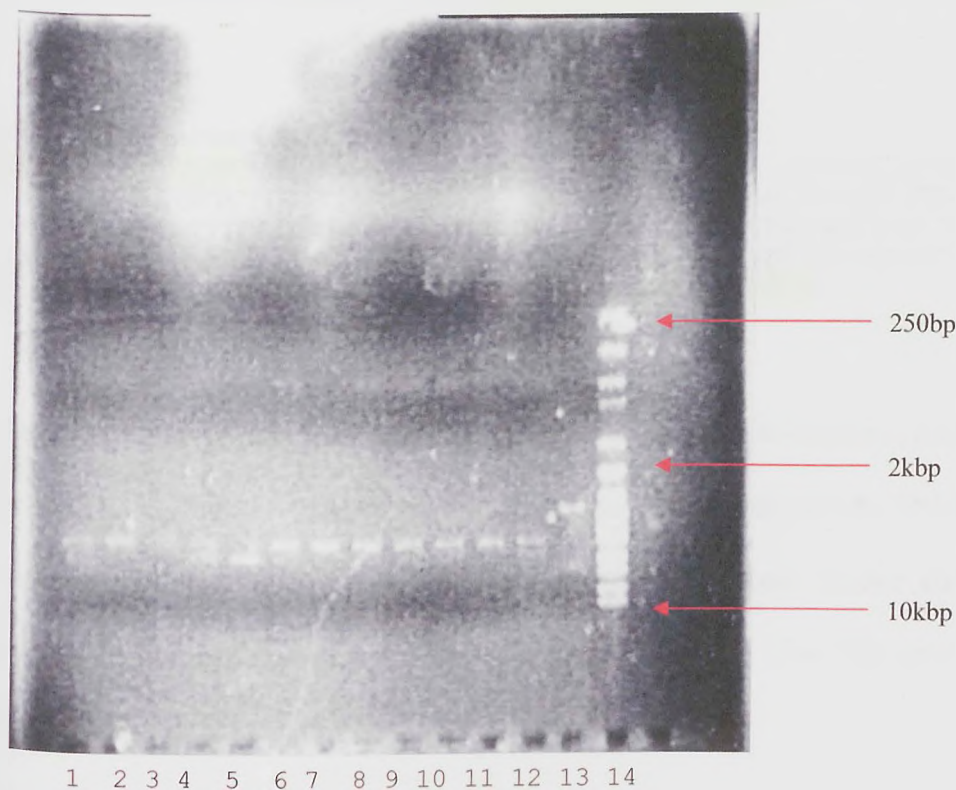
Figure 5.15. 0.8% gel showing products of restriction digest with *Eco*R1 of DNA from twelve white colonies from two agar plates of *E. coli* Topo cloned with PCR product from SAP primers N-terminus 1 C-terminus 4 and N-terminus 4 C-terminus 4. Lane 1)colony 1 N-terminus 1 C-terminus 4 2)colony 2 N-terminus 1 C-terminus 4 3)colony 3 N-terminus 1 C-terminus 4 4)colony 4 N-terminus 1 C-terminus 4 5)colony 5 N-terminus 1 C-terminus 4 6)colony 6 N-terminus 1 C-terminus 4 7)colony 1 N-terminus 4 C-terminus 4 8)colony 2 N-terminus 4 C-terminus 4 9) colony 3 N-terminus 4 C-terminus 4 10)colony 4 N-terminus 4 C-terminus 4 11)colony 5 N-terminus 4 C-terminus 4 12)colony 6 N-terminus 4 C-terminus 4 13)uncut plasmid 14)1kbp ladder

The isolated DNA was sent for sequencing to identify whether the fragment did actually

belong to dogfish SAP (see section 5.2.12). Sequencing at MWG was performed using two

primers: M13 reverse and T7 which are either side of where the insert area lies on the Topo

vector (see figure 5.2) and therefore would both allow sequencing of the insert fragment.

Of the 8 sequencing reactions, only 5 worked and only two had a primer sequence at

both ends. However, the primer sequences were both for the C-terminus primer as shown

by the sequence below where the primer is highlighted in yellow:

N-terminus 1 C-terminus 4 PCR product:
5' TGTACATCTGTCATCTC TCCACCCAAGACTCCAACCGATCTATGTCCTGCTGTATCCTCTGATG
GTCCTCATCACTATCTGCAATTCCACCAACCTTTGTGTCGTCCGCAAACTTACTAATCAAACCAG
TTACGTTTTCCTCCAAATCATTTATATATATTACAAACAGCAAAGGTGCCAGCACTGTTCTCTGA
GGAACGCCACATACCCCAGCCCTCCAATCAGAAACACACCCTTGCACTGCTACCCTCTGTCTTC
TATGATCAAGCCAGTTCTGTATCCATCTTGCCAGTTCACTCCTGATCCCGTGCGACTTCACCTTC
TGTACCAGTCTGGCATGAGAGACCTTATGTG GAGATGACCGACGTGCA 3'

N-terminus 4 C-terminus 4 PCR product
5' TGAACATCGGTCATCTC TTCGAAAGATTCAATCAAGTTGGTCAGACATGACTTCCCCTTAACA
AAACCACGCTGACTGTTCTTGATTAATCTCTGCCTCTCCAAATGCAAACTAACTCTGTTTCTCAG
AATTGCTTCCAATAGTTTCCCCACCACTGAGGTCAGACTGACTGGCCTGTAGTTTCCTGGATTAT
CCCTTCCTCCTTGTTTGAATAATGTTACCACGTTGACTGTCCTCCAGTCCTCCAGCACTTCTCCTG
TGACCCAGAGAGGGATTGAAACTTATTATGTG GAGATGACAGACGTTCA 3'

As the C-terminus primer had acted as both C-terminus and N-terminus primer rather

than the N-terminus and the C-terminus both annealing, it is likely that the DNA that was

amplified was not that for dogfish SAP because the N-terminus primer should have

annealed to a complimentary sequence in the dogfish SAP gene. The PCR products were

both translated into all six reading frames.

Of the translated protein sequences, two of the reading frames matched the amino acid

sequence corresponding to the primer (highlighted in yellow):

N-terminus 1 C-terminus 4 PCR product translated 5'-3' frame 1:
CTSVISPPKTPTDLCPAVSSDGPHHYLQFHQPLCRPQTY-SNQLRFPPNHLYILQTAKVP
ALFSEERHIPQPSNQKHTLALLPSVFYDQASSVSILPVHS-SRATSPSVPVWHERPYV EM
TDV

N-terminus 1 C-terminus 4 PCR product translated 3'-5' frame 1:
CTSVIST-GLSCQTGTEGEVARDQE-TGKMDTELA-S-KTEGSSARVCF-
LEGWGMWRSSENSAGTFAVCNIYK-FGGKRNWFD--VCGRHKGWWNCR---
GPSEDTAGHRSVGVLGG==EMTDV==

N-terminus 4 C-terminus 4 PCR product translated 5'-3' frame 1:
-TSVISSKDSIKLVRHDFPLTKPR-LFLINLCLSKCKLTLFLRIASNSFPTTEVRLTGL-
FPGLSLPPCLNNVTTLTVLQSSSTSPVTQRGIETYYV==EMTDV==

N-terminus 4 C-terminus 4 PCR product translated 3'-5' frame 1:
-TSVIST-VSIPLWVTGEVLEDWRTVNVVTLFKQGGRDNPGNYRPVSLTSVVGKLLEAI
LRNRVSLHLERQRLIKNSQRGFVKGKSCLTNLIESFE==EMTDV==

A nucleotide-nucleotide BLAST search was performed with both the PCR product sequences, and also protein-protein BLAST and protein-translated nucleotide BLAST searches with the translated protein sequences from all six reading frames of both sequences. In parts of the sequences, there was high similarity to non-pentraxin DNA from other elasmobranchs such as *Ginglymostoma cirratum* (Nurseshark) DNA clone which showed 32-73% identity over 6-31% of the query sequence; *Scyliorhinus torazane* (Cloudy catshark) DNA for HER1 line which showed 30%-59% identity over 9-30% of the query sequence; and *Callorhinchus milli* (Elephant shark) DNA clone that showed 30-52% identity over 24-30% of the query sequence. Therefore, it appeared that the primers had amplified dogfish DNA, but it was not that of pentraxin origin.

5.3.6 Design of new *M. canis* CRP primers

As previous primers designed for amplifying the CRP genes in dogfish DNA were unsuccessful, design of another set of primers using peptides from a trypsin digest of dogfish CRP was attempted. The N-terminal amino acid sequences of peptides produced from trypsin cleavage (see section 5.2.4) are shown below in table 5.4. There were several N-terminal sequences shown for fragment 1 as it was not a single band on the gel.

Table 5.4 N-terminal sequences of peptide fragments from trypsin digest of dogfish CRP.

Fragment 1			Fragment 2
Sequence 1	Sequence 2	Sequence 3	Sequence 1
Valine	Phenylalanine	Glycine	Serine
Alanine	Threonine		Leucine
Asparagine	Histidine	Glycine	Alanine
Serine	Glutamic acid	Lysine	Threonine
Aspartic acid	Proline	Leucine	Alanine
Glutamine	Threonine	Glycine	Glycine
Glutamic acid	Serine		Leucine
Asparagine	Valine		Asparagine
Asparagine	Isoleucine	Leucine	Glycine
Alanine	Leucine		Lysine
Aspartic acid	Glycine	Tryptophan	
Aspartic acid			
Asparagine	Threonine		
Valine	Asparagine	Methionine	
Leucine			

These fragments were aligned with some known pentraxin fragments and those already discovered for dogfish CRP and SAP (Robey *et al.*, 1983) as performed previously in section 5.3.1. A combination of the three sequences from fragment 1 of the trypsin digest was performed to obtain the best alignment with a conserved pentraxin sequence near the C-terminus end with few conserved residues and inclusion of gaps (see figure 5.16). Fragment 2 from the trypsin digest aligned at around a similar area as the dogfish CRP and SAP peptide sequences previously discovered by Robey *et al.*, (1983) (see figure 5.16).

```
rabbit_crp                -------------AGMHKKAFVFPKESDN    ASIILGQDQDSFGGSFEKQQSLVGDIGNVNMWDYALSP
human_crp                 ------------QTDMSRKAFVFPKESDT    ASIILGQEQDSFGGNFEGSQSLVGDIGNVNMWDFVLSP
guinea_pig_crp            ------------GTDMSKKTFVFPKETDN    AMIILGQDQDSFGGSFDANQSFVGDIGDVNMWDFVLSP
rat_crp                   ------------HEDMSKQAFVFPGVSAT    ASIILGQEQDSYGGGFDANQSLVGDIGDVNMWDFVLSP
golden_hamster_crp        ------------QKDMSKTAFVFPKESAN    ASIILGQEQDSYGGGFDANQSLVGDIGDCNMWDIVLSP
pig_crp                   ------------QTDMIGKAFVFPKESEN    ASIILGQEQDAFAGGFEKNQCLVGDIGDVNMWDYVLSP
pig_sap                   ------------HRDLTGKVFVFPRESAT    PRIVLGQEQDSYGGGFDKTQSFVGEIGDLYMWGSVLSP
cattle_sap                ------------QTDLRGKVFVFPRESST    PKIVLGQEQDSYGGKFDRSQSFVGEIGDLYMWDSVLSP
human_sap                 ------------HTDLSGKVFVFPRESVT    PKIVLGQEQDSYGGGFDKNQSFMGEIGDLYMWDSVLPP
guinea_pig_sap            ------------QTDLDKKVFVFPRESSS    PKIILGQEQDSYGGKFDRGQSFLGEIGDVYMWDSVLSP
golden_hamster_sap        ------------QTDLTGKVFVFPRQSET    PSIILGQEQDNYGGGFDNYQSFVGEIGDLNMWDSVLTP
armenian_hamster_sap      ------------QTDLTGKVFVFPRESES    PSIILGQEQDNYGGGFDKSQSFVGEMGDLNMWDSVLTP
rat_sap                   ------------QTDLNQKVFVFPRESET    PSIVLGQEQDTYGGGFDKTQSFVGEIADLYMWDSVLTP
mouse_sap                 ------------QTDLKRKVFVFPRESET    PSIVLGQEQDNYGGGFQRSQSFVGEFSDLYMWDYVLTP
salmon_ptxn               -----------EHQDLSGKVFVIPMATST    PSIMLVQEQDSYGGGFDVSQSFVGEVTDVHFWDSVISO
trout_ptxn                -----------DLQDLSGKVFVIPMTTST    PSIMLVQEQDSYGGGFDSSQSFVGEVTDVHFWDSVISP
carp_pentraxin            FFCLLLSLTAAATEVGLVGKVLLFPTKTNT  GTVLLGQDPDNYLGAFEVEQSFVGEITDVHMWDHVLSG
Limulus_CRP1.1            -----------LEEGEITSKVKFPPSSSPS   GLVVLGQDQDSVGGKFDATQSLEGELSELNLWNTVLNH
Limulus_CRP3.3            -----------LEEGEITSKVKFPPSSSPS   GLVVLGQDQDSVGGKFDATQSLEGELSELNLWNTVLNH
Limulin_fragment          -----------LEEGEITSKVKFPPSSSPS   -------------------------------------
Limulus_CRP1.4            -----------LEEGEITSKVKFPPSSSPS   GLVVLGQEQDSVGGEYDAEQSLEGELSELNLWNTVLNH
Tachypleus_1              ----------------KVKFPPSSSPS      GLVVLGQDQDTVGGGFDAKQSLEGELSELNLWDTVLNH
Tachypleus_8              ----------------KVKFPPSSSPS      GLVVIGQDQDSVGGGFDEKQSLVGELSELNLWDMVLNH
Tachypleus_16             ----------------KVKFPPSTASS      GTVVLGQDQDTVGGGFVASESMEGELSELNMWNSVLNS
Limulus_SAP               -----------AVDIRDVKISFPGTQNPK    GTVVIGQEQDKIGGGFEEQESWSGELSDLQVWDEALTT
dogfish_sap_Robey et al   --------------GFPGKSLIF------    -------------------------------------
dogfish_crp_Robey et al   ----SPVAASYRATAGLAGKALDF------  -------------------------------------
dogfish_CRP_tryp1         ----------------------------    ------------------FAGELGEVNLWDTVL--
dogfish_CRP_tryp2         ---------SL-ATAGLNGK----------   -------------------------------------
human CRP numbering       ...............1.............17  136..................................173
```

Figure 5.16. Alignment shows sequences 1-17 and 136-173 (human CRP numbering) where dogfish CRP fragments 1 and 2 from trypsin digest align to known pentraxins. The green background indicates identical residues, whilst the yellow background indicates similar residues. Fragments 1 and 2 from trypsin digest of dogfish DNA are highlighted in grey as are the previously discovered dogfish CRP and SAP peptide sequences by Robey *et al*., (1983).

As the fragment 1 from the trypsin cleavage was poorly aligned, and fragment 2 aligned close to an area of sequence previously used unsuccessfully for primer design, it was decided that another set of primers were to be designed from the areas of sequence that had been sequenced previously and areas of sequence deduced from electron density maps of dogfish CRP. The areas of sequence from which primers were designed are as follows:

- Amino acids (15-21) Glycine, Lysine, Alanine, Leucine, Aspartic acid, Phenylalanine, Alanine: have been sequenced several times in this present work and in previous work (Robey *et al.*, 1983) and will form the basis for N-terminus primer 1.

- Amino acids (44-50) Threonine, Histidine/Phenylalanine, Cysteine, Methionine, Arginine/Lysine, Tryptophan, Alanine: have not been sequenced but deduced from electron density maps of dogfish CRP, with the TXC being a conserved pentraxin sequence. These will form the basis for N-terminus primers 2-5

- Amino acids (106-113) Glutamic acid/Glutamine, Histidine, Tyrosine, Cysteine, Leucine, Threonine, Tryptophan, Aspartic acid: has not been sequenced but deduced from electron density maps of dogfish CRP with HXCXS/TW being a conserved pentraxin motif. These will form the basis for N-terminus primer 6

- Amino acids (146-153) Methionine, Methionine, Valine, Glycine, Glutamic acid, Aspartic acid, Glutamine, Aspartic acid: have been N-terminal amino acid sequenced a few times in this work, and agree with conservation throughout the pentraxins and the electron density map of dogfish CRP. These will form the basis for the C-terminus primer.

Although some of the primers to be designed were similar to those designed previously, the addition of new primers and the use of inosine as a nucleotide base substitution in place of the third codon position where there was high degeneracy, was hoped to increase the chance of amplification of the dogfish CRP gene. The primers designed are shown in table 5.5.

Primer	Sequence							
M. canis CRP N-terminus sequence 1	Gly	Lys	Ala	Leu	Asp	Phe	Ala	
M. canis CRP primer N-terminus 1 5'-3'	GGI	AA(AG)	GCI	(CT)TI	GA(CT)	TT(CT)	GCI	
M. canis CRP N-terminus sequence 2	Thr	Phe	Cys	Met	Arg	Trp	Ala	
M. canis CRP primer N-terminus 2 5'-3'	ACI	TT(CT)	TG(CT)	ATG	(AC)GI	TGG	GCI	
M. canis CRP N-terminus sequence 3	Thr	Phe	Cys	Met	Lys	Trp	Ala	
M. canis CRP primer N-terminus 3 5'-3'	ACI	TT(CT)	TG(CT)	ATG	AA(AG)	TGG	GCI	
M. canis CRP N-terminus sequence 4	Thr	His	Cys	Met	Arg	Trp	Ala	
M. canis CRP primer N-terminus 4 5'-3'	ACI	CA(CT)	TG(CT)	ATG	(AC)GI	TGG	GCI	
M. canis CRP N-terminus sequence 5	Thr	His	Cys	Met	Lys	Trp	Ala	
M. canis CRP primer N-terminus 5 5'-3'	ACI	CA(CT)	TG(CT)	ATG	AA(AG)	TGG	GCI	
M. canis CRP N-terminus sequence 6	Glu/Gln	His	Tyr	Cys	Leu	Thr	Trp	Asp
M. canis CRP primer N-terminus 6 5'-3'	(G/C)A(G/A)	CA(CT)	TA(CT)	TG(CT)	(CT)TI	ACI	TGG	GA(CT)
M. canis CRP C-terminus sequence	Met	Met	Val	Gly	Glu	Asp	Gln	Asp
M. canis CRP primer C-terminus 3'-5'	TAC	TAC	CA(A/C/G/T)	CCI	CT(C/T)	CT(A/G)	GT(C/T)	CT(A/G)

Table 5.5 Primers designed for amplification of *M. canis* CRP genes. For single letter codes please see table 5.2.

5.3.7 PCR using new *M. canis* CRP primers

Gradient PCR was performed with the primers in table 5.5 using the general PCR program described in section 5.2.8. The annealing temperatures for gradient PCR were determined from the melting temperatures of the primers.

Gradient PCR was performed using the *M. canis* primers N-terminus 1 and C-terminus as detailed in table 5.5. The annealing temperature for the PCR ranged from 50-65°C. A 1% gel was run of the PCR products for each annealing temperature, which showed bands at approximately 300-400 base pairs up to 60.3°C (lane 8) and ~700bp-800bp (lanes 1-6) (see figure 5.17).
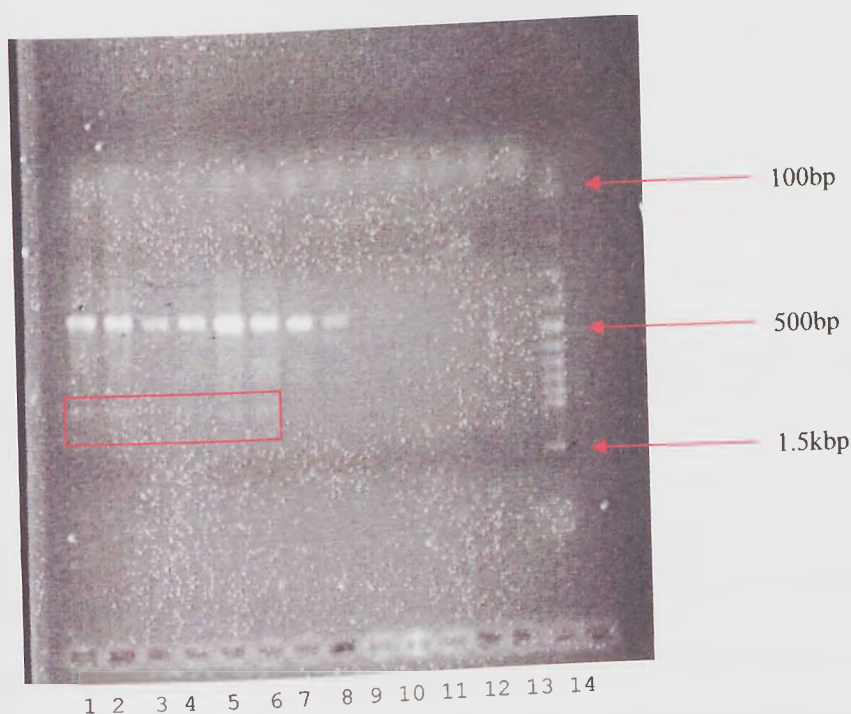
Figure 5.17. A 1% gel of PCR products from gradient PCR from 50-65°C, using *M. canis* CRP primers N-terminus 1 and C-terminus. Lanes show PCR products for PCR performed at the following annealing temperatures: 1) 50.1°C, 2) 50.5°C, 3) 51.3°C, 4) 52.7°C, 5) 54.4°C, 6) 56.4°C, 7) 58.3°C, 8) 60.3°C, 9) 62.6°C 10) 64.0°C, 11) 64.7°C, 12) 65.2°C, 13) empty, 14) 100bp ladder. The red box shows where there are products 700-800bp in size.

The bands of 300-400bp were approximately the expected size for a PCR product that corresponded to dogfish CRP amplified between the N-terminus 1 primer and C-terminus primer which was approximately 140 amino acids and therefore approximately 420 base pairs.

Gradient PCR was performed using the *M. canis* primers N-terminus 2 and C-terminus as detailed in table 5.5. The annealing temperatures for the PCR ranged from 55-65°C. A 1% gel of the PCR products for each annealing temperature showed bands between 300 and 400 base pairs in size up to annealing temperatures of 59.2°C (lane 6) (see figure 5.18).
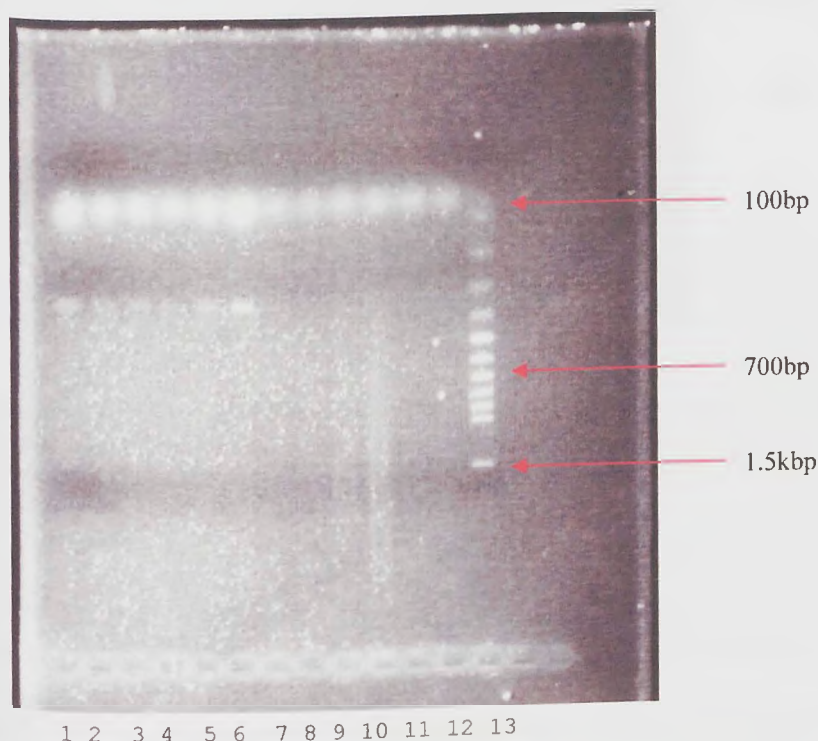
Figure 5.18. A 1% gel of PCR products from gradient PCR from 55-65°C, using *M. canis* CRP primers N-terminus 2 and C-terminus. Lanes show PCR products for PCR performed at the following annealing temperatures: 1) 55.0°C, 2) 55.3°C, 3) 55.9°C, 4) 56.8°C, 5) 57.9°C, 6) 59.2°C, 7) 60.5°C, 8) 61.9°C, 9) 63.4°C 10) 64.3°C, 11) 64.8°C, 12) 65.1°C, 13) 100bp ladder

The bands were approximately the expected size for a PCR product that corresponded to dogfish CRP amplified between the N-terminus 2 primer and C-terminus primer which was approximately 110 amino acids long and therefore approximately 330 base pairs in size.

Gradient PCR was performed using the *M. canis* primers N-terminus 3 and C-terminus as detailed in table 5.5. The annealing temperature for the PCR ranged from 50-65°C. A 1% gel of the PCR products for each annealing temperature showed bands between 300 and 400 base pairs in size up to annealing temperatures of 60.3°C (lane 8) (see figure 5.19).
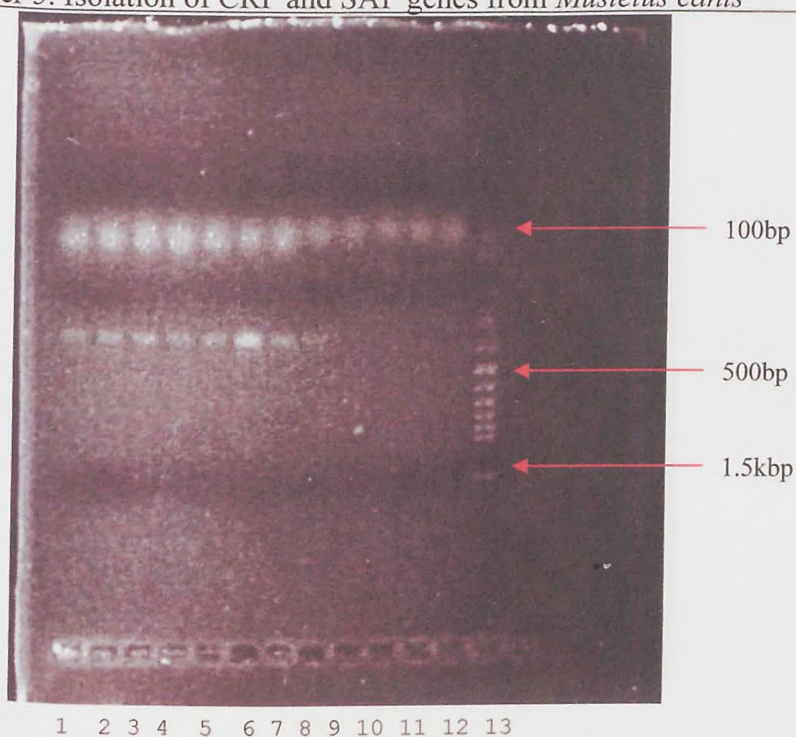
Figure 5.19. A 1% gel of PCR products from gradient PCR from 50-65°C, using *M. canis* CRP primers N-terminus 3 and C-terminus. Lanes show PCR products for PCR performed at the following annealing temperatures: 1) 50.1°C, 2) 50.5°C, 3) 51.3°C, 4) 52.7°C, 5) 54.4°C, 6) 56.4°C, 7) 58.3°C, 8) 60.3°C, 9) 62.6°C 10) 64.0°C, 11) 64.7°C, 12) 65.2°C, 13) 100bp ladder

The bands were approximately the expected size for a PCR product that corresponded to dogfish CRP amplified between the N-terminus 3 primer and C-terminus primer which was approximately 110 amino acids long and therefore approximately 330 base pairs in size.

Gradient PCR was performed using the *M. canis* primers N-terminus 4 and C-terminus as detailed in table 5.5. The annealing temperature for the PCR ranged from 55-65°C. A 1% gel of the PCR product for each annealing temperature showed bands between 300 and 400 base pairs in size (see figure 5.20), at annealing temperatures 55.0°C (lane 1), 56.8°C (lane 4) and 57.9°C (lane 5) with doublet bands appearing at the last two temperatures.
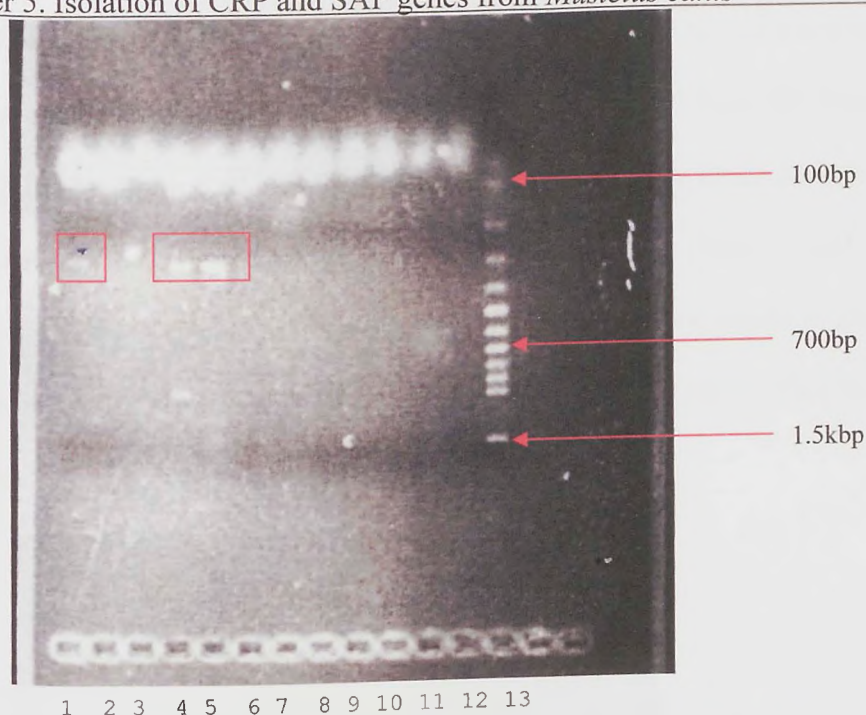
Figure 5.20. A 1% gel of PCR products from gradient PCR from 55-65°C, using *M. canis* CRP primers N-terminus 4 and C-terminus. Lanes show PCR products for PCR performed at the following annealing temperatures: 1) 55.0°C, 2) 55.3°C, 3) 55.9°C, 4) 56.8°C, 5) 57.9°C, 6) 59.2°C, 7) 60.5°C, 8) 61.9°C, 9) 63.4°C 10) 64.3°C, 11) 64.8°C, 12) 65.1°C, 13) 100bp ladder. Red boxes show where the bands at ~300-400bp are.

The bands were approximately the expected size for a PCR product that corresponded to dogfish CRP amplified between the N-terminus 4 primer and C-terminus primer which was approximately 110 amino acids long and therefore approximately 330 base pairs in size.

Gradient PCR was performed using the *M. canis* primers N-terminus 5 and C-terminus as detailed in table 5.5 at annealing temperatures from 50-65°C, and for *M. canis* primers N-terminus 6 and C-terminus as detailed in table 5.5 at annealing temperatures from 55-65°C. Neither of these PCR reactions showed any bands on gels.

As the PCR reactions that worked (i.e. with N-terminus 1-4 and C-terminus primers) showed bands of an approximate correct size for dogfish CRP DNA, samples of the PCR products were used in Topo cloning reactions as described previously in section 5.2.10. Three bacterial colonies for each PCR product were selected and DNA was extracted from the *E .coli* using the methods described in section 5.2.11. Extracted DNA was used in a

restriction digest with *Eco*R1 to determine whether the PCR product had been inserted into the plasmid DNA (described in section 5.2.6). The products from the digest were run on a 0.8% gel (see figure 5.21) which showed that only two colonies had plasmid with PCR product insert: colonies N-terminus 1 C-terminus colony 1 (lane 1), and N-terminus 1 C-terminus colony 3 (lane 3). None of the other lanes showed bands at approximately 300 base pairs in size (i.e. the size of the PCR product insert). DNA from these two colonies was sent for sequencing as described in section 5.2.12.
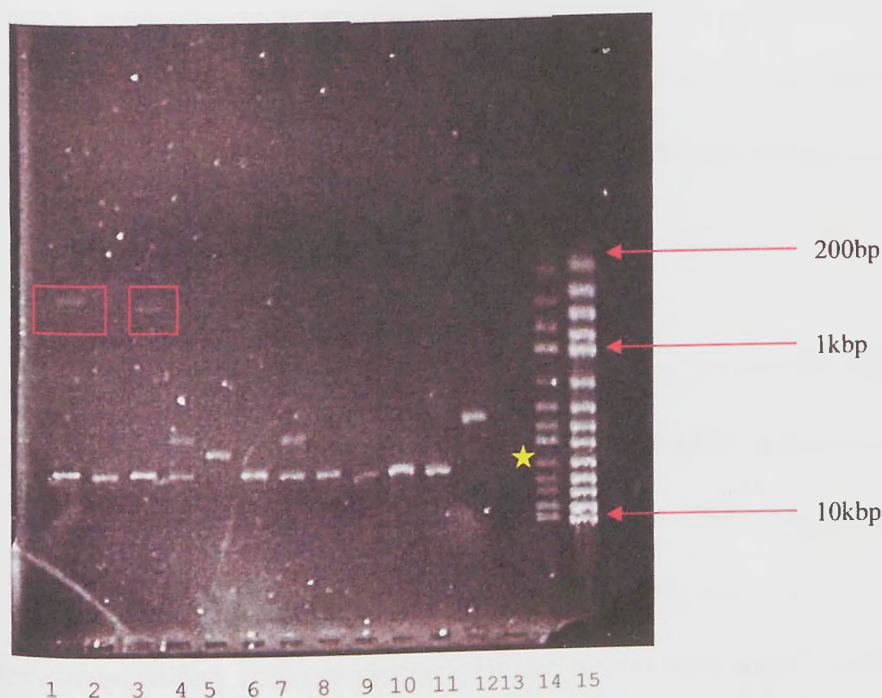
Figure 5.21. 0.8% gel showing products of restriction digest with *Eco*R1 of DNA from twelve white colonies from two agar plates of *E. coli* Topo cloned with PCR product from *M. canis* CRP primers N-terminus 1-4 and C-terminus primer 4. Lane 1)colony 1 N-terminus 1 C-terminus 2)colony 2 N-terminus 1 C-terminus 3)colony 3 N-terminus 1 C-terminus 4)colony 1 N-terminus 2 C-terminus 5)colony 2 N-terminus 2 C-terminus 6)colony 3 N-terminus 2 C-terminus 7)colony 1 N-terminus 3 C-terminus 8)colony 2 N-terminus 3 C-terminus 9)colony 3 N-terminus 3 C-terminus 10)colony 1 N-terminus 4 C-terminus 11)colony 2 N-terminus 4 C-terminus 12)colony 3 N-terminus 4 C-terminus 13)uncut plasmid 14)1kbp ladder 15) hypderladder. Red boxes show PCR product ~300bp.
Although uncut plasmid was loaded onto the gel it has not stained well and is not visible. Comparison with a previous gel (figure 5.14) indicates that the band for uncut plasmid should be at approximately 4kbp as marked on the gel with ★

5.3.8 Sequencing of PCR products using new *M. canis* primers

PCR product from colony 3.

The DNA sequences of the PCR product inserts showed only one complete sequence between the two primers (N-terminus 1 C-terminus) from DNA from colony 3. This sequence is shown below where the area highlighted in pink is that corresponding to the N-

254

terminus 1 primer, whilst that highlighted in yellow corresponds to the C-terminus primer:

5'GTCTTGGTCTTCGCCTACCATCATGCAGGTATGACCCAACCTCCGACCTACAGGCTCAACCCA
CTCTGCCAGGCCAACCATCCCCACCCTGGTGTCTGATCACCATCACTTGCCTGATGTATGAATTC
CCAGCCGCCCCACTCTCCATCGCCAAATGCCTGACTTTCCTCAACCCTCATGTCACAACCTCCAA
TCCCCACCCCCCTACATACCCCACAAATCATATAAATTTGCCTTAGAGGCTTACCTTCTCTCTGG
CCTGTGAATTTCACTGGACCAAAAAAAACATAGAATCCCTACAGTGTAGAAAGTGGCCATTTGG
CCCAATGAGTCTGCACTGACTCTCCAACAGAGCATGTTCCCCAGGCCCAAATCCCTGCCCTATT
CCCATAACCCAGTGCATTTATAATAGCTAATCCACATAAACTTCACATCTTTGGAGACTAAGGG
GCAATTTAGCATGGCCAGTCCACCTAACCCACACATCTTTGGACCGTGGGAGGAAACCTGCGCA
CTCAGAGGAAACGCACACAAACATGGGGAAAACATAAAAAATCCACACAATCATCCAAGG▒▒
AAAGTCCAACGCTTTCCC 3'

The entire nucleotide sequence between and including the two primers (N-terminus 1 and C-terminus) from DNA insert from colony 3, was used in a nucleotide BLAST search, which showed 75% identity over 38% of the query sequence to nucleotide sequence from *Ginglymostoma cirratum* (nurse shark).

The nucleotide sequence between and including the two primers was translated into all six reading frames. As the primers were designed from peptide sequence and an electron density map showing amino acid composition of dogfish CRP, then it must be assumed that any translated amino acid sequence for dogfish CRP must include the amino acid sequence corresponding to these primers. Of the amino acid sequences from all the six reading frames, only one had the translated amino acid sequence for the N-terminus and C-terminus primers used. Both protein-protein and protein-translated nucleotide BLAST searches were performed on the sequences from the other 5 reading frames but no pentraxin related protein sequences were found. The amino acid sequence determined from reading frame 1 3'-5' that include both primers is shown below with the sequences corresponding to the N-terminus primer and C-terminus primer highlighted in pink and yellow respectively.

N-terminus 1 C-terminus colony 3 PCR product translate 3'-5' frame 1:
GKALDFALG-LCGFFMFSPCLCAFPLSAQVSSHGPKMCGLGGLAMLNCPLVSKDVKF
MWISYYKCTGLWE-GRDLGLGNMLCWRVSADSLGQMATFYTVGILCFFWSSEIHRPE
RR-ASK ANLYDLWGM-GGGDWRL-HEG-GKSGIWRWRVGRLGIHTSGK-W-SDTRVG
MVGLAEWVE PVGRRLGHTCMMVGEDQD

There are 8 stop signals between the primer sequences shown as (-), which is unlikely to occur if the dogfish CRP gene shows homology to other CRP genes that show either none or a single intron. A protein-protein BLAST search using the above sequence identified only a single sequence from *Acinetobacter baumannii*, but this was not pentraxin related and did not encompass the N-terminus or C-terminus primer area. A protein-translated nucleotide BLAST search using the above sequence identified DNA from: *Ginglymostoma cirratum* (nurse shark) DNA clone and MHC class 1 antigen that showed 36-80% identity over 15-74 amino acids of the query sequence shown above; *Scyliorhinus torazame* (cloudy catshark) DNA from HER1 line and HE1 sine that showed 47-71% identity over 19-47 amino acids of the query sequence; *Danio rerio* (zebrafish) clone DNA that showed 42-65% identity to 22-49 amino acids of the query sequence; *Triakis scyllium* (leopard shark) macrophage inflammatory protein 3 that showed 42-87% identity over 18-51 amino acids of the query sequence, and interleukin 1 gene that showed 42-85% identity over 16-55 amino acids of the query sequence; *Homo sapien* (human) DNA BAC clone that showed 39-66% identity over 20-48 amino acids of the query sequence; *Sphyrna lewini* (scalloped hammerhead) gene for preproghrelin that showed 40-82% identity over 14-36 amino acids of the query sequence; *Mustelus manazo* (starspotted dogfish) DNA clone that showed 40-81% identity over 11-55 amino acids of the query sequence; *Scliorhinus canicula* (small spotted catshark) interleukin 1 gene that showed 45-81% identity over 23-37 amino acids of the query sequence; and *Polypterus bichir* clone DNA that showed 42-65% identity over 16-45 amino acids of the query sequence. Therefore, even though the sequence matches those of elasmobranchs, the whole of the protein sequence is not similar, just parts of it and similarity is not consistent with any particular protein.

To determine whether the sequence obtained from sequencing contained introns and exons coding for pentraxin, a length of sequence before the first stop codon, and that after the last stop codon after an AG (signifying the end of an intron) were joined together and

used in a protein-protein BLAST search and a protein-translated nucleotide BLAST search

(see below for sequence used highlighted in blue).

N-terminus 1 C-terminus colony 3 PCR product translated 3'-5' frame 1:
GKALDFALG-LCGFFMFSPCLCAFPLSAQVSSHGPKMCGLGGLAMLNCPLVSKDVKF
MWISYYKCTGLWE-GRDLGLGNMLCWRVSADSLGQMATFYTVGILCFFWSSEIHRPE
RR-ASK ANLYDLWGM-GGGDWRL-HEG-GKSGIWRWRVGRLGIHTSGK-W-SDTRVG
MVGLAEWVE PVGRRLGHTCMMVGEDQD

The protein-protein BLAST search gave sequences with no similarity to pentraxins,

whilst the protein-translated nucleotide BLAST search gave no sequences at all that were

similar. This was to be expected because the primers were designed to cover sequence

from residues corresponding to amino acids 15-21 and those from 146-153 therefore, as the

PCR product was around 300 base pairs which was the similar to the predicted size of the

sequence between the two primers, then there should have been no introns in the PCR

product.

PCR product from colony 1.

The sequence retrieved from DNA from colony 1 did not span the entire sequence

between the two primers as shown below where the sequence for the N-terminus primer is

shown in pink. Nucleotides that corresponded to plasmid DNA from either end of the

sequence were removed leaving the sequence shown below:

5'AAGGCGTTGGATTTTGCAAAGGCAAAGAACCCAGACAACATTCTGGTAATAGTACTGAAGAC
TTGTGTCACAAGACACAAGTCAGGAGTGTGATGGATGGGTGGCTGCAGCTCCAACAACACTCA
AGAAGCTCAACAGAAGAACAAAGCAGCCCGCTTGATTGACACCCCCATTCACCACTTTTGATAT
ACACTCCCTCCACCACCAACATACAGTGGAATCAATATATGATGCCTTCAAGATGCATAGCAGC
AACTCCCAAGGCTCCTTCAACAGTACATTCCAAACCCCTGACCTTTACCACCTAGTAAGAAGCT
TAACAACACCAGGTTAAAGTCCAACAGGTTTATTTGGTAGCAAAATCCAACGCTTTCCC-3'

This sequence was used in a nucleotide-nucleotide BLAST search which showed

sequence identity to only one gene: *Scyliorhinus canicula* interleukin-1 beta gene. There

was 93% sequence identity between the two nucleotide sequences but only covering 11%

of the query sequence.

The nucleotide sequence was translated into all six possible reading frames and only two showed amino acid sequence that corresponds to the primer. The reading frames that show the amino acid sequence that corresponds to the primer used is shown below where the sequence corresponding to the N-terminus primer is highlighted in pink, and – indicates a stop codon:

N-terminus 1 C-terminus colony 1 PCR product translated 3'-5' frame 3:
QKALDFATK-TCWTLTWCC-ASY-VVKVRGLECTVEGALGVAAMHLEGIIY-
FHCMLVVEGVYIKSGEWGCQSSGLLCSSVELLECCWSCSHPSITLLTCVL-
HKSSVLLPECCLGSLPLQNPTP

N-terminus 1 C-terminus colony 1 PCR product translated 5'-3' frame 1:
QKALDFAKAKNPDNILVIVLKTCVTRHKSGV-WMGGCSSNNTQEAQQ
KNKAARLIDTPIHHF-YTLPPPPTYSGINI-CLQDA-QQLPRLLQQYIPNP-PLPPSK
KLNNTRLKSNRFIW-QNPTLSQGRIPAHWRPLLVDPSSVPSLA-SWS-LFPV-NCYP

Protein-protein BLAST searches with these sequences showed no similar sequences, and the protein-translated nucleotide BLAST search showed similarity to elasmobranchs only for N-terminus 1 C-terminus colony 1 PCR product translated 3'-5' frame 3 sequence. The other sequence for 5'-3' frame 1 showed similarity to no proteins that were from elasmobranchs or were pentraxin. The sequences that showed some similarity to the first sequence were: *Heterodontus francisci* (horn shark) HOX genes showing 52% identity to 55 amino acids of the query sequence shown above; *Carcharhinus plumbeus* (sandbar shark) recombination activating gene showing 38% identity to 55 amino acids of the query sequence; *Ginglymostoma cirratum* (nurse shark) clone DNA showing 43% identity to 67 amino acids of the query sequence; *Sphyrna lewini* (scalloped hammerhead) gene for preproghrelin showing 57% identity to 33 amino acids of the query sequence; and human genetic DNA clones which showed 27-37% identity over 45-111 amino acids of the query sequence. There were no pentraxin-like sequences in the results of the BLAST search.

To determine whether these sequences contained introns that did not code for pentraxin, and exons that did code for pentraxin, the lengths of sequence before the stop codon and

after the last stop codon (after an AG signifying the end of an intron) (see below where

sequence used is highlighted in blue) were used in a protein-protein BLAST search and a

protein-translated nucleotide BLAST search.

N-terminus 1 C-terminus colony 1 PCR product translated 3'-5' frame 3:
GKALDFATE-TCWTLTWCC-ASY-VVKVRGLECTVEGALGVAAMHLEGIIY-
FHCMLVVEGVYIKSGEWGCQSSGLLCSSVELLECCWSCSHPSITLLTCVL-
HKSSVLLPECCLGSLPLQNPTP

N-terminus 1 C-terminus colony 1 PCR product translated 5'-3' frame 1:
GKALDFAKAKNPDNILVIVLKTCVTRHKSGV-WMGGCSSNNTQEAQQ
KNKAARLIDTPIHHF-YTLPPPPTYSGINI-CLQDA-QQLPRLLQQYIPNP-PLPPSK
KLNNTRLKSNRFIW-QNPTLSQGRIPAHWRPLLVDPSSVPSLA-SWS-LFPV-N P

For both of the sequences the protein-protein BLAST search gave no pentraxin related

proteins that were similar, and the protein-translated nucleotide BLAST search only gave a

single sequence result for the latter of the above sequences, but this was not pentraxin

related. A multiple sequence alignment of the above amino acids sequences with CRP and

SAP like protein sequences showed some homology with other sequences (see figure

5.22).

As can be seen from the multiple sequence alignment, there are some areas of homology

between sequences 1 and 2 with other pentraxin like sequences, for example the

phenylalanine residue (residue number 11) that is conserved in all sequences except the

salmon and trout pentraxin. This was part of the designed primer so it is supposed to align

with the other pentraxins. In sequence 2 there is also a cysteine (residue number 37) that

appears as conserved throughout all the pentraxin sequences, but this is not present in

sequence 1. This cysteine was part of the length of sequence after the possible intron(s)

sequence had been removed. In consideration of this, and the absence of the conserved

threonine (residue number 35) in the pentraxin sequences that appears two residues before

the cysteine, it is likely that this cysteine does not belong at this position. The

determination of placement of introns was difficult due to the number of stop codons in the

sequence and also due to the fact that the protein and mRNA sequence are unknown.

However, the amino acid sequence length before the first stop codon for sequence 2 shows slightly more sequence similarity to the other pentraxin sequences than sequence 1 does and it may be possible that this corresponds to dogfish CRP although the pentraxin conserved proline (residue number 12) and leucine (residue number 22) residues are not shown in the sequence.

```
pig_sap                 ----------HRDLTGKVFVFPRESATDHVKLITKLEKP-LQNFTLCFRAYSDLS--RGYS-----LFSYNLQGKDNELLVFKHRIGE
cattle_sap              ----------QTDLRGKVFVFPRESSTDHVTLITKLEKP-LKNLTLCLRAYSDLS--RGYS-----LFSYNIHSKDNELLVFKNGIGE
human_sap_              ----------HTDLSGKVFVFPRESVTDHVNLITPLEKP-LQNFTLCFRAYSDLS--RAYS-----LFSYNTQGRDNELLVYKERVGE
guinea_pig_sap          ----------QTDLDKKVFVFPRESSSDHVNLITKLETP-LQEFTVCLRAYSDLS--RHYS-----LFSYNTPGKDNELLIYKEKLGE
golden_hamster_sap      ----------QTDLTGKVFVFPRQSETDYVKLIPRLDKP-LQNFTVCFRAYSDLS--RPHS-----LFSYNAEYGENELLIYKERIGE
armenian_hamster_sap    ----------QTDLTGKVFVFPRESESDYVKLIPRLEKP-LENFTLCFRTYTDLS--RPHS-----LFSYNTKNKDNELLIYKERMGE
rat_sap                 ----------QTDLNQKVFVFPRESETDYVKLIPWLEKP-LQNFTLCFRAYSDLS--RSQS-----LFSYSVNSRDNELLIYKDKVGQ
mouse_sap               ----------QTDLKRKVFVFPRESETDHVKLIPHLEKP-LQNFTLCFRTYSDLS--RSQS-----LFSYSVKGRDNELLIYKEKVGE
rabbit_crp              ----------QAGMHKKAFVFPKESDNSYVSLNAQLKKP-LKAFTVCLYFYTDLSMTRGYS-----IFSYATRRQFNEILLFWSKDIG
human_crp               ----------QTDMSRKAFVFPKESDTSYVSLKAPLTKP-LKAFTVCLHFYTELSSTRGYS-----IFSYATKRQDNEILIFWSKDIG
guinea_pig_crp          ----------GTDMSKKTFVFPKETDNSYVSLKAQLKKP-LSAFTVCLHIYTELFMTRGYS-----IFSYATEKEANEILIFWSKDRG
rat_crp_                ----------HEDMSKQAFVFPGVSATAYVSLEAESKKP-LEAFTVCLYAHADVS--RSFS-----IFSYATKTSFNEILLFWTRGQG
golden_hamster_crp      ----------QKDMSKTAFVFPKESANSYVSLEAQSKKT-LKAFTVCLHIFTELSTTRSFS-----IFSYATKNSPNEILIFWSKDRG
pig_crp_                ----------QTDMIGKAFVFPKESENSYVSLTARLTKP-LTAFTVCLRVYTDLN--RDYS-----LFSYATKTQYNELLFRGKTAV
salmon_ptxn_            ----------EHQDLSGKVFVIPMATSTSHVKLHANVSEP-ISAMTMCQRFNSEQE--RGQS-----LFSLATQSHDNDLLLYKRSMGV
trout_ptxn_             ----------DLQDLSGKVFVIPMTTSTSHVKLHANVSKP-ISAMTMCQRFNSEQE--RGQS-----LFSLATQSHDNDLLLYKRSMGV
Limulus_CRP1.1          ----------LEEGEITSKVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKVNRLKGTL------HMFSYATAKKDNELLTLIDEQGD
Limulus_CRP3.3          ----------LEEGEITSKVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKVNCLKGTL------HMFSYATAKKDNELLTLLDEQGD
Limulin_fragment        ----------LEEGEITSKVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKVNQLKGTLGHSRVLBMFSYATAKKDNELLTFLDEQGD
Limulus_CRP1.4          ----------LEEGEITSKVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKVNHLKSTL------TIFSYNTAKNDNELLTSLEKQGA
Tachypleus_1            ----------------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKIHRLKASL------HMFSYATTEKDNEILTFLNQQGD
Tachypleus_8            ----------------KVKFPPSSSPSFPRLVMVGTLPDLHEITLCYWFKLHRLNGAP------HIFSYATSETDNEILTSLNENGD
Tachypleus_16           ----------------KVKFPPSTASSFPRLVMDGTLPNLHELTLCYWFKAHRLDRRL------HVFSYAQNATDNEILAFVEAGNT
Limulus_SAP             ----------AVDIRDVKISFPGTQNPKFPHLRFMQTLPAVRQLTVCQRIKPFHRNTGY-------IFSCATSNQDNQFITSMYVKSD
carp_pentraxin          CLLLSLTAAATEVGLVGKVLLFPTKTN-TSFVALTPEKPLSLSAFTLCMRVATELQGERETI-----LFAYRTQDYDELNVWREKDGRL
dogfish_sap_Robey       ------------------GFPGKSLIF-----------------------------------------------------------------
dogfish_crp_Robey       AA-----SYRATAGLAGKALDF-----------------------------------------------------------------------
seq_1                   ---------------GKALDFATKSS-----VLLPECCLGSLPLQNPTP------------------------------------------
seq_2                   ---------------GKALDFAKAKNPDNILVIVLKTCVTRHKSGVCYP------------------------------------------
human_CRP_numbering     ...........1........10........20........30........40........50........60........70......
```

Figure 5.22. Multiple sequence alignment of sequences N-terminus 1 C-terminus colony 1 PCR product translated 3'-5' frame 3 labeled "seq 1", and N-terminus C-terminus colony 1 PCR product translated 5'-3' frame 1 labeled "seq 2" both highlighted in grey. Sequences found previously for dogfish CRP and SAP by Robey *et al.*, (1983) are highlighted in grey also. The sequences are edited from the full sequences obtained by removal on sequence length that may correspond to introns. Green represents identical residues and yellow similar residues.

5.4 Discussion

In this study we attempted to determine dogfish CRP and SAP sequences by amplifying DNA using degenerate PCR primers designed from limited peptide sequence data and electron density map data from crystallized dogfish CRP and SAP. This was unsuccessful, primarily due to the degeneracy of the primers that were designed. This could not be avoided due to use of limited peptide sequences used for primer design. Therefore, during PCR, the areas of DNA that were amplified did not correspond to the dogfish CRP and SAP genes. The low melting temperatures of some of the primers used also affected annealing as the PCR programs used could not have had a high annealing temperature which would have meant more specificity for primer binding. Using pools of degenerate primers also caused problems as only one of the primers should be complimentary for the dogfish CRP or SAP gene sequence, but the other primers could be complementary for other areas of DNA sequence. PCR would amplify any regions of DNA that the primers annealed to, which seems to be the case for this work. Therefore, the primers had less specific annealing which led to annealing at areas of DNA that did not correspond to the CRP and SAP genes.

PCR would have had more success in amplifying dogfish CRP and SAP sequences from DNA if the primers were longer and less degenerate meaning that they would be able to anneal more specifically to the genomic DNA at the areas of CRP and SAP genes. From the amino acid sequences available the primers have been made as least degenerate as possible. If more and longer sequences were known, then longer and less degenerate primers could be designed. Longer peptides could be found from a repeat of the cyanogen bromide or trypsin digest or from cleavage using another chemical or enzymatic procedure that targets different amino acids. Even then, there would be no guarantee that the additional peptide sequence would help to design a less degenerate primer, as shown by the use of a trypsin digest in addition to the cyanogen bromide digest.

The sequence of the dogfish CRP and SAP could also have been found using a cDNA library from mRNA of dogfish liver similar to that done for *Limulus* CRP genes 1.1, 1.4, and 3.3 (Nguyen *et al.*, 1986b). cDNA fragments cloned in *E.coli* would be screened with multiple DNA probes of non-overlapping sites designed from the amino acid sequence of dogfish CRP or SAP. One problem would be that the probes used to hybridize to the cDNA would be designed from the amino acid sequence that the initial degenerate primers were designed from. As the primers did not work in annealing to the correct areas of DNA for dogfish CRP and SAP, they would also hybridize with the same areas in the cDNA library. The effect of this in selection of the clones with dogfish CRP and SAP DNA would be minimised by using multiple probes with non-overlapping sites. This method requires a lot of time and screening of bacterial colonies and so was inappropriate for this research.

As the primers used in this study were designed from limited amino acid sequence data, not only were they degenerate, but they were also shorter than those used in the previous study by Tharia *et al.*, (2002). As previously discussed, longer primers could be designed from longer peptide fragments. However, another possibility would be to design primers from conserved regions of nucleotide sequence throughout the CRP and SAP pentraxins. A multiple sequence alignment of some CRP and SAP nucleotide sequences from invertebrates, fish, and mammals was performed (see figure 5.23) to highlight some areas of conservation and hence potential primer design. As can be observed from the alignment, there appear to be several conserved regions (see regions 101-120; 400-426; 434-473; 509-531; 539-554; 560-601) some with nucleotide conservation throughout all the species listed. A number of primers could be designed from these regions of high conservation and used to attempt to isolate CRP and/or SAP genes from *Mustelus canis*.

In addition to new primer design either from a more detailed peptide sequence or from nucleotide sequence conservation throughout the CRP and SAP pentraxins of different species, the template pool of dogfish CRP and SAP DNA could be made larger by reverse

263

transcription PCR. Reverse transcription PCR would involve isolating RNA from the dogfish liver and converting it to cDNA using reverse transcriptase. As RNA is the product of genes being actively transcribed to produce proteins that the cell or body needs to use, it would be safe to assume that the RNA for the CRP and SAP genes were present. As the levels of CRP and SAP in dogfish are consistently high (~400μg/ml and ~100μg/ml respectively) (Robey *et al.*, 1983) then the CRP gene should be transcribed often and RNA present consistently too. Therefore, not only would the chance of primers binding be increased due to the larger template pool of DNA, but also the possibility of introns being amplified during PCR is eliminated.

Therefore, although the approach taken to isolate dogfish CRP and SAP genes used in this study was unsuccessful due to the amplification of a non-specific product during PCR, other techniques may be tried in further studies.

```
limulus_crp_1.1       --------------------------------------------------------------------------------------------
limulus_crp_3.3       --------------------------------------------------------------------------------------------
limulus_crp_1.4       --------------------------------------------------------------------------------------------
tachypleus_1_tCRP2    --------------------------------------------------------------------------------------------
tachypleus_8_tCRP1    --------------------------------------------------------------------------------------------
tachypleus_16_tCRP3   --------------------------------------------------------------------------------------------
human_crp             --------ATGGAGAAGCTG------TTGTGTTTCTTGGTCTTGAC---CAGCCTCTCTCATGCTTTTGGCCAG-------------ACAGACATGTCGA
rabbit_crp            ------------------------------------------------------------------------------------GCAGGCATGCACA
guinea_pig_crp        ------TGGCGAAGCTGCTG------TTGTATTTCCTGCTCTTAAC---CAGTCTTTCTGATGTGTTTGGCGGG-------------ACAGACATGTCTA
rat_crp               -----ATGGAGAAGCTACTA------TGGTGTCTTCTGATCACGAT---AAGCTTCTCTCAGGCTTTTGGTCAT-------------GAAGACATGTCTA
golden_hamster_crp    -----ATGGAGAAGCTGCTG------TGGTGCTCACTGGTCATGAT---CGGCTTCTCTCAGGCTTTTGCCCAG-------------AAAGACATGTCTA
golden_hamster_sap    ------TGGACAAGCTGCTGTCCCTGCTTGGGGTGTCTATACTTGCAGGCCTGCTTTTAGAAGCCTTTGCTCAG-------------ACAGACCTCACAG
armenian_hamster_sap  -----ATGGACAAGATGCTGCTGCTGCTTGGGGTATCTAT---------CCTCCTTTCAGAAGTCTTTGCTCAG----------- --ACAGACCTCACAG
rat_sap               -----ATGGACAAGCTGCTG------CTTTGGATGTCTGTCTTCACCAGCCTTCTTTCAGAAGCCTTTGCTCAG-------------ACAGACCTCAATC
mouse_sap             -----ATGGACAAGCTACTG------CTTTGGATGTTTGTCTTCACCAGCCTTCTTTCAGAAGCCTTTTGTCAG-------------ACAGACCTCAAGA
human_sap             -----ATGAACAAGCCGCTG------CTTTGGATCTCTGTCCTCAC---CAGCCTCCTGGAAGCCTTTGCTCA-------------CACAGACCTCAGTG
guinea_pig_sap        -----ATGGACAAGATGCTT------TTTTGGGTCTCTGTCTTTAC---CATCTTCCTGGATGTCTTTGCTCAGACAGGTTACCCCCCACAGACCTCGACA
pig_sap               -----------------------------------------------------------------------------------CACCGAGACCTCACGG
cattle_sap            -----ATGGAGAAGCTTCTT------CTGGGTGTCCTGCTCCTCGC---CTTTCTCCCAGAGGGCATGACACAG-----‘--------AAAGACTTGAGAG
pig_crp               ----------------------------------------------------------------------------------CAGACAGACATGATCG
salmon_pentraxin      AAAATATGGAGAGTGCACTCAATCTAATGGCGAAGCTGGTGTTTTTGCTGCCCCTGATCTATGGCTGTTATGGTGA--------ACATCAAGATCTCTCAG
trout_pentraxin       ---------------------------------------------------------------------GA--------CCTTCAAGATCTCTCAG
carp_pentraxin        --------AGCATGATGTTGGTGCCAGTGGTTTTATTTTTCTGTCTTC-TGCTCTCTCTGACAGCAGCGGCTACTG---------AAGTGGGTCTTGTTG
limulus_sap           ------------------------------ATGGTTTCTGGATACATCGTGATTTTTGTAATATTGGCCTTGACGAAGGCAGCGGTTGATATTCGGGATGTTA
                      1........10........20........30........40........50........60........70........80........90.......100
```

```
limulus_crp_1.1       --------------------------------TCGTTCCCACGACTAG-TAATGGTGGGGACGTTGCCTGATCTGCAAGAAATTACCTTATGTTACTGGTTC
limulus_crp_3.3       --------------------------------TCGTTCCCACGACTAG-TAATGGTGGGGACGTTGCCTGATCTGCAAGAAATTACCTTATGTTACTGGTTC
limulus_crp_1.4       --------------------------------TCGTTCCCGCGACTAG-TAATGGTGGGGACGTTGCCTGATCTGCAAGAAATTACCTTATGTTACTGGTTC
tachypleus_1_tCRP2    -AAGGTTAAGTTCCCGCCGTCTAGTTCTCCGTCATTCCCGCGACTAG-TAATGGTGGGACGTTACCTGATCTGCAAGAAATTACCTTATGTTACTGGTTC
tachypleus_8_tCRP1    -AAGGTAAAGTTCCCACCGTCAAGTTCTCCGTCATTCCCACGACTGG-TAATGATAGGAACGTTGCCTGATCTACACGAAATTACCTTATGTTACTGGTTC
tachypleus_16_tCRP3   -AAGGTCAAGTTCCCGCCTTCTACAGCCAGTTCCTTCCCTCGTTTGG-TGATGGATGGTACGTTACCAAATCTCCATGAATTAACCCTCTGTTACTGGTTC
human_crp            GGAAGGCTTTTGTGTTTCCCAAAGAGTCGGATACTTCCTATGTATCCCTCAAAGCACCGTTAACGAAGCCTCTCAAAGCCTTCACTGTGTGCCTCCACTTC
rabbit_crp           AGAAGGCCTTTGTGTTTCCCCAAAGAGTCAGATAATTCCTACGTGTCCCTCAACGCACAGTTAAAGAAGCCACTCAAAGCCTTCACTGTGTGCCTCTACTTC
guinea_pig_crp       AGAAGACTTTTGTGTTTCCTAAAGAGACGGATAACTCCTATGTGTCTCTGAAAGCACAGTTGAAGAAGCCACTGAGTGCCTTCACTGTGTGCCTCCATATC
rat_crp             AACAGGCCTTCGTATTTCCCGGAGTGTCAGCTACTGCCTATGTGTCCCTGGAAGCAGAGTCAAAGAAGCCACTGGAAGCCTTCACTGTGTGTCTCTATGCC
golden_hamster_crp   AAACAGCCTTCGTATTTCCCAAAGAGTCAGCCAATTCCTATGTATCCCTGGAAGCACAGTCAAAGAAGCATTGAAAGCCTTCACTGTGTGTCTCCACATC
golden_hamster_sap   GAAAGGTATTTGTGTTCCCCAGACAGTCTGAAACTGATTATGTGAAGCTGATACCACGATTGGACAAACCTCTGCAGAAACTTTACAGTGTGTTTTCGAGCC
armenian_hamster_sap GAAAGGTATTTGTGTTCCCCAGAGAGTCTGAAAGTGATTATGTGAAGCTGATCCCACGTCTGGAAAAACCACTGGAGAACTTCACACTGTGTTTTCGAACC
rat_sap             AGAAGGTATTTGTGTTCCCCAGAGAATCTGAAACTGATTATGTGAAGCTGATCCCATGGCTAGAAAAACCGCTGCAGAATTTTACACTGTGTTTCCGAGCC
mouse_sap           GGAAAGTATTTGTGTTCCCCAGAGAATCTGAAACTGATCATGTGAAGCTGATCCCACATCTAGAGAAACCTCTGCAGAATTTTACACTGTGTTTCCGAACC
human_sap           GGAAGGTGTTTGTATTTCCTAGAGAATCTGTTACTGATCGTAAACTTGATCACACCGCTGGAGAAGCCTCTACAGAACTTTACCTTGTGTTTTCGAGCC
guinea_pig_sap      AGAAGGTATTTGTGTTCCCTAGAGAATCTTCCAGTGATCACGTGAATCTGATAACAAAATTGGAGACACCTCTGCAGGAATTTACCGTGTGTTTGCGAGCT
pig_sap             GGAAGGTGTTTGTGTTCCCTAGAGAGTCTGCCACTGACCATGTGAAACTGATCACAAAGCTGGAGAAACCTCTGCAGAACTTTACCTTGTGTTTTCGAGCC
cattle_sap          GGAAGGTGTTTATTTTCCCTGAACAATCAGACACAGCCTACGTGACCCTGATCCCCAGGGTGAGGAAGCCGCTGAGGAGTTTCACTCTGTGCCTGAAAGCC
pig_crp             GAAAGGCCTTTGTCTTCCCCAAAGAGTCGGAGAACTCCTATGTGAGTCTGACTGCAAGGCTCACCAAACCACTCACGGCCTTTACTGTGTGCCTGCGTGTC
salmon_pentraxin    GTAAAGTGTTCGTAATCCCAATGGCGACAAGCACCTCACATGTAAAGCTCCATGCGAACGTCTCAGAGCCCATTTCTGCTATGACCATGTGTCAGAGGTTC
trout_pentraxin     GTAAGGTGTTCGTAATCCCAATGACGACAAGCACCTCACACGTAAAGCTCCATGCCAACGTCTCAAAGCCCATTTCTGCTATGACCATGTGTCAGAGGTTC
carp_pentraxin      GTAAAGTGCTTCTTTTTCCAACCAAGACTAACACCAGTTTTGTTGCACTCACTCCTGAAAAGCCACTGAGTCTTTCAGCTTTTACTCTCTGCATGCGTGTT
limulus_sap         AGATTTCATTTCCTGGAACCCAAAACCCTAAATTTCCCCACTTGCGATTTA---TGCAGACTTTACCCGCTGTACGTCAGCTTACAGTTTGTCAACGGATT
                    .......110.......120.......130.......140.......150.......160.......170......180.......190.......200.
```

```
limulus_crp_1.1     AAGGTTAATCGCTTA---AAGGGCACCCTTCACATGTTTTCGTACGCTACTGCTAAAAAAGACAATGAGCTTTTGACGTTAATCGACGAA------CAAGG
limulus_crp_3.3     AAGGTTAATTGCTTA---AAGGGCACCCTTCACATGTTTTCGTACGCTACTGCTAAAAAAGACAATGAGCTTTTGACGTTACTCGACGAA------CAAGG
limulus_crp_1.4     AAGGTTAATCACTTA---AAGAGCACCCTTACCATATTTTCGTACAATACAGCTAAAAACGACAATGAGCTTCTGACATCTCTCGAAAAA------CAAGG
tachypleus_1_tCRP2  AAGATTCATCGCTTA---AAAGCCTCACTTCATATGTTTTCGTACGCTACCACTGAGAAAGACAATGAGATTCTGACATTTCTAAACCAA------CAAGG
tachypleus_8_tCRP1  AAGATTCATCACTTA---AAGGCCTCACTTACCATGTTTTCATACGCTTCCACTGAAACAGACTATGAGATTGCAACATCTCTGGATGAA------CAAGG
tachypleus_16_tCRP3 AAAGCCCATCGTTTG---GATAGAAGACTCCATGTATTTTCATATGCCCAAAATGCAACAGATAATGAGATTCTAGCTTTTGTGGAAGCA------GGGAA
human_crp           TACACGGAACTGTCCTCGACCCGTGGGTACAGTATTTTCTCGTATGCCACCAAGAGACAAGACAATGAGATTCTCATATTTTGGTCTAAG------GATAT
rabbit_crp          TACACGTGATCTGTCCATGACTCGTGGGTACAGTATTTTCTCCTATGCCACCAGGAGACAATTTAACGAGATCCTCCTGTTTTGGTCCAAG------GACAT
guinea_pig_crp      TACACTGAACTGTTTATGACGCGTGGATACAGCATTTTTCTCTTATGCCACCAAAGAAGCTAATGAGATCCTCATATTTTGGTCTAAA------GATAG
rat_crp             CACGCTGATGTGAGC------CGAAGCTTCAGCATCTTCTCTTACGCTACCAAGACGAGCTTTAACGAGATTCTTCTGTTTTGGACTAGG------GGTCA
golden_hamster_crp  TTCACGGAACTGAGCACAACCCGCAGCTTCAGTATTTTTCTCTTATGCTACCAAGAACAGCCCTAACGAAATCCTCATATTTTGGTCTAAG------GATAG
golden_hamster_sap  TATAGTGACCTTTCC------CGCCCTCACAGTCTTTTCTCCTACAATGCTGAGTACGGAGAAAATGAGCTGCTAATTTATAAGGAAAGA------ATTGG
armenian_hamster_sap TATACTGACCTTTCC------CGCCCTCACAGTCTTTTCTCCTACAATACAAAGAACAAAGACAATGAGCTGCTAATTTATAAGGAAAGA------ATGGG
rat_sap             TACAGTGACCTTTCC------CGCTCTCAGAGTCTTTTCTCCTACAGTGTCAACAGCAGAGACAATGAGCTACTAATTTATAAAGACAAA------GTTGG
mouse_sap           TACAGTGACCTTTCC------CGCTCTCAGAGTCTTTTCTCCTACAGTGTCAAGGGCAGAGACAATGAGCTACTAATTTATAAAGAAAAA------GTTGG
human_sap           TATAGTGATCTCTCT------CGTGCCTACAGCCTCTTCTCCTACAATACCCAAGGCAGGGATAATGAGCTACTAGTTTATAAAGAAAGA------GTTGG
guinea_pig_sap      TATAGTGATCTTTCT------CGACATTACAGCCTCTTCTCCTACAACACCCCAGGCAAGGATAATGAGCTACTCATTTATAAAGAAAAG------CTTGG
pig_sap             TATAGTGACCTCTCT------CGTGGTTACAGCCTCTTCTCCTATAACCTCCAGGGCAAGGACAATGAGCTCCTGGTTTTTAAACATAGA------ATTGG
cattle_sap          TTCACAGACCTCACT------CGCCCCTACAGCCTCTTCTCCTACAGCACTAAATTCTAAAGACAATGAGCTGCTTCTCTTTGTCAACAAA------GTGGG
pig_crp             TATACCGACCTGAAC------CGTGACTATAGCCTCTTCTCTTACGCCACAAAGACACAGTATAACGAGATCCTCCTCTTCAGGGGAAAG------ACTGC
salmon_pentraxin    AACTCTGAGCAAGAA------CGAGGCCAGTCCCTTTTTTCTCTGGCAACCCAGTCTCATGACAATGATTTGTTGTTGTACAAACGCTCC------ATGGG
trout_pentraxin     AACTCTGAGCAAGAA------CGAGGCCAGTCCCTTTTTTCTCTGGCAACCCAGTCTCATGACAATGATTTGTTGTTGTACAAACGCTCC------ATGGG
carp_pentraxin      GCGACGGAGCTCCAGGGCGAGAGGGAGACCATTCTGTTCGCCTACCGCACACAGGACTA---TGACGAGCTCAATGTGTGGAGAGAGAAA------GATGG
limulus_sap         AAACCGTTTCACAGA------AACACAGGATACATCTTCTCCTGCGCCACATCTAATCAAGACAACCAGTTCATTACGTCTATGTATGTAAAATCTGACGG
                    ......210.......220.......230.......240.......250.......260......270.......280.......290.......300..
```

```
limulus_crp_1.1       TGATTTCCTCTTCAACGTTCATGGAGCGCCC----CAGCTGAAAGT--ACAATGCCCTAATAAGATACACATTGGAAAGTGGCATCATGTATGCCACACGT
limulus_crp_3.3       TGATTTCCTCTTCAACGTTCATGGAGCGCCC----CAGCTGAAAGT--ACAATGCCCTAATAAGATACACATTGGAAAGTGGCATCATGTATGCCACACGT
limulus_crp_1.4       TGCTTTCCACATGAACGTTCATGGAGCCCCC----CAGCTGAAAGT--ACAATGCCCTAATAAAATACACATTGGAAAGTGGCATCATGTATGTCACACGT
tachypleus_1_tCRP2    TGATTTTCTTTTCAACGTTCATGGGACTCCC----ATGCTGAAAGT--ACAATGTCCAAATAAAATACACATTGGAAGGTGGCATCATGCATGTCACACGT
tachypleus_8_tCRP1    TGATTTTTTCTTAAACGTTCATGGGAATCCC----CAGCTGAAGGT--ACAGTGCAAAAATAAAATACATATTGGAAGGTGGTATCATGTATGTCACACGT
tachypleus_16_tCRP3   TACTATAGGTTTAATTATTGGCGGTAAAACA----CAGCTACATAT--TGCATGTCCCTCTGACATAGAGGTTGGAAAATGGCATCATGTTTGCTACCTGT
human_crp             AGGATACAGTTTTACAGTG---GGTGGGTCT----GAAATATTATT--CGAGGTTCCTGAAGTCACAGTAGCT------CCAGTACACATTTGTACAAGCT
rabbit_crp            AGGATATAGTTTTTCAGTG---GGTGGGAGAT---GAAATAATATT--CAAGGTTTCTGACGTCCCTGTGGAT------CCAACTCACCTCTGTGCAAGCT
guinea_pig_crp        AGGATATATTCTTGGAGTG---GGTGGGATT----GAAATGCCTTT--CAAGGCTCCTGAAATTCCCAGTGCT------CCAGTACACATTTGTACCAGCT
rat_crp               AGGGTTTAGTATTGCAGTA---GGTGGGCCT----GAAATACTGTT--CAGTGCTTCAGAAATTCCTGAGGTA------CCAACATCTGTGCCACCT
golden_hamster_crp    AGGGTATGCTTTTGGAGTG---GGTGGGCCT----GAAGTACTATT--CAAGGCTTCTGAAATTCCTGAAGTT------CCAACACACATCTGTGCCAGCT
golden_hamster_sap    AGAATATGAACTGTACATT---GGAAATCAG----GGAACCAAAGT--CCATGGGGTGGAAGAATTTGCTTCT------CCAGTACATTTCTGTACTAGTT
armenian_hamster_sap  AGAATACGGATTGTACATT---GAAAATGTG----GGAGCCATAGT--CCGTGGGGTGGAAGAATTTGCTTCT------CCGGTGCATTTCTGCACCAGTT
rat_sap               ACAATATAGCCTATACATT---GGAAATTCA----AAAGTCACAGT--CCGTGGTTTAGAAGAATTCCCTTCT------CCAATACATTTCTGTACCAGCT
mouse_sap             AGAATACAGCCTATACATC---GGACAATCA----AAAGTCACAGT--CCGTGGTATGGAAGAATACCTTTCT------CCAGTACACCTATGTACCACTT
human_sap             AGAGTATAGTCTATACATT---GGAAGACAC----AAAGTTACATC--CAAAGTTATCGAAAAGTTCCCGGCT------CCAGTGCACATCTGTGTGAGCT
guinea_pig_sap        AGAATACAGTCTGTACATT---GGAGGGACC----AAAGTCACAGC--CAGGGTTCCCGAAGAGATTCTAGCT------CCAGTGCATATCTGTACCAGCT
pig_sap               AGAGTACAGTCTATACATT---GGAAAAACC----AAAGTTACCTT--CAAAGTTAGGGAGGTGTTCCCCAGG------CCGGTGCACATCTGTACCAGCT
cattle_sap            AGAGTACGAGTTGCACATC---GGGAACACT----AAGGTCACTTT--CAAGGTCCCCCGACCCCCTTACGGC------CCGGTCCATCTCTGTGTCAGCT
pig_crp               TGTGTACAGTATATCCGTG---GGTGGTGCC----GATGTCGTTTT--CAAGCCTCATCAGAGTTCTGAA--------CCCATGCACTTCTGTATGACGT
salmon_pentraxin      TGTGTACCGAGTGCATATC---AAGGGAGCG----TCACTGGATTT--CATCAGTTTGCCAGATTCAAAAAATGAA---TGGATCTCCATCTGCTGGACCT
trout_pentraxin       TGTGTACCGAGTGCATATC---AGGGGAGAT----GTACTGGATTT--CTTCAGTTTGCCAGATTCAAAAAATGAA---TGGATCTCCATCTGCTGGACCT
carp_pentraxin        CCGTTTGTCCTTCTACCTC---AGCGGCAGT----GGAGCGTTTTT--CAACCTGCCTGCTCTCTCCACCTTT------GGGACCCATCTGTGCCTCACCT
limulus_sap           GACTCTGAATCTTGGCCTTCAAGTAAATGCTTCTTCAAATAAATATATCTCTTGCCCTATCGAAATCGAACTAGGTCAGTGGTACCACGTGTGTCACGTAT
                      .....310.......320.......330.......340........350......360.......370.......380.......390.......400...
```

```
limulus_crp_1.1       GGTCATCATGGGAAGGTGAGGCGACTATTGCCGTGGATGGTTTCCATTGCAAAGGTAACGCAACTGGGATCGCCGTGGGACGTACGCTTAGTCAAGGTGGC
limulus_crp_3.3       GGTCATCATGGGAAGGTGAGGCGACTATTGCCGTGGATGGTTTCCATTGCAAAGGTAACGCAACTGGGATCGCCGTGGGACGTACGCTTAGTCAAGGTGGC
limulus_crp_1.4       GGTCATCATGGGAAGGTGAGGCGACTATAGGCGTGGATGGTTTCCATTGCAAAGGTAACGCAACTGGGATCGCCATGGGAGTTACGCTTAGTCAAGGAGGC
tachypleus_1_tCRP2    GGTCATCATGGAAAGGTGAGGCGACTACAGCCGTGGATGGTTTCCATTGTAAAGGTAACGCAACTGGGATCGCCATGGGAGCTACCCTTCGTCAAGGTGGC
tachypleus_8_tCRP1    GGTCATCATGGGAAGGTGAGACGACTATAGCCGTGGATGGTTTCCATTGTGAAGGTAACGCAACTGGGATCGCCACGGGAGCTACCCTTCGTCAAGGTGGC
tachypleus_16_tCRP3   GGTCTTCTTGGGAAGGAAGCTCATCTCTGATGATGAATGGTTATCGTTGCTATGGCAATACTACTGGGGTAACTAAAGGAAAAACAGTTCAAGCAGGAGGT
human_crp             GGGAGTCCGCCTCAGGGATCGTGGAGTTCTGGGTAGATGGGAAGCCCAGGGTGAGGAA-GAGTCTGAAG-----AAGGGATACACTGTGGGGGCAGAAGCA
rabbit_crp            GGGAGTCCAGCCAGAGGCATTGCAGAGCTCTGGGTAGATGGGAAGCCCATGGTGAGGAA-GAGTCTGAAG-----AAGGGCTACATTTTGGGGCCAGAGGCA
guinea_pig_crp        GGGAGTCTGTCTCAGGGATCATAGAGCTGTGGGTAGATGGGAAAGCCCAGGTGAGGAA-GAGTCTGCAG-----AAAGGGTACTTTGTGGGTACAGAGGCA
rat_crp               GGGAGTCTGCTACAGGAATTGTAGAGCTTTGGCTTGACGGGAAACCCAGGGTGCGGAA-AAGTCTGCAG-----AAGGGCTACATTGTGGGGACAAATGCA
golden_hamster_crp    GGGAGTCTGCTACAGGGATTGCAGAGTTATGGGTTGATGGGAAACCCAAGGTGCGGAA-AATTCTGCAG-----AAGGGCTACACTGTGGGGACAGATGCA
golden_hamster_sap    GGGAGTCCTCTTCTGGCATTGCTGAATTTTGGGTCAATGGGAAACCTTGGGTAAAAAA-GGGTTTGCAG-----AAGGGGTACACTGTGAAAAACAAACCC
armenian_hamster_sap  GGGAGTCCTCTTCTGGCATTGCTGATTTTTGGGTCAATGGGATACCTTGGGTAAAAAA-GGGTCTGAAG-----AAGGGGTACACTGTGAAAACCCAACCC
rat_sap               GGGAGTCCTCCTCTGGTATTGCTGAATTTTGGGTCAATGGAAAGCCTTGGGTAAAAAA-GGGTTTGCAG-----AAGGGATACACTGTGAAATCCTCACCC
mouse_sap             GGGAGTCCTCCTCTGGCATTGTTGAATTTTGGGTCAATGGAAAGCCTTGGGTAAAAAA-GTCTCTGCAG-----AGGGAATACACTGTGAAAGCCCCACCC
human_sap             GGGAGTCCTCATCAGGTATTGCTGAATTTTGGATCAATGGGACACCTTTGGTGAAAAA-GGGTCTGCGA-----CAGGGTTACTTTGTAGAAGCTCAGCCC
guinea_pig_sap        GGGAGTCCTCCTCTGGCATTGCTGAGTTTTGGATCAATGGGAAGCCTTTGGTGAAAAA-AGGTCTGAAG-----AGGGGGTACTCTGTGGCAGCTCACCCC
pig_sap               GGGAGTCTTCCACAGGCATTGCTGAGTTTTGGATCAATGGAGAGCCGCTGGTGAAAAA-GGGCCTGAGG-----CAGGGTTACTCTGTTGGAGCTTATCCC
cattle_sap            GGGAGTCTGTCTCTGGGATTGCTGAACTCTGGATGAACAGCAGGCCCGTGGGGAGGAA-GGGCTTGAGG-----AGAGGGTACACTTTGGGACAAGATGCA
pig_crp               GGGAGTCCACCTCAGGGATTACAGAGCTCTGGGTGGACGGGAAGCCCATGGTGAGGAG-AAGTCTGAAG-----AGGGGCTACTCTCTGGGGACACAGGCA
salmon_pentraxin      GGGACTCTAAAAGTGGTCTGACCCAGCTGTGGGTTAATGGGAAGCGAAGTGCACGGAG-GATTCTTAAA-----CCTGATACATCTGTAACTGGCACACCA
trout_pentraxin       GGGACTCTACAACGGGTCTGACCCAGCTGTGGGTTAATGGGAAGCGAAGTGCACGGAG-GATTCTTAAA-----CCTGATACATCTGTAACTGGCACACCA
carp_pentraxin        GGGACTCTGAGACTGGTCTTTCTGCCTTCTGGATGAATGGACATCGCAGTACATTCCA-GTTGTATAGA-----AAAGGTCACTCAATTCGTCCTGGAGGC
limulus_sap           GGTCCGGGGTGGATGGTCGTATGGCAGTTTATGCGAACGGAAGTCCGTGTGGAACAATGGAAAATGTGGGA---AAAGGCCATCAGATATCTGCAGGTGGT
                      ....410.......420.......430.......440.......450.......460.......470.......480.......490.......500....
```

```
limulus_crp_1.1       TTAGTTGTTCTTGGACAAGACCAGGACAGTGTCGGTGGTAAGTTTGATGCGACACAAAGTTTGGAAGGAGAACTGAGCGAACTTAATCTTTGGAACACGGT
limulus_crp_3.3       TTAGTTGTTCTTGGACAAGACCAGGACAGTGTCGGTGGTAAGTTTGATGCGACACAAAGTTTGGAAGGAGAACTGAGCGAACTTAATCTTTGGAACACGGT
limulus_crp_1.4       TTAGTTGTTCTTGGACAAGAGCAGGACAGTGTCGGTGGTGAGTATGATGCGGAACAAAGTTTGGAAGGAGAACTGAGCGAACTTAATCTTTGGAATACGGT
tachypleus_1_tCRP2    TTAGTTGTTCTTGGACAAGACCAGGATACTGTCGGTGGTGGGTTTGATGCAAAGCAAAGTTTGGAAGGCGAACTGAGCGAACTTAATCTTTGGGACACGGT
tachypleus_8_tCRP1    TTAGTTGTTCTTGGACAAGACCAGGATACTGTCGGTGGTGGGTTTGATGGAGATCAAAGTTGGGAAGGCGAACTGAGCGAACTTAATCTTTGGGACACGGT
tachypleus_16_tCRP3   ACTGTTGTTCTTGGACAAGACCAGGATACAGTTGGTGGTGGATTTGTTGCTTCTGAAAGTATGGAAGGGGAGCTGAGTGAGTTGAATATGTGGAACTCAGT
human_crp             AGCATCATCTTGGGGCAGGAGCAGGATTCCTTCGGTGGGAACTTTGAAGGAAGCCAGTCCCTGGTGGGAGACATTGGAAATGTGAACATGTGGGACTTTGT
rabbit_crp            AGCATTATTCTGGGGCAGGATCAGGATTCGTTTGGTGGAAGCTTTGAGAAACAACAGTCTTTGGTTGGAGACATTGGAAATGTGAACATGTGGGACTATGC
guinea_pig_crp        ATGATTATCCTTGGGCAGGATCAGGATTCATTTGGTGGGAGTTTTGATGCAAACCAGTCTTTCGTGGGAGACATTGGAGATGTGAACATGTGGGACTTTGT
rat_crp               AGCATCATCTTGGGGCAGGAGCAGGACTCGTATGGCGGTGGCTTTGACGCGAATCAGTCTTTGGTGGGAGACATTGGAGATGTGAACATGTGGGACTTTGT
golden_hamster_crp    AGCATCATCCTGGGGCAGGAGCAGGACTCATATGGTGGTGGTTTTGATGCAAACCAGTCTTTGGTAGGAGACATTGGAGATGTGAACATGTGGGACATTGT
golden_hamster_sap    AGTATTATCTTAGGACAGGAACAGGATAATTATGGAGGAGGGTTTGATAATTACCAGTCCTTCGTGGGAGAGATTGGGGATTTGAACATGTGGGACTCTGT
armenian_hamster_sap  AGTATTATCCTAGGACAGGAGCAGGATAATTATGGAGGAGGGTTTGATAAGTCCCAGTCCTTTGTAGGGGAGATGGGGGATTTGAACATGTGGGACTCTGT
rat_sap               AGTATTGTCCTGGGACAGGAGCAGGATACGTATGGAGGAGGGTTTGATAAGACACAGTCCTTTGTGGGAGAGATTGCAGATTTGTACATGTGGGACAGTGT
mouse_sap             AGTATAGTCCTGGGACAGGAGCAGGATAACTACGGAGGAGGGTTTCAAAGGTCACAGTCCTTTGTAGGAGAGTTTTCAGATTTATACATGTGGGACTATGT
human_sap             AAGATTGTCCTGGGGCAGGAACAGGATTCCTATGGGGGCAAGTTTGATAGGAGCCAGTCCTTTGTGGGAGAGATTGGGGATTTGTACATGTGGGACTCTGT
guinea_pig_sap        AAGATTATCTTGGGGCAGGAGCAGGATTCCTATGGAGGAAAGTTTGATCGAGGTCAGTCATTTTTAGGAGAGATTGGGGACGTGTACATGTGGGACTCTGT
pig_sap               AGGATTGTCCTGGGACAGGAGCAGGACAGCTATGGAGGAGGGTTTGATAAGACCCAGTCCTTTGTGGGAGAGATTGGGGATCTGTACATGTGGGGCTCTGT
cattle_sap            AGAATCATCTTGGGACAGGAGCAAGATTCATTTGGGGGAAAATTTGATGCCAAACAGTCCTTTGTTGGGGAGATCTGGGATGTGTCCTTGTGGGATCATGT
pig_crp               AGCATCATCCTGGGGCAGGAGCAAGATGCATTTGCTGGGGGCTTTGAGAAGAACCAGTGTTTGGTGGGAGACATTGGAGATGTGAACATGTGGGACTATGT
salmon_pentraxin      AGTATAATGTTAGTTCAAGAGCAAGACAGTTATGGCGGAGGTTTTGATGTCTCACAATCCTTTGTGGGGGAGGTTACTGACGTCCACTTCTGGGACAGTGT
trout_pentraxin       AGTATAATGTTAGTTCAAGAGCAAGACAGTTATGGCGGAGGTTTTGATTCCTCACAATCCTTTGTGGGGGAAGTTACTGACGTCCACTTCTGGGACAGTGT
carp_pentraxin        ACCGTCCTGCTCGGTCAAGACCCTGATAACTACCTGGGTGCCTTTGAAGTAGAGCAGAGTTTTGTAGGAGAAATTACAGATGTGCACATGTGGGATCATGT
limulus_sap           ACTGTGGTCATAGGTCAAGAACAAGACAAAATAGGTGGTGGATTCGAAGCAGGAATCTTGGAGTGGGGAGTTATCTGATCTGCAAGTCTGGGACGAAGC
                      ...510.......520.......530.......540.......550.......560.......570.......580.......590.......600.....
```

```
limulus_crp_1.1      TCTGAATCACGAACAGATTAAATATTTGAGTAAGTGTGCGCATCCTTCGGAAAGACATATCTATGGAAACATAATTCAATGGGATAAAACCCAA
limulus_crp_3.3      TCTGAATCACGAACAGATTAAATATTTGAGTAAGTGTGCGCATCCTTCGGAAAGACATATCTATGGAAACATAATTCAATGGGATAAAACACAA
limulus_crp_1.4      TCTGAATCACGAACAGATTAAACATTTGAGTAAATGTGCGCATCCTTCGGAAAGACATATCCATGGAAACATAATTCAATGGGATAAAACCCAA
tachypleus_1_tCRP2   TCTGAATCACGAACAAATTAAACACTTGAGTAAATGTGTCCATCATTCGGAACGACATATCTATGGAAACATAATTCACTGGAATAAAACACAA
tachypleus_8_tCRP1   TCTGAATCACGAACAAATTAAACACTTGAGTAAATGTGTCCATCATTCGGAACGACATATCTATGGAAACATAATTCACTGGGGTAAAACACAA
tachypleus_16_tCRP3  ACTGAACAGTAACCAGATATTACACCTAAGTAATTGTGCTGATGTATCGGAAAGACATCTGTACGGAAACATAATTCAGTGGGAAAAGACATCT
human_crp            GCTGTCAC CAGATGAGATTAACACCATCTATCTTGG----CGGGCCC-----------TTCAGTCCTAATGTCCTGAACTGGCGGGCACTGAAG
rabbit_crp           ACTTTCAC CAGAAGAGATTAATACCGTCTATGCTGG----TGGGACC-----------TTTAGTCCCAATGTCCTAGACTGGCGCGAGCTGACA
guinea_pig_crp       TCTGTCAC CTAAGGAGATTGACATGGTGTATTCCGG----TGGGACC-----------TTCAGTCCTAATGTGCTGAGCTGGCGATCGCTGACA
rat_crp              GCTATCTC CAGAACAGATCAATGCAGTCTATGTTGG----TAGGGTA-----------TTCAGCCCCAATGTTTTGAACTGGCGGGCACTGAAG
golden_hamster_crp   GCTGTCTC CAGAACAGATAAACACAGTTTGTGTTGG----TGGGACA-----------CTAGACCCCAGTGTTTTAAACTGGCAGGCACTGAAG
golden_hamster_sap   GCTGACCC CAGAAGAAATTAAATCTGTATACCAAGG----TGTCCCC-----------CTTGAGCCTAACATTCTGGATTGGCAGGCTCTGAAC
armenian_hamster_sap GCTGACCC CAGAAGAAATTAAATCTGTATATGAAGG----TTCCTGG-----------CTTGAGCCTAACATTCTAGATTGGCGGGCTTTGAAC
rat_sap              GCTGACCC CAGAGAACATTCATTCTGTGGACAGAGG----TTTCCCA-----------CCCAATCCTAATATTTTGGATTGGCGGGCCCTGAAT
mouse_sap            GCTGACCC CACAAGACATTCTATTTGTGTACAGAGA----TTCCCCT-----------GTCAATCCTAATATTTTGAATTGGCAGGCTCTTAAC
human_sap            GCTGCCCC CAGAAAATATCCTGTCTGCCTATCAGGG----TACCCCT-----------CTCCCTGCCAATATCCTGGACTGGCAGGCTCTGAAC
guinea_pig_sap       ACTGTCTC CAGATGACGTCCAGGCTGTGTATTATGG----GTCCTAT-----------GTCAATGGTAGTATCCTGAACTGGCAGGCTCTGAAC
pig_sap              GTTGTCTC CAAATGAGATCCGGCTCGTATATCAGGGTTT-GTCCTTC-----------CCCCATCCCACCATCCTGGACTGGCAGGCACTGAAC
cattle_sap           GGTCTCCT TAAAGAAC-------TTGTGCTTCA---------CCTGT-----------TACACCAGCAACATCTTGAACTGGAAGGCCCTGATT
pig_crp              GTTGTCAC CGGAGGAGATTAACACTGTCTATGCTGG----TGGGACC-----------TTCAGTCCTAATGTCCTTAACTGGAGGGCGCTGAGG
salmon_pentraxin     CATCTCTC CCTGTGAAATACA-ATTGTATATGCAGTTGAATAGATTT-----------ACTCCAGGAAATATTCTCAACTGGAAAGCATTGGAG
trout_pentraxin      CATCTCTC CCTGTGAAATACA-AATGTATATGGAGTTGAATAAATTT----------ACTGCAGGAAATATTCTCAACTGGAAAGACTTGCAG
carp_pentraxin       TCTCTCCG GCAGTCAGATTATGGCGGTTTATTCAAACCAGGAACCGTATGTG------CCGAAGGGAAATGTGTTTGACTGGAACACAATCAAA
limulus_sap          CCTAACGA CACATCAGGTGTCCACAGTTGCATCTTTGTAA-TGGCATCAGA--------CCACGTGGAAATGTCATCTCTTGGATGGAAGATTCA
                     ..610.......620.......630.......640.......650.......660.......670.......680.......690.......700
```

```
limulus_crp_1.1      TTTAAGGCTTACGACGGAGTTGTTCTTTCTCCTAACGAAATCTGTGCT------------------------
limulus_crp_3.3      TTTAAGGCTTACGACGGAGTTGTTCTTTCTCCTAACGAAATCTGTGCT------------------------
limulus_crp_1.4      TTTCAGGCTTACGACGGAGTTGTTCTTTCTCCTAACGAAATATGTGCT------------------------
tachypleus_1_tCRP2   TTTCGGGCTTATGATGGTGTCGT-------------------------------------------------
tachypleus_8_tCRP1   TTTCGGGTTTACGATGGCGTTGT-------------------------------------------------
tachypleus_16_tCRP3  TTTATGTTTTATGACGGAGTAGT -----------------------------------------------
human_crp            TATGAAGTGCAAGGCGAAGTGTTCACCAAACCCCA--GCTGTGGCCCTGA----------------------
rabbit_crp           TATCAAGTACGTGGTGAAGTACATGTCAAGCCCCA--GCTATGGCCC------------------------
guinea_pig_crp       TATGAGACACACGGTGAAGTCTTCATCAAGCCCCA--GCTGTGGCCCTGA---------------------
rat_crp              TATGAAACACACGGTGATGTGTTTATCAAGCCGCA--GCTGTGGCCCTTGACTGACTGTTGTGAGTCCTGA--
golden_hamster_crp   TATAAAGTACAGGGTGATGTGTTTATTAAGCCCCA--GCTGTGGCCCTGA---------------------
golden_hamster_sap   TATGAAATGAATGGCTATGCAGTCATCAGACCCCGA-TGTGTGGCATTAAGCTCTTACAACAAAATCTCATGA
armenian_hamster_sap TATGAAATGAGTGGCTATGCAGTTATCAGGCCTCG--TGTGTGGCATTAA---------------------
rat_sap              TATGAAATAAATGGTTATGTAGTCATCAAGCCCCG--TATGTGGGACAACAAAAGCTCATGA----------
mouse_sap            TATGAAATAAATGGCTACGTAGTCATCAGGCCCCG--TGTCTGGGATTGA--------------------
human_sap            TATGAAATCAGAGGATATGTCATCATCAAACCCTT--GGTGTGGGTCTGA--------------------
guinea_pig_sap       TATGAGTTGAATGACTATGTCATCATCAAACCTCG--AGTGTGGGACTGA--------------------
pig_sap              TACGAAATGAATGGATATGTTGTCATCAAGCCCCG--GGTGTGGAGC------------------------
cattle_sap           TATCAAGCTAAGGGCTATGTGGTGGTGAAGCCCAA--GTTGTAG--------------------------
pig_crp              TATGAAATGAGTGGCGAGGTGTATGTCAAGCCCCA--GCTGTGGCCC-----------------------
salmon_pentraxin     TTCACTGTTGAGGGAAAAGTGTTCATAGAGAGAAGTGAATTTAGAAGTGAATGTTATAATTATTAG-------
trout_pentraxin      TTCAGTATTGAGGGAAAAGTGTTCATAGAGAAAAGTGAATTTAGAAATGAATGTTATAATTAT----------
carp_pentraxin       TATGAGATTAATGGAAGTGTGTTAGTGGTGCAAGAGAGCTGA---------------------------
limulus_sap          TTTGTAGCAGACGATGGAGTGATTGTTGGAATTTCCCACATGTGTAGTTTATAA-----------------
                     ........710.......720.......730.......740.......750.......760........770.
```

Figure 5.23 Multiple sequence alignment of nucleotide sequences of CRP and SAP like proteins. Green represents identical nucleotides and yellow similar nucleotides.

Chapter 6. Phylogenetic and structural analysis of *Limulus* CRP and SAP

6.1 Introduction

CRP-like proteins have been found in horseshoe crabs *Limulus polyphemus* (Nguyen *et al.*, 1986a), *Tachypleus tridentatus* (Iwaki *et al.*, 1999) and *Carcinoscorpius rotundicauda* (Ng *et al.*, 2004). To date, 22 CRP-like proteins have been identified in *Tachypleus tridentatus* based on analysis of mRNA using reverse transcription PCR (Iwaki *et al.*, 1999). The *T. tridentatus* CRP-like proteins can be grouped into three classes tCRP1-3, based on their ligand binding, haemolytic and haemagglutinating properties as summarised in table 6.1 (Iwaki *et al.*, 1999).

Table 6.1 Properties of tCRPs as described by Iwaki *et al.*, (1999). ++ and + represent intensity of activity, whilst – represents no activity.

	PE binding	Sialic acid binding	Fetuin binding	Colominic acid binding	Haemolytic	Haemagglutinating	*E. coli* agglutination
tCRP1	++	-	-	-	-	-	-
tCRP2	+	+	+	++	+	+	+
tCRP3	-	++	++	-	++	++	-

Phylogenetic analysis of the *T. tridentatus* CRP-like protein amino acid sequences supported these groupings (see figure 6.1).

Figure 6.1 Phylogenetic tree of tCRPs 1-22. The scale bar represents a branch length of 0.05. Taken from Iwaki *et al.*, (1999).

The *C. rotundicauda* CRP-like protein sequences were cloned from genes using degenerate primers based on peptide sequences from mass spectrometry of LPS isolated CRP (Ng *et al.*, 2004). A phylogenetic tree (see figure 6.2) showed that three of the *C. rotundicauda* CRP sequences are homologous to tCRP1, and five are homologous to tCRP2. Only three CRP-like proteins have so far been identified in *L. polyphemus* (Nguyen *et al.*, 1986a), but there have been no attempts to quantify the number of CRP-like proteins in *L. polyphemus* so it is likely that more exist. The diversity of CRP-like proteins in the horseshoe crab contrasts with the presence of a single gene in humans, and might indicate the basis of a sophisticated immune response working in co-ordination with

other immune proteins. Such an example is CrOctin, a pentraxin in *C. rotundicauda* which

interacts with *Carcinoscorpius* CRP during infection and may affect CRP-LPS interaction

(Li *et al.*, 2007).

Figure 6.2. Phylogenetic tree of *C. rotundicauda* CRPs with tCRP1-3. Taken from Ng *et al.*, (2004).

The first aim of this analysis is to identify how the horseshoe crab pentraxins CRP and

SAP are phylogenetically related to CRP and SAP like proteins from other invertebrates,

amphibians, fish, and mammals. Although this has been performed previously (Hughes

1998; Tharia *et al.*, 2002), more CRP and SAP sequences have since been identified. A

phylogenetic analysis of horseshoe crab pentraxins with human CRP is also performed to

identify how the CRP and SAP like proteins from *L. polyphemus* and the CRP-like proteins

from *C. rotundicauda* and *T. tridentatus* are phylogenetically related to each other and

human CRP. An attempt is made to infer whether the proteins were homologues: i.e. have

a common ancestor, and if so whether they are orthologues: proteins which are

homologues with the same function arisen through speciation; or paralogues: proteins

which are homologues but different with perhaps related function arisen through gene

duplication within the same species.

The second aim of this analysis is to map the horseshoe crab CRP and SAP sequences onto the known structure of human CRP or *Limulus* SAP to determine areas of sequence conservation between the sequences that correspond to areas of calcium, ligand, and C1q binding for human CRP, and calcium binding for *Limulus* SAP.

6.2 Materials and Methods

6.2.1 Obtaining sequences

CRP and SAP sequences were obtained from 1) the European Bioinformatics Institute (EBI) database, which is part of the European Molecular Biology Laboratory (EMBL), and the National Centre for Biotechnology Information (NCBI) database using search terms such as pentraxin, and from 2) a Basic Local Alignment Search Tool (BLAST) search using various known pentraxin sequences as the query sequence. The EMBL-EBI covers the databases: EMBL nucleotide database, UniProt Knowledgebase (including SwissProt and TrEMBL), Macromolecular Structure Database, Array Express, EnsEMBL and IntACT. NCBI covers the databases: non-redundant GenBank CDC translations, PDB, SwissProt PIR, and PRF. BLAST was used to search for protein sequences using many known pentraxin sequences such as human CRP or *Limulus* CRP, by "WU-BLAST2-PROTEIN" in the EMBL-EBI database, and for protein and nucleic acid sequences by "protein blast" and "tblastn" in NCBI database. BLAST scans the desired database(s) for sequences which align with the query (pentraxin) sequence. The alignment of the sequence and the pentraxin sequence is given a score which is calculated from substitutions and gaps between the two sequences (Altschul *et al.*, 1997). Alignments with scores above a threshold score are shown as results on the BLAST search (Altschul *et al.*, 1997) represented by P and E values. The P value is the probability that the sequence similarity score would be obtained by a chance match between unrelated sequences of similar composition and length. The E value is the number of different alignments expected to occur by chance with the same or better score than the alignment has. Therefore, the lower the E-value, the more significant the score.

6.2.2. Translation of nucleotide sequences

Some of the pentraxin sequences obtained from the NCBI database were nucleotide sequences which were translated into an amino acid sequence using the ExPASy translation tool (http://www.expasy.ch/tools/dna.html). The nucleotide sequence was entered, and the output format chosen as compact (i.e. single letter amino acid codes and no spaces between letters) so that the output sequences could be used without further modification in sequence alignment.

6.2.3. Sequence alignment

Phylogenetic analysis of the CRP and SAP sequences was carried out using multiple sequence alignments and phylogenetic trees. Clustal is a general alignment tool that uses a variety of protein weight matrices to determine evolutionary distances between related sequences (Thompson, Higgins, Gibson 1994). Clustal works by aligning pairs of sequences and calculating their divergence shown as percentage differences in a distance matrix. The distance matrix is used to calculate a phylogenetic tree using the neighbour-joining method. This constructs a phylogenetic tree by connecting sequences with greatest similarity (i.e. neighbours) to minimize branch length (Saitou and Nei 1987; Thompson, Higgins, Gibson 1994). The Clustal program was used in both a downloaded version (ClustalX) and a web version (ClustalW: http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.py?form=clustalw-multialign) which gave identical results. In both programs, sequences were loaded into the program in FASTA format and the multiple alignment mode chosen. The output format for a complete alignment was "Clustal", which was performed using a Gonnet series protein weight matrix which was the default. The alignment file produced, was loaded into the boxshade program: http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.py?form=boxshade to allow visualisation of identical, similar and conserved residues between sequences. Identical

residues are those which have a single, fully conserved residue throughout all sequences. Similar residues within sequences have amino acids corresponding to one of the following groups: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, and FYW. Conserved residues have amino acids corresponding to one of the following groups: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, FVLIM, and HFY. The background colours chosen were: different residues as white, identical as green, similar as yellow, and conserved as red.

MUSCLE (http://www.ebi.ac.uk/muscle/) was also used to align sequences. MUSCLE is a more recently developed program than Clustal which creates alignments with the accuracy of Clustal and other alignment methods (Edgar, 2004). MUSCLE works on a similar method to Clustal using a distance matrix, but uses the unweighted pair group method with arithmetic mean (UPGMA) to create phylogenetic trees instead of neighbour-joining (Edgar 2004). The sequences loaded into the program and also the output format were chosen as FASTA. The alignment was shown in JalView, which also shows conservation, quality and consensus.

Sequences were aligned using two the programs: Clustal and MUSCLE. Clustal was mainly used, although MUSCLE was used to double check alignments and subsequent phylogenetic trees.

6.2.4 Phylogenetic trees

Once multiple alignments had been performed using Clustal and MUSCLE, trees were viewed with JalView in MUSCLE and TreeView for Clustal. Jalview in MUSCLE shows the multiple sequence alignment from which the phylogenetic tree can be viewed. Clustal produces tree files in Newick/New Hampshire format that can be interpreted by TreeView program to produce a phylogenetic tree in PHYLIP format (Page 1996). PHYLIP is a package of programs (PHYLogeny Inference Package) for inferring evolutionary trees

which   are   written   in   Newick/New   Hampshire   format (http://evolution.genetics.washington.edu/phylpi/general.html).   Bootstrap values for the phylogenetic trees were calculated using Clustal and displayed in Treeview. Bootstrap is basically a statistical analysis of inferred variability in a set of data by re-sampling the data (Felsenstein 1985). For a phylogenetic tree this means repeatedly re-sampling data that formed topology and branch lengths and giving them a confidence value. The default value used by Clustal is based upon the number of times out of 1000 re-sampling iterations, that each branch has the same length and topology to the one produced on screen. The higher the value the greater the confidence.

In Clustal, after performing a complete alignment as detailed previously, the output parameters for a bootstrapped neighbour-joining tree were set as: PHYLIP format tree with bootstrap labels on the node. The tree file created was viewed using the Treeview program. For a rooted tree, the sequence used for the root was defined as an outgroup, and the tree rooted with the outgroup.

6.2.5 Structures

The structures of human CRP and *Limulus* SAP used in this work were obtained from the Research Collabatory for Structural Bioinformatics (RCSB) protein data bank (http://www.rcsb.org/pdb/home/home.do), and from my supervisor Dr A. K. Shrive respectively. The protein data bank code for human CRP pentamer with calcium and ligand bound is: 1B09.

6.2.6 Pymol

Subunit A of human CRP and *Limulus* SAP was isolated using the Pymol program, and saved as a pdb file. Pymol was also used for the representations of the human CRP calcium, ligand, and C1q binding sites, and the representations of pentameric human CRP and hexadecameric *Limulus* SAP.

6.2.7 Protein Explorer

Protein Explorer can be used to visualise protein structure and identify features such as disulphide bonds, surface hydrophobicity, salt bridges, ligand interaction, and secondary structure (Martz 2002). Protein Explorer was also used to visualise regions of sequence conservation between the horseshoe crab CRP and SAP-like proteins and human CRP, or CRP-like proteins and *Limulus* SAP by mapping aligned sequences onto the structure of human CRP chain A, or *Limulus* SAP chain A respectively. This is often called homology mapping. The "Protein Explorer in Chime" program was used and an empty Protein Explorer program started where the pdb file of human CRP chain A or *Limulus* SAP chain A was uploaded. Advanced Explorer was used to load a multiple sequence alignment of the horseshoe crab CRP and SAP, and human CRP sequences using the "MSA3D" form. The sequence of the uploaded structure that was in the multiple sequence alignment was identified on the "MSA3D" form. The consensus percentage in advanced options on the same form was 100%, and the colours for identical, similar, and different amino acids were changed to green, yellow and red respectively. Areas of "no information" where there were gaps in the alignment or extensions of the sequence of the 3D structure beyond the ends of the alignment were coloured white.

Once the parameters were submitted and the program run, a multiple sequence alignment was provided with colours for each residue as chosen for identical, similar and different. The uploaded structure of human CRP chain A or *Limulus* SAP chain A, was coloured according to the amino acid identity: identical (green), similar (yellow) or different (red). Therefore, the areas of conservation in the multiple sequence alignment are coloured using the colour scheme mentioned which is then applied to the 3D protein structure (Martz 2002). The process using the "MSA3D" alignment form was repeated for consensus percentages from 100% to 30% in steps of 10%.

6.2.8 ConSurf

Horseshoe crab CRP and SAP sequences were also mapped onto the structure of human CRP or *Limulus* SAP subunit A using ConSurf. ConSurf was developed later than Protein Explorer and improved upon Protein Explorer by "closely approximating the evolutionary process by taking into account the phylogenetic relationships between sequences and the similarity between amino acids" (Glaser *et al.*, 2003). The basis of ConSurf is to calculate conservation scores based on evolutionary relations between protein sequences that are translated into a colouring scale used to map them onto a known 3D protein structure (Landau *et al.*, 2005). This identifies functionally important regions of proteins by identifying areas of conservation among sequence homologues (Glaser *et al.*, 2003).

ConSurf was used to identify regions of conservation between the horseshoe crab CRP and SAP, and human CRP sequences mapped onto the structure of human CRP chain A, and between the horseshoe crab CRP and *Limulus* SAP mapped onto the structure of *Limulus* SAP chain A. The ConSurf server was used online at http://consurf.tau.ac.il/. The process involved uploading the pdb file of human CRP or *Limulus* SAP chain A, from which the amino acid sequence was extracted. The program constructs a phylogenetic tree of the uploaded multiple sequence alignment of horseshoe crab CRP and SAP, using the neighbour-joining algorithm. Conservation scores from the tree were computed, which were divided into a scale of 1-9 of which grade 1 contained the most variable positions and grade 9 contained the most conserved positions. The 9 colours of the conservation grades were then projected onto the 3D structure of human CRP or *Limulus* SAP chain A. Amino acids for which there were insufficient data due to inadequate diversity in the multiple sequence alignment, were assigned conservation with a low confidence level and coloured yellow as they were areas of "no-information" (Landau *et al.*, 2005).

6.3 Sequence homology in the pentraxins CRP and SAP using multiple sequence alignments and phylogenetic trees

Multiple sequence alignments and phylogenetic trees allow visualisation of identity between sequences and can be used to deduce homology of a phylogenetic relationship. Branch length and topology on the phylogenetic tree give an indication of when the sequences on the tree evolved in relation to each other. Using multiple sequence alignments and phylogenetic trees, we can gain an insight into the homology and evolution of the pentraxins CRP and SAP between species. The multiple sequence alignments were carried out using Clustal, and the phylogenetic trees deduced from these alignments were visualised using Treeview program. MUSCLE was also used to align the pentraxin sequences which gave the same alignments as Clustal.

6.3.1 Phylogenetic relationship of CRP and SAP pentraxins

To determine how the horseshoe crab CRP and SAP-like proteins are phylogenetically related to pentraxins from other invertebrates, fish, amphibians, and mammals, phylogenetic trees were generated using a multiple sequence alignment (see figure 6.3) of sequences from organisms shown in table 6.2. The sequences from organisms in table 6.2 are composed of both whole sequences (i.e. the entire sequence for that protein), and fragments of sequences where only a part of a sequence has been determined which shows similarity to pentraxin sequences. Phylogenetic trees were generated using all sequences in table 6.2 (see figure 6.4), and just the whole sequences in table 6.2 (i.e. without sequence fragments) (see figure 6.5). Not all of the horseshoe crab pentraxins are included in either of the phylogenetic trees (figures 6.4 and 6.5), but instead a representative sequence of each of the CRP groups identified by Iwaki *et al.*, (1999) and Ng *et al.*, (2004) was used. Therefore, to cut down on the size of the phylogenetic tree, only the representative sequences from each tCRP and cCRP group were needed to visualise their placement and therefore phylogenetic relationship with other pentraxins.

Table 6.2 Organism and accession numbers from specified databases for sequences used in phylogenetic analysis. Fragments are indicated a *

Common name	Latin name	Accession number	Database
African clawed frog CRP	*Xenopus laevis*	Q07203	EMBL-EBI
Armenian hamster SAP	*Cricetulus migratorius*	P15697	EMBL-EBI
Asian horseshoe crab tCRP1 14	*Tachypleus tridentatus*	Q9U8Z4	EMBL-EBI
Asian horseshoe crab tCRP2 1	*Tachypleus tridentatus*	Q9U8Z0	EMBL-EBI
Asian horseshoe crab tCRP3 17	*Tachypleus tridentatus*	Q9U8Y1	EMBL-EBI
Atlantic horseshoe crab CRP 1.1	*Limulus polyphemus*	P06205	EMBL-EBI
Atlantic horseshoe crab CRP 1.4	*Limulus polyphemus*	P06206	EMBL-EBI
Atlantic horseshoe crab CRP 3.3	*Limulus polyphemus*	P06207	EMBL-EBI
Atlantic horseshoe crab Limulin*	*Limulus polyphemus*	P02744	EMBL-EBI
Atlantic horseshoe crab SAP	*Limulus polyphemus*	Q8WQK3	EMBL-EBI
Blue catfish CRP	*Ictalarus furcatus*	CK401961	NCBI
Budgetts frog egg jelly pentraxin	*Lepidobatrachus laevis*	Q9PTT2	EMBL-EBI
Carp pentraxin	*Cyprinus carpio*	BAB69039	NCBI
Carp pentraxin*	*Cyprinus carpio*	Cartwright *et al.*, 2003	n/a
Cattle CRP	*Bos taurus*	XP605946	NCBI
Cattle SAP	*Bos taurus*	Q3T004	EMBL-EBI
Channel catfish CRP	*Ictalarus punctatus*	CK412093	NCBI
Cherry salmon CRP	*Oncoryhnchus masou*	BP999834	NCBI
Chicken pentraxin	*Gallus gallus*	Q2EJU6	EMBL-EBI
Chimpanzee CRP	*Pan troglodytes*	XP001170785	NCBI
Cod pentraxin*	*Gadus morhua*	Lund and Olafsen 1998a	n/a
Common wolffish pentraxin*	*Anarhicas lupus*	Lund and Olafsen 1998a	n/a
Dog CRP	*Canis lupus familiaris*	XP545746	NCBI
European flounder CRP	*Platicthys flesu*	DV569956	NCBI
Golden hamster CRP	*Mesocricetus auratus*	P49262	EMBL-EBI
Golden hamster SAP	*Mesocricetus auratus*	P07629	EMBL-EBI
Grass carp CRP	*Ctenopharyngodon idella*	CK233056	NCBI
Guinea pig CRP	*Cavia porcellus*	P49254	EMBL-EBI
Guinea pig SAP	*Cavia porcellus*	P49255	EMBL-EBI
Halibut pentraxin*	*Hippoglossus hippoglossus*	Lund and Olafsen 1998a	n/a
Harbour porpoise CRP*	*Phoca vitulina*	Q4VHC7	EMBL-EBI
Horse CRP	*Equus caballus*	XP001504452	NCBI
Horse SAP	*Equus caballus*	XP001504450	NCBI
Human CRP	*Homo sapien*	P02741	EMBL-EBI
Human SAP	*Homo sapien*	P02743	EMBL-EBI
Killifish CRP	*Fundulus heteroclitus*	CN962268	NCBI
Mink SAP	*Neovison vison*	P18575	NCBI

Mosquito pentraxin	*Anopheles gambiae*	EAA06854	NCBI
Mouse CRP	*Mus musculus*	P14847	EMBL-EBI
Mouse SAP	*Mus musculus*	P12246	EMBL-EBI
Opossum SAP	*Monodelphis domestica*	XP001369849	NCBI
Pig CRP	*Sus scrofa*	O19062	EMBL-EBI
Pig SAP	*Sus scrofa*	O19063	EMBL-EBI
Plaice CRP*	*Pleuronectes platessa*	P12245	EMBL-EBI
Plaice SAP*	*Pleuronectes platessa*	P20677	EMBL-EBI
Platypus CRP	*Ornithorhynchus anatinus*	XP001517720	NCBI
Rabbit CRP	*Orytolagus cuniculus*	P02742	EMBL-EBI
Rainbow trout CRP	*Oncorhynchus mykiss*	BE669045	NCBI
Rainbow trout pentraxin*	*Oncorhynchus mykiss*	Lund and Olafsen 1998a	n/a
Rat CRP	*Rattus norvegicus*	P48199	EMBL-EBI
Rat SAP	*Rattus norvegicus*	P23680	EMBL-EBI
Rhesus monkey CRP	*Macaca mulatta*	XP001117239	NCBI
Rhesus monkey SAP	*Macaca mulatta*	XP001117234	NCBI
Salmon pentraxin*	*Salmo salar*	Lund and Olafsen 1998a	n/a
Salmon pentraxin	*Salmo salar*	CAA67765	NCBI
Seasquirt CRP a	*Halocynthia roretzi*	A4F2Q4	EMBL-EBI
Seasquirt CRP b	*Halocynthia roretzi*	A4F2Q6	EMBL-EBI
Seasquirt CRP c	*Halocynthia roretzi*	A4F2Q7	EMBL-EBI
Smooth dogfish CRP*	*Mustelus canis*	P19094	NCBI
Smooth dogfish CRP* N-terminus	*Mustelus canis*	Current work	n/a
Smooth dogfish CRP* sequence	*Mustelus canis*	Current work	n/a
Smooth dogfish SAP*	*Mustelus canis*	P19095	NCBI
Southeast Asian horseshoe crab CRP1 7	*Carcinoscorpius rotundicauda*	Q2TS37	EMBL-EBI
Southeast Asian horseshoe crab CRP2 1	*Carcinoscorpius rotundicauda*	Q2TS31	EMBL-EBI
Spiny dogfish CRP	*Squalus acanthias*	E5452672	NCBI
Trout pentraxin	*Oncorhynchus mykiss*	CAA67764	NCBI
Western clawed frog CRP	*Xenopus tropicalis*	Q6DIV9	EMBL-EBI
Western honeybee pentraxin	*Apis melifera*	XP397302	NCBI

Figure 6.3 Sequence alignment of pentraxins shown in table 6.2. The green background denotes identical residues, yellow background denotes similar residues and the white background denotes different residues. Signal peptides have been removed from the sequences where they are known.

```
human_crp                         ----------------------------------QTDMSRKAFVFPKESDTSYVSLKAPLTKPLKAFTVCLHFYTELSSTRGYSIF
chimpanzee_crp                    ----------------------------------QTDMSRKAFVFPKESDTSYVSLKAPLTKPLKAFTVCLHFYTELSSTRGYSIF
rhesus_monkey_crp                 ----------------------------------QTDMSMKAFVFPKESDNSYVTLKARLTKPLKAFTVCLHFYTELSSTRGPSVL
cattle_crp                        ----------------------------------QTDLHKKAFVFPKETENSYVSLKTQLRTPLKAFTVCLYFYTELARTHSYSIF
guinea_pig_crp                    ----------------------------------GTDMSKKTFVFPKETDNSYVSLKAQLKKPLSAFTVCLHIYTELFMTRGYSIF
pig_crp                           ----------------------------------QTDMIGKAFVFPKESENSYVSLTARLTKPLTAFTVCLRVYTDLN--RDYSLF
horse_crp                         ----------------------------------QRDMSKQAFVFPKESENSYVSLTAQLRKPLIAFTVCLQAYTDLT--RDYSLF
golden_hamster_crp                ----------------------------------QKDMSKTAFVFPKESANSYVSLEAQSKKTLKAFTVCLHIFTELSTTRSFSIF
mouse_crp                         --------------------------------HEDMFKKAFVFPKESDTSYVSLEAESKKPLNTFTVCLHFYTALSTVRSFSVF
rat_crp                           --------------------------------HEDMSKQAFVFPGVSATAYVSLEAESKKPLEAFTVCLYAHADVS--RSFSIF
rabbit_crp                        ---------------------------------AGMHKKAFVFPKESDNSYVSLNAQLKKPLKAFTVCLYFYTDLSMTRGYSIF
harbour_porpoise_crp_fragment     ----------------------------------------------------------------------------------------
dog_crp                           -----------------------------PTDLDEKAFVFPRESENSYVILFPQLQKPMKAFTVCLQVYTDLT--RPHSLF
rainbow_trout_crp_fragment        ----------------------------ERPMRRSLVFPMETDNSYVELVPQK--------------------------------
plaice_crp_fragment               ---------------------------VVIKTLVFQSESNNSFVELIPMKPLNLRAFXL---------------
pig_sap                           ------------------------------HRDLTGKVFVFPRESATDHVKLITKLEKPLQNFTLCFRAYSDLS--RGYSLF
cattle_sap                        ------------------------------QTDLRGKVFVFPRESSTDHVTLITKLEKPLKNLTLCLRAYSDLS--RGYSLF
human_sap                         ------------------------------HTDLSGKVFVFPRESVTDHVNLITPLEKPLQNFTLCFRAYSDLS--RAYSLF
rhesus_monkey_sap                 ------------------------------HTDLSGKVFVFPRESVTDHVNLITQLEKPLQNFTLCFRAYTDLS--RPYSLF
horse_sap                         ------------------------------QTYLSGKVFVFPRESSSDHVNLITKLEKPLQKFTVCFRAYSDLS--RAYSFF
guinea_pig_sap                    ------------------------------QTDLDKKVFVFPRESSSDHVNLITKLETPLQEFTVCLRAYSDLS--RHYSLF
mink_sap                          ------------------------------HTDLSGKVFVFPQESSTDHVTLTPKLEKPLKSFTLCFRAYSDLL--RAHSLF
golden_hamster_sap                ------------------------------QTDLTGKVFVFPRQSETDYVKLIPRLDKPLQNFTVCFRAYSDLS--RPHSLF
armenian_hamster_sap              ------------------------------QTDLTGKVFVFPRESESDYVKLIPRLEKPLENFTLCFRTYTDLS--RPHSLF
rat_sap                           ------------------------------QTDLNQKVFVFPRESETDYVKLIPWLEKPLQNFTLCFRAYSDLS--RSQSLF
mouse_sap                         ------------------------------QTDLKRKVFVFPRESETDHVKLIPHLEKPLQNFTLCFRTYSDLS--RSQSLF
opossum_sap                       ------------------------------QTDMSQQVMVFPKESSDAHVLLIPQEKKPLKNLTVCLRAYTELT--RAYSLF
chicken_pentraxin                 ------------------------------QEDLYRKVFVFREDPSDAYVLLQVQLERPLLNFTVCLRSYTDLT--RPHSLF
platypus_crp                      ----------------------SSTAVQEAINLGGKVFIFPRTSSDSYVTLSPQLDKALRNLTVCLRAYSDLN--RPYSLF
spiny_dogfish                     ----M-SIAAGNRMKIFIPFVLVACIYLPASVS--GGFAGKSLIFPTKTTANYVELKPTHFPNLYALTLCVRAASDEK--RSYALF
smooth_dogfish_sap_Robey_et_al    -------------------------------------GFPGKSLIF--------------------------------
little_skate                      -ETCLL-TDTVRDRMKSFVTIVLVLCIYLSGSDSAGLM-QKSVIFPTKTATSFVKLNAANFSDLTAFTVCLRAASEEN--RNYALF
smooth_dogfish_N-ter_current_w    --------------------------SPPAASSRATAGL-GKALDFP--------------------------------
smooth_dogfish_crp_Robey_et_al    --------------------------SPVAASYRATAGLAGKALDF--------------------------------
channel_catfish                   ASAWGLSLSLSLSLSLSLGGTVTVMLLLLLVLASGAGLGKVFVFPEESRSAYEQLTPM-PLNLQDPHTLYERCYTNLVGPRDHPVR
```

```
carp_crp_cartwright_                 ----------------------------------EVGLGGKVLLFPTETDTSFAELTPEKPLSLS-AFTLCMRVATELQGERDIILF
carp_ptxn                            ------------MMLVPVVLFFCLLLSLTAAATEVGLVGKVLLFPTKTNTSFVALTPEKPLSLS-AFTLCMRVATELQGERETILF
killifish                            ------------------------------EWGGS-------PGISTAMWR--SLRNVSLK-AFTLCMRVATELSGEREIILF
western_clawed_frog_crp              ------------MAIAIIGALLFIISTGLCATEGNQDLQGKSLSFEKQTFSNYAILHPKKPLNLD-AFTLCLRVSAEPPGNRDMILF
cherry_salmon                        --------------------------------------------------------TVCL-VGTGS---------
grass_carp                           ---------------------------------------------------------ILVVSSIF---------
limulus_crp3.3                       ----------------------------LEEGEITSKVKFPPSSSPSFPRLVMVG-TLPDLQEITLCYWFKVNCLKGTL----
limulus_crp1.1                       ----------------------------LEEGEITSKVKFPPSSSPSFPRLVMVG-TLPDLQEITLCYWFKVNRLKGTL----
limulin__frag_                       ----------------------------LEEGEITSKVKFPPSSSPSFPRLVMVG-TLPDLQEITLCYWFKVNQLKGTLGHSR
limulus_crp1.4                       ----------------------------LEEGEITSKVKFPPSSSPSFPRLVMVG-TLPDLQEITLCYWFKVNHLKSTL----
carcinoscorpius_crp2_1               --------------------------------KVKFPPSSSPSFPRLVMVG-TLPDLQEITLCYWFKIHRLKASL----
Tachypleus_1_tCRP2                   --------------------------------KVKFPPSSSPSFPRLVMVG-TLPDLQEITLCYWFKIHRLKASL----
carcinoscorpius_crp1_7               --------------------------------KVKFPPSSSPSFPRLVMVG-TLPDLQEITLCYWFKLHRLKGTP----
Tachypleus_14_tCRP1                  --------------------------------KVKFPPSSSPSFPRLVMVG-TLPDLHEITLCYWFKLHRLNGAP----
Tachypleus_17_tCRP3                  --------------------------------KVKFPPSTASSFPRLVMDG-TLPNLHELTLCYWFKAHRLDRRL----
limulus_sap                          ------------------------------AVDIRDVKISFPGTQNPKFPHLRFMQ-TLPAVRQLTVCQRIKPFHRNTGY----
honeybee                             ------------------------------VIHFTKSGTTDYVSMKG---PTKNLLKLTACLWLQSKNTFNYG----
mosquito                             ----------------------------LLLSAAFELEFPRTGVQAYVKLAADSSDATDLRAISMCGWYRTDDDFNYG----
seasquirt_crp_c                      -----------------------------CHQTRFYLPRKTQTSDYIRIKTALPELNDFTVCTRIKGSGTNGKFVG--
seasquirt_crp_b                      -----------------------------CHQTRFYLPRKTQTSDYFRIKTALPELHDFTVCTRIKGSGTNGKIDG--
seasquirt_crp_a                      -----------------------------CLQTRFYLPRKTQTSDYIRIKTALPELHDFTVCTRIKASGTNRKIDG--
smooth_dogfish_seq_current_wor       -----------------------------------------------------------------
budgetts_frog_jeltraxin              ----------------------------TGKTIMLFPQKTDTDYVTLKP-TERVLNQITVCLKSYTELIKEHSLFSL
african_clawed_frog_crp              --------------------------QEDLVGNVFLFPKPSVTTYAILKPEVEKPLKNLTVCLRSYTTLTRFHSLLSL
european_flounder                    ---------------------------------------------------------------M
salmon_pentraxin_fragment            ------------------AHQDLSGXVFVIPMATSTSHVXLHARVSEPISAMTM-----------------
rainbow_trout                        ------------------DLQDLSGKVFVIPMTTSTSHVKLHANVSKPISAMTMCQRFNSEQERGQSLFSL
salmon_ptxn                          ------------------EHQDLSGKVFVIPMATSTSHVKLHANVSEPISAMTMCQRFNSEQERGQSLFSL
trout_ptxn                           ------------------DLQDLSGKVFVIPMTTSTSHVKLHANVSKPISAMTMCQRFNSEQERGQSLFSL
rainbow_trout_sap_fragment           ------------------TGQDLSGKVFVSPMTSS---------VGK------------------
common_wolffish_pentraxin_frag       ------------------APQDLSGKMFIFPQETSTANV-LTARSQDF------------------
cod_pentraxin_fragment               ------------------IPQDLSGKMLTFPK-EDDDDVKLMTPK------------------
halibut_pentraxin_fragment           ------------------EPIDLLGKVFVFSKESNVDDVKLLTPQTE------------------
plaice_sap_fragment                  ------------------ZPIDLMGKVFVFDKE--------LSPBI------------------
blue_catfish                         --------TRPRDMIREMKGLVVLFSLFLLNAAERQDLSGKMFTFPVESNTHHVVLSPETNKIFFAVTVCLRAFSDISRAQSLFSL
                                     1........10........20........30........40........50........60........70........80.....
```

```
human_crp                          SYATKRQDNEILIFWSK------------DIGYSFTVGGSEILFEVPEVTVAPVHICTSWESASGIVEFWVD--GKPRVRKSLKK
chimpanzee_crp                     SYATKRQDNEILIFWSK------------GIGYSFTVGGSEILFEVPEVTVAPVHICTSWESASGIVEFWVD--GKPRVRKSLKK
rhesus_monkey_crp                  YWRALK-----------------------YEVQG--------EVFIKP--------------QLWS-------------
cattle_crp                         SYATKQQPNEILIFWSK------------GKGYIFGVHGKQVLFQSPENTQAPTHICASWESTSGIAELWVN--GKPRMRKSLEK
guinea_pig_crp                     SYATEKEANEILIFWSK------------DRGYILGVGGIEMPFKAPEIPSAPVHICSWESVSGIIELWVD--GKAQVRKSLQK
pig_crp                            SYATKTQYNEILLFRGK------------TAVYSISVGGADVVFK-PHQSSEPMHFCMTWESTSGITELWVD--GKPMVRRSLKR
horse_crp                          SYATKKQNNEILLFKGQ------------KGAYSVSVGGADVFFKLPERSPAPMHFCVTWESHSGITELWVD--GKPMVRRSLKK
golden_hamster_crp                 SYATKNSPNEILIFWSK------------DRGYAFGVGGPEVLFKASEIPEVPTHICASWESATGIAELWVD--GKPKVRKILQK
mouse_crp                          SYATKKNSNDILIFWNK------------DKQYTFGVGGAEVRFMVSEIPEAPTHICASWESATGIVEFWID--GKAKVRKSLHK
rat_crp                            SYATKTSFNEILLFWTR------------GQGFSIAVGGPEILFSASEIPEVPTHICATWESATGIVELWLD--GKPRVRKSLQK
rabbit_crp                         SYATRRQFNEILLFWSK------------DIGYSFSVGGDEIIFKVSDIPVDPTHLCASWESSTGIAELWVD--GKPMVRKSLKK
harbour_porpoise_crp_fragment      ----------------------------------------------------------WESATGIAEFWVD--GKPKVRRSLKK
dog_crp                            SYATKSQSNEILLFKER------------PGLFSVSVGGSDAFINFPQKFYAPQHFCVTWESVTGLTELWVD--GKPMVRASLRR
rainbow_trout_crp_fragment         ---------------------------------------------------------------------------------
plaice_crp_fragment                ---------------------------------------------------------------------------------
pig_sap                            SYNLQGKDNELLVFKHR------------IGEYSLYIGKTKVTFKVREVFPRPVHICTSWESSTGIAEFWIN--GEPLVKKGLRQ
cattle_sap                         SYNIHSKDNELLVFKNG------------IGEYSLYIGKTKVTVRATEKFPSPVHICTSWESSTGIAEFWIN--GKPLVKRGLKQ
human_sap                          SYNTQGRDNELLVYKER------------VGEYSLYIGRHKVTSKVIEKFPAPVHICVSWESSSGIAEFWIN--GTPLVKKGLRQ
rhesus_monkey_sap                  SYNTQGKDNELVVYKER------------VGDYSLYIGRQKVTFKVIEKLPAPVHICVSWESSSGIAECWIN--GTPLVKKGLRQ
horse_sap                          SYNIRTKDNELLIYKDK------------VGQYALHIGGTKVIAKAIEQFPSPVHICTSWESSSGIAEFWVN--GKPLVKKGLKQ
guinea_pig_sap                     SYNTPGKDNELLIYKEK------------LGEYSLYIGGTKVTARVPEEILAPVHICTSWESSSGIAEFWIN--GKPLVKKGLKR
mink_sap                           SYNAQGKDNEILVFKEK------------AGEYSLHIGKTKVTFKVAEVFPSPTHICVTWESSTGIAEFWVN--GKPLVKKGLKK
golden_hamster_sap                 SYNAEYGENELLIYKER------------IGEYELYIGNQGTKVHGVEEFASPVHFCTSWESSSGIAEFWVN--GKPWVKKGLQK
armenian_hamster_sap               SYNTKNKDNELLIYKER------------MGEYGLYIENLGAIVRGVEEFASPVHFCTSWESSSGIAEFWVN--GIPWVKKGLKK
rat_sap                            SYSVNSRDNELLIYKDK------------VGQYSLYIGNSKVTVRGLEEFPSPIHFCTSWESSSGIAEFWVN--GKPWVKKGLQK
mouse_sap                          SYSVKGRDNELLIYKEK------------VGEYSLYIGQSKVTVRGMEEYLSPVHLCTTWESSSGIVEFWVN--GKPWVKKSLQR
opossum_sap                        SYSTKGKDNELLLFKSK------------VEEYNLFIGNEKVSFRALLDFGSPVHVCASWESASGIAELWIN--GKPLVRKGLKK
chicken_pentraxin                  SYATKAQDNEILLFKPK------------PGEYRFYVGGKYVTFRVPENRGEWEHVCASWESGSGIAEFWLN--GRPWPRKGLQK
platypus_crp                       SLATQGNSNDVLLFKEAG-----------SAVFSVSIGGETVYYTVKETYPAPRQYCFTWESASGLVQLFLD--GNPLVRKGLQK
spiny_dogfish                      SYAT-SNHHNELMLWHRASGEMSLYFGR-----------HV--FSLPEMNALLKHICVTWRSLDGQTTFWVN--GERSEQKYFKR
smooth_dogfish_sap_Robey_et_al     ---------------------------------------------------------------------------------
little_skate                       SYAT-SRSNNELLIWQKANAQLDLYLGS-----------VITGFSLPKMNAWLRHICVSWESQNGEITVWFN--GGRSLGKVSSK
smooth_dogfish_N-ter_current_w     ---------------------------------------------------------------------------------
smooth_dogfish_crp_Robey_et_al     ---------------------------------------------------------------------------------
channel_catfish                    LSHSMG-PAQCVERKKWNTFSILEKQ-R-----------WCISLLSRSSLFYTMHMCGSSYSSTGYNILGVD--GRLNVLTICMH
carp_crp_cartwright_               AYRTPEVDELNVWREKDGRVSLYIQSSS-----------NATFFRLPPLSTFQTHLCITWESETGLTAFWVD--GRRSLFQLYRR
carp_ptxn                          AYRTQDYDELNVWREKDGRLSFYLSGS------------GAFFNLPALSTFGTHLCLTWDSETGLSAFWMN--GHRSTFQLYRK
killifish                          AYRTPDFDELNVWREVDGRLSFYLSGD------------GVFFQVPGLGPLQTRLCVTWDSSSGAAALFID--GRKSLTKIYRK
```

```
western_clawed_frog_crp             SYFGKGADQLNLWRERDGRFSLYLASSV-----------EGVFFSLPQISTFGTQLCVTWESSSGTTAFWVD--GKMSVRKTYRQ
cherry_salmon                       --------------------------------------------YSAYLSSTPSRPTSVSPGTRSQVWLRST-YGKRSIRKVYKP
grass_carp                          ------------------------------------------HLPPLSIFHTHMCVTWDSAPGLAAFWVD--GRRSSFQLYRK
limulus_crp3.3                      --HMFSYATAKKDNELLTLLDEQGD----FLFNVHGAPQLKVQCPNKIHIGKWHHVCHTWSSWEGEATIAVDGFHCKGNATGIAV
limulus_crp1.1                      --HMFSYATAKKDNELLTLIDEQGD----FLFNVHGAPQLKVQCPNKIHIGKWHHVCHTWSSWEGEATIAVDGFHCKGNATGIAV
limulin__fragment                   VLBMFSYATAKKDNELLTFLDEQGD----FLFNV--------------------------------------------
limulus_crp1.4                      --TIFSYNTAKNDNELLTSLEKQGA----FHMNVHGAPQLKVQCPNKIHIGKWHHVCHTWSSWEGEATIGVDGFHCKGNATGIAM
carcinoscorpius_crp2_1             --HMFSYATTGKDNEILTFINQQGD----FLFNVHGSPMLKVQCPNKIHIGRWHHACHTWSSWKGEATNVDGFHCVGNATGIAT
Tachypleus_1_tCRP2                  --HMFSYATTEKDNEILTFLNQQGD----FLFNVHGTPMLKVQCPNKIHIGRWHHACHTWSSWKGEATTAVDGFHCKGNATGIAM
carcinoscorpius_crp1_7             --HIFSYATSETDNEILTSLNEQND----FLFNIHGKTQLNVQCNNKIHAGRWHHVCHTWSSWEGEATTVVDGFRCKGNATGKAM
Tachypleus_14_tCRP1                --HIFSYATSETGNEILTSLNENGD----FLFNIHGNTQLNVQCNNKILAGRWHHVCHTWSSWEGEATIAVDGFHCKGNATEKAT
Tachypleus_17_tCRP3                --HVFSYAQNATDNEILAYVEAGNT----IGLIIGGKTQLHIACPSDIEVGKWHHVCYLWSSWEGSSSLMMNGYRCYGNTTGVTK
limulus_sap                         ---IFSCATSNQDNQFITSMYVKSDGTLNLGLQVNASSNKYISCPIEIELGQWYHVCHVWSGVDGRMAVYANGSPC-GTMENVGK
honeybee                            --TVLSYATYYYDN--AFTLIDYNG----LVIYLNGE---KIITDIKVNDGHWHFVCFTWEAENGSWNIFIDGILR-DNGIYLAK
mosquito                            --TLLSYATATVDN--AFTLTDYSG----LVLYVNGA---HVVTNVSLNDGEWHFVCVSWTSAGGRYALYVDGERS-AYGSRLSE
seasquirt_crp_c                     --TVISYAVGAKYNSLLLIGNQQNVN-------IYMNEQHSSATNIMQNGVEYHACFTRSIAPRSVKWYLNGVMKKDLTSSYSQ
seasquirt_crp_b                     --TVISYAVGAKYNRILLIADSQNVIG-------VNINDQYGQVTNIMQNGVEYHACFTRSIAPRSVKWYVNGVMKKDLTSSYSQ
seasquirt_crp_a                     --TVISPAVGAKSNRILLIADNQNVIT-------VNINDQSGQVTNIMQNGVEYHACFTRTIAPRSVKWYVNGVMKKDLTSSYSQ
smooth_dogfish_seq_current_wor      ----------------------------------------------------------------------------
budgetts_frog_jeltraxin            AMQGSGKDNTLLIYPYPPNNISISIHNE-----------DIYFKVDPEVLQWKRTCVTWDSKTGLLQLWINGKLYPRRIT--KS
african_clawed_frog_crp            ATSNPLQDNAFLLFSKPPNQCSIYINQE-----------ENVFKVDPTAVEWKHTCVSWDSVSGVVELWIDGKLYPRTVS--KK
european_flounder                   RLHAADGGTDFLSLSFPPN----------------------TWHSMCATWRSDNGLAQLWVDGVPTIKRFV--KS
salmon_pentraxin_fragment           ----------------------------------------------------------------------------
rainbow_trout                       ATQSHDN-DLLLYKRSMGVYRVHIRGD-------------------------------VWIS---SVCQIQ--KM
salmon_ptxn                         ATQSHDN-DLLLYKRSMGVYRVHIKGAS-----------LDFISLPDSKNEWISICWTWDSKSGLTQLWVNGKRSARRIL--KP
trout_ptxn                          ATQSHDN-DLLLYKRSMGVYRVHIRGDV-----------LDFFSLPDSKNEWISICWTWDSTTGLTQLWVNGKRSARRIL--KP
rainbow_trout_sap_fragment          ----------------------------------------------------------------------------
common_wolffish_pentraxin_frag      ----------------------------------------------------------------------------
cod_pentraxin_fragment              ----------------------------------------------------------------------------
halibut_pentraxin_fragment          ----------------------------------------------------------------------------
plaice_sap_fragment                 ----------------------------------------------------------------------------
blue_catfish                        SLPSTDN-GFLIYKQKQGAYSVYVLGQS-----------AEFWGLQDESNVWNSVCATWDASTGLAQLWVNGIPSSRKGI--KA
                                    ...90.......100.......110.......120.......130.......140.......150.......160......170
```

```
human_crp                        GYTVGAEASIILGQEQDSFGGNFEGSQSLVGDIGNVNMWD-------------FVLSPDEINTIYLG---------
chimpanzee_crp                   GYTVGAEASIILGQEQDSFGGNFEESQSLVGDIG--NMR----------------LTP----SILAG---------
rhesus_monkey_crp                --------------------------------------------------------------------------
cattle_crp                       GAILGTEASIILGQEQDAFAGGFDRNQCLVGDIGDVNMWD-------------FVLSPEEINAVYLG---------
guinea_pig_crp                   GYFVGTEAMIILGQDDSFGGSFDANQSFVGDIGDVNMWD------------FVLSPKEIDMVYSG---------
pig_crp                          GYSLGTQASIILGQEQDAFAGGFEKNQCLVGDIGDVNMWD-------------YVLSPEEINTVYAG---------
horse_crp                        GYNVETDASIILGQEQDSFGGGFDKGQSLVGDIGDVNMWD-------------FVLSPDEINNVYQG---------
golden_hamster_crp               GYTVGTDASIILGQEQDSYGGGFDAKQSLVGDIGDVNMWD----------IVLSPEQINTVCVG---------
mouse_crp                        GYTVGPDASIILGQEQDSYGGDFDAKQSLVGDIGDVNMWD-------------FVLSPEQINTVYVG---------
rat_crp                          GYIVGTNASIILGQEQDSYGGGFDANQSLVGDIGDVNMWD-------------FVLSPEQINAVYVG---------
rabbit_crp                       GYILGPEASIILGQDQDSFGGSFEKQQSLVGDIGNVNMWD----------YALSPEEINTIYAG---------
harbour_porpoise_crp_fragment    GYSLGSEASIVLGQEQDSFAGGFEKKQSLVGDIEDVGTC-----------------------------------
dog_crp                          GYTVGSGASIVLGQEQDSFGGGFDKNQSLVGDIEDVNMWD-------------FVLSPSQILTLYTT---------
rainbow_trout_crp_fragment       --------------------------------------------------------------------------
plaice_crp_fragment              --------------------------------------------------------------------------
pig_sap                          GYSVGAYPRIV LGQEQDSYGGGFDKTQSFVGEIGDLYMWG------------SVLSPNEIRLVYQGL--------
cattle_sap                       GYAVGAHPKIVLGQEQDSYGGGFDKNQSFMGEIGDLYMWD------------SVLSPEEILLVYQGS--------
human_sap                        GYFVEAQPKIVLGQEQDSYGGKFDRSQSFVGEIGDLYMWD------------SVLPPENILSAYQG---------
rhesus_monkey_sap                GYSVEAHPKIVLGQEQDSYGGKFDKSQSFVGEIGDLYMWN------------SVLPPEEILSAYQG---------
horse_sap                        GYSVGVNPKIVLGQEQDSCGGGFNQNQSFVGEIGDLYMWD------------SVLSPEEILSVYKG---------
guinea_pig_sap                   GYSVAAHPKIILGQEQDSYGGKFDRGQSFLGEIGDVYMWD------------SVLSPDDVQAVYYG---------
mink_sap                         GYSVATQPKIVLGQEQDSYGGGFDVNQSFVGEIGDLYMWD------------SVLSAQQILSVKHGS--------
golden_hamster_sap               GYTVKNKPSIILGQEQDNYGGGFDNYQSFVGEIGDLNMWD------------SVLTPEEIKSVYQG---------
armenian_hamster_sap             GYTVKTQPSIILGQEQDNYGGGFDKSQSFVGEMGDLNMWD------------SVLTPEEIKSVYEG---------
rat_sap                          GYTVKSSPSIVLGQEQDTYGGGFDKTQSFVGEIADLYMWD------------SVLTPENIHSVDRG---------
mouse_sap                        EYTVKAPPSIVLGQEQDNYGGGFQRSQSFVGEFSDLYMWD------------YVLTPQDILFVYRD---------
opossum_sap                      GYSLGAEPKIVLGQDQDSFGGSFDMQQSFVGEIGDIGMWD------------FVLSPHQINSLYRG---------
chicken_pentraxin                GYEVGNEAVVMLGQEQDAYGGGFDVYNSFTGEMADVHLWD------------AGLSPDKMRSAYLA---------
platypus_crp                     GYSVSRGARIILGQDQDSWGGGFKEKKSFVGEIGDVFMWD------------SVLTLDQIRAVDSG---------
spiny_dogfish                    GGCVLTGGAVILGQDQDSVRGGFKADESFVGEITDVNLWD------------HVLTASEIKTVNQGCH-------
smooth_dogfish_sap_Robey_et_al   --------------------------------------------------------------------------
little_skate                     GGVVKNGGQFFIGQEQDSVGGTFDIEQSFVGEITNVNMWD------------RVLKSNEIELISQGCY-------
smooth_dogfish_N-ter_current_w   --------------------------------------------------------------------------
smooth_dogfish_crp_Robey_et_al   --------------------------------------------------------------------------
channel_catfish                  GLNVHPNGAIIIGHDTHSTLGLFDASQSIVGVFTDVHMLDGVLRESLIERQYGKGPHALEWERALLGLS-VTISGE
carp_crp_cartwright              GHSIRPDGIVLLGQDADRFVGGFDASQSFVGEITELHMWD-------------HVLSGSQIKAVYSNQE------T
carp_ptxn                        GHSIRPGGTVLLGQDPDNYLGAFEVEQSFVGEITDVHMWD-------------HVLSGSQIMAVYSNQE------P
killifish                        GHTIRAGGSVLLGQDPDNYLGGFDINQSFVGE------------------------------------------
```

```
western_clawed_frog_crp            HHKVPGGGTVILGQDQDLQGGGFDASQSFVGEVTNVNLWD--------------YVLPCEFTSEAFALG--------
cherry_salmon                      GHKVQNGGKVILGQDPDTFLKDFESKQSFIGEITDVNMWD------------YVLSDAMIEALHVGKR-------
grass_carp                         GHSVLPGGTILLGQDADSCPGSFDAEQSFVGEITDVKMWN------------YVLSGSQIKAVYSNQE-------P
limulus_crp3.3                     GRTLSQGGLVVLGQDDSVGGKFDATQSLEGELSELNLWN------------TVLNHEQIKYLSKCAH----PSE
limulus_crp1.1                     GRTLSQGGLVVLGQDQDSVGGKFDATQSLEGELSELNLWN-----------TVLNHEQIKYLSKCAH----PSE
limulin__fragment                  ---------------------------------------------------------------------------
limulus_crp1.4                     GVTLSQGGLVVLGQEQDSVGGEYDAEQSLEGELSELNLWN-----------TVLNHEQIKHLSKCAH----PSE
carcinoscorpius_crp2_1             GATLRQGGLVVLGQDQDTVGGGFGANQSLEGELSELNLWD-----------AVLNHEQIKHLSKCVH----HSE
Tachypleus_1_tCRP2                 GATLRQGGLVVLGQDQDTVGGGFDAKQSLEGELSELNLWD-----------TVLNHEQIKHLSKCVH----HSE
carcinoscorpius_crp1_7             GVTFRQGGLVVLGQDQDSVGGGFDAKQSLVGELSELNLWD-----------MVLNHEQIKHLSECVH----PSE
Tachypleus_14_tCRP1                GVTFRQGGLVVIGQDQDSVGGGFDEKQSLVGELSELNLWD-----------MVLNHEQIKHLSKCAH----PSE
Tachypleus_17_tCRP3               GKTVQAGGTVVLGQDQDTVGGGFVASESMEGELSELNMWN------------SVLNSNQILHLSNCAD----VSE
limulus_sap                        GHQISAGGTVVIGQEQDKIGGGFEEQESWSGELSDLQVWD-----------EALTTHQVSTVASCNG----IRP
honeybee                           ETSIEGNGTFVIGQEQDRIGGGFSELESFLGKLTLLDIWS-----------TVLAAKDIKYLLN--------TC
mosquito                           GAPIQSGGLYVFGQEQDVLGGGFSETESYRGRMAYVDLWK-----------RELELEEVGRLYR--------TC
seasquirt_crp_c                    YEPIPAGGVFILGQDQDALAGGFVSSEAFAGEIQNPTVWS-----------RALSDTEIRILSDSKC----SCI
seasquirt_crp_b                    YDPIPAGGVFILGQEQDGLAGGFDSTQAFAGEIQNPTVWS-----------RALSDTEIRILSDSKC----SCI
seasquirt_crp_a                    YEPIPAGGVFILGQDQDASAGGFDSTQAFAGEFQNPTVWS-----------RALSDTEISILSDSKC----SCI
smooth_dogfish_seq_current_wor     ---------MMVGEDSLGGSFEHS---------------------------YVLPPENIKAYFSDDY-------
budgetts_frog_jeltraxin            RSPIGPQISVILGQEQDSYGGSFDINQAFVGEMSDVNVWD-----------YVLTPDHIQKVLFANM-------
african_clawed_frog_crp            ASSIGFPSSIIQGQEQDSFGGGFNIDQSFVGEISDVHMWD-----------YVIPDSEIKRYMKDRC-------
european_flounder                  GQPIRGKPSTILGQEQDSYGGGFDADQSFIGMIGKVHMWN----------------------------------
salmon_pentraxin_fragment          ---------------------------------------------------------------------------
rainbow_trout                      NGS-----------------------------------------------------------------------
salmon_ptxn                        DTSVTGTPSIMLVQEQDSYGGGFDVSQSFVGEVTDVHFWD------------SVISPCEIQLYMQLNR-------
trout_ptxn                         DTSVTGTPSIMLVQEQDSYGGGFDSSQSFVGEVTDVHFWD------------SVISPCEIQMYMELNK-------
rainbow_trout_sap_fragment         ---------------------------------------------------------------------------
common_wolffish_pentraxin_frag     ---------------------------------------------------------------------------
cod_pentraxin_fragment             ---------------------------------------------------------------------------
halibut_pentraxin_fragment         ---------------------------------------------------------------------------
plaice_sap_fragment                ---------------------------------------------------------------------------
blue_catfish                       GSSLTGAPKIILGQEQDSYGGKFDTVQSFVGMLTDVHMWD------------FVLSPHQITYYTYGGM-------
                                   .........180.......190.......200.......210.......220.......230.......240.....
```

```
human_crp                          GPFSPNVLNWRALKYEVQGEVFTKPQLWP-------------------------------------------
chimpanzee_crp                     ----PSVLMS------------------------------------------------------------
rhesus_monkey_crp                  ----------------------------------------------------------------------
cattle_crp                         STLSPNVLNWQALNYEAKGEVFIKPQLW------------------------------------------
guinea_pig_crp                     GTFSPNVLSWRSLTYETHGEVFIKPQLWP-----------------------------------------
pig_crp                            GTFSPNVLNWRALRYEMSGEVYVKPQLWP-----------------------------------------
horse_crp                          SIFSPNVIDWQELKYKAYGEVFIKPQLWS-----------------------------------------
golden_hamster_crp                 GTLDPSVLNWQALKYKVQGDVFIKPQLWP-----------------------------------------
mouse_crp                          GTLSPNVLNWRALNYKAQGDVFIKPQLWS-----------------------------------------
rat_crp                            RVFSPNVLNWRALKYETHGDVFIKPQLWPLTDCCES----------------------------------
rabbit_crp                         GTFSPNVLDWRELTYQVRGEVHVKPQLWP-----------------------------------------
harbour_porpoise_crp_fragment      ----------------------------------------------------------------------
dog_crp                            RALSPNVLNWRNLRYETRGEVFLKKELWS-----------------------------------------
rainbow_trout_crp_fragment         ----------------------------------------------------------------------
plaice_crp_fragment                ----------------------------------------------------------------------
pig_sap                            SFPHPTILDWQALNYEMNGYVVIKPRVWS-----------------------------------------
cattle_sap                         SSISPTILDWQALKYEIKGYVIVKPMVWG-----------------------------------------
human_sap                          TPLPANILDWQALNYEIRGYVIIKPLVWV-----------------------------------------
rhesus_monkey_sap                  TPVPANILDWRALNYEIKGYVIIKPLVWV-----------------------------------------
horse_sap                          SPLNPNILDWRALSYEVEGYVVIKPLV-------------------------------------------
guinea_pig_sap                     SYVNGSILNWQALNYELNDYVIIKPRVWD-----------------------------------------
mink_sap                           YLLDPNVLDWSALDYKSQGYVVIKPLVWD-----------------------------------------
golden_hamster_sap                 VPLEPNILDWQALNYEMNGYAVIRPRCVALSSYNKIS---------------------------------
armenian_hamster_sap               SWLEANILDWRTLNYEMSGYAVIRPRCVALSSYNKIS---------------------------------
rat_sap                            FPPNPNILDWRALNYEINGYVVIKPRMWDNKSS-------------------------------------
mouse_sap                          SPVNPNILNWQALNYEINGYVVIRPRVWD-----------------------------------------
opossum_sap                        ETLNANILNWKALDYEIKGYVVIKPQVWS-----------------------------------------
chicken_pentraxin                  LRLPPAPLAWGRLRYEAKGDVVVKPRLREALGA-------------------------------------
platypus_crp                       GITSPNFLNWYEMTYEEKGSVVTAPHMWY-----------------------------------------
spiny_dogfish                      -SPGGNIIDWANIAYRKAGNLIIKTIRTAQFKIF------------------------------------
smooth_dogfish_sap_Robey_et_al     ----------------------------------------------------------------------
little_skate                       -NIGGNLIDWGSTTFTQGGNVIIKDNNDCTL-M FRLQKAFKC---------------------------
smooth_dogfish_N-ter_current_w     ----------------------------------------------------------------------
smooth_dogfish_crp_Robey_et_al     ----------------------------------------------------------------------
channel_catfish                    CCL-GPMQETPCKSH-HPADEFCSPINCPTQLTI-Y-LSVLTHSMFRD-NLKCILEKRSPDFEYVNKSXTQYLENPRTWGV
carp_crp_cartwright                CVS-GNVFDWNTINYEITGNVLVTQD--------------------------------------------
carp_ptxn                          YVPKGNVFDWNTIKYEINGSVLVVQES-------------------------------------------
killifish                          ----------------------------------------------------------------------
```

```
western_clawed_frog_crp            ----GNVINWGEVQCKIVGEVKMYESGGKYIEEQIVE-------------------------------------
cherry_salmon                      -MPRGNVFDWXTIELIVNGNVKVITEEL-GVPRHQQHALHPGVLHCXPDS--------------------------
grass_carp                         YVPKGNVFDWSTIKYETRGNVLVVENN--S-SVVTSHTLSSTQLCQSL-SVHLF-SVTKLYGVNK-IFLSIKKKK------
limulus_crp3.3                     RHIYGNIIQWDKTQFKAYDGVVLSPNEICA---------------------------------------------
limulus_crp1.1                     RHIYGNIIQWDKTQFKAYDGVVLSPNEICA---------------------------------------------
limulin__fragment                  -----------------------------------------------------------------------------
limulus_crp1.4                     RHIHGNIIQWDKTQFQAYDGVVLSPNEICA---------------------------------------------
carcinoscorpius_crp2_1             RHIYGNIIRWDKTQFRAYDGVVLSPNEICA---------------------------------------------
Tachypleus_1_tCRP2                 RHIYGNIIHWNKTQFRAYDG-------------------------------------------------------
carcinoscorpius_crp1_7             RHIYGNVIHWDKTQFQAYDGVVLSPNEICA---------------------------------------------
Tachypleus_14_tCRP1               RHIYGNIIHWDKTQFRVYDG-------------------------------------------------------
Tachypleus_17_tCRP3               RHLYGNIIQWEKTSFMFYDG-------------------------------------------------------
limulus_sap                        R---GNVISWMEDSFVADDGVIVGISHMCSL--------------------------------------------
honeybee                           KKYHGDIIAWAQVQQHIHGDVTILNSPFC---------------------------------------------
mosquito                           APYSGDLVRWIDLRLQTVGLVRIVASSFC---------------------------------------------
seasquirt_crp_c                    PNAVVLKATPDVVQLEGGAALLQPDDCF----------------------------------------------
seasquirt_crp_b                    PNAVVLTATPDVVQLEGGAALLLPDDCF----------------------------------------------
seasquirt_crp_a                    PDAVVLEATPDAVQLEGAAALLLPDDCF----------------------------------------------
smooth_dogfish_seq_current_wor     -----------------------------------------------------------------------------
budgetts_frog_jeltraxin           -TLDGNFYSWDGGNYTINGLIVVLRNQFIP-------------KL-----------------------------
african_clawed_frog_crp            -DFNGNIISWRSLQYELRGQATTQPKRQCK-------------TLEHHYGLFA----KCYK----------------
european_flounder                  -FTPGNVFNWRSLDYVISGQVYEDDPHSV-HGVK--MNR-SAHLLQHPRSCE-FDL-RLLKSFWW---------------
salmon_pentraxin_fragment          -----------------------------------------------------------------------------
rainbow_trout                      -----------------------------------------------------------------------------
salmon_ptxn                        -FTPGNILNWKALEFTVEGKVFIERSEFRSECYNY---------------------------------------
trout_ptxn                         -FTAGNILNWKDLQFSIEGKVFIEKSEFRNECYNY---------------------------------------
rainbow_trout_sap_fragment         -----------------------------------------------------------------------------
common_wolffish_pentraxin_frag     -----------------------------------------------------------------------------
cod_pentraxin_fragment             -----------------------------------------------------------------------------
halibut_pentraxin_fragment         -----------------------------------------------------------------------------
plaice_sap_fragment                -----------------------------------------------------------------------------
blue_catfish                       -IQPGNVLKWNSLEFSRNGYVVIENKETSRKAGNG-TLEQLRLDVVQTRCFL-ILYIVRLYFFFPPYIVFCMIFFPLENGF
                                   ..250.......260.......270.......280.......290.......300.......310.......320......
```

Figure 6.3 Sequence alignment of pentraxins shown in table 6.2. The green background denotes identical residues, yellow background denotes similar residues and the white background denotes different residues. Signal peptides have been removed from the sequences where they are known.

Figure 6.4. Unrooted phylogenetic tree of CRP and SAP-like whole sequences and sequence fragments from various organisms. Trichotomy is the point where the ancestor would be located in a rooted tree. The scale represents the number of amino acid substitutions per site. Bootstrap numbers are displayed on the nodes of the branches
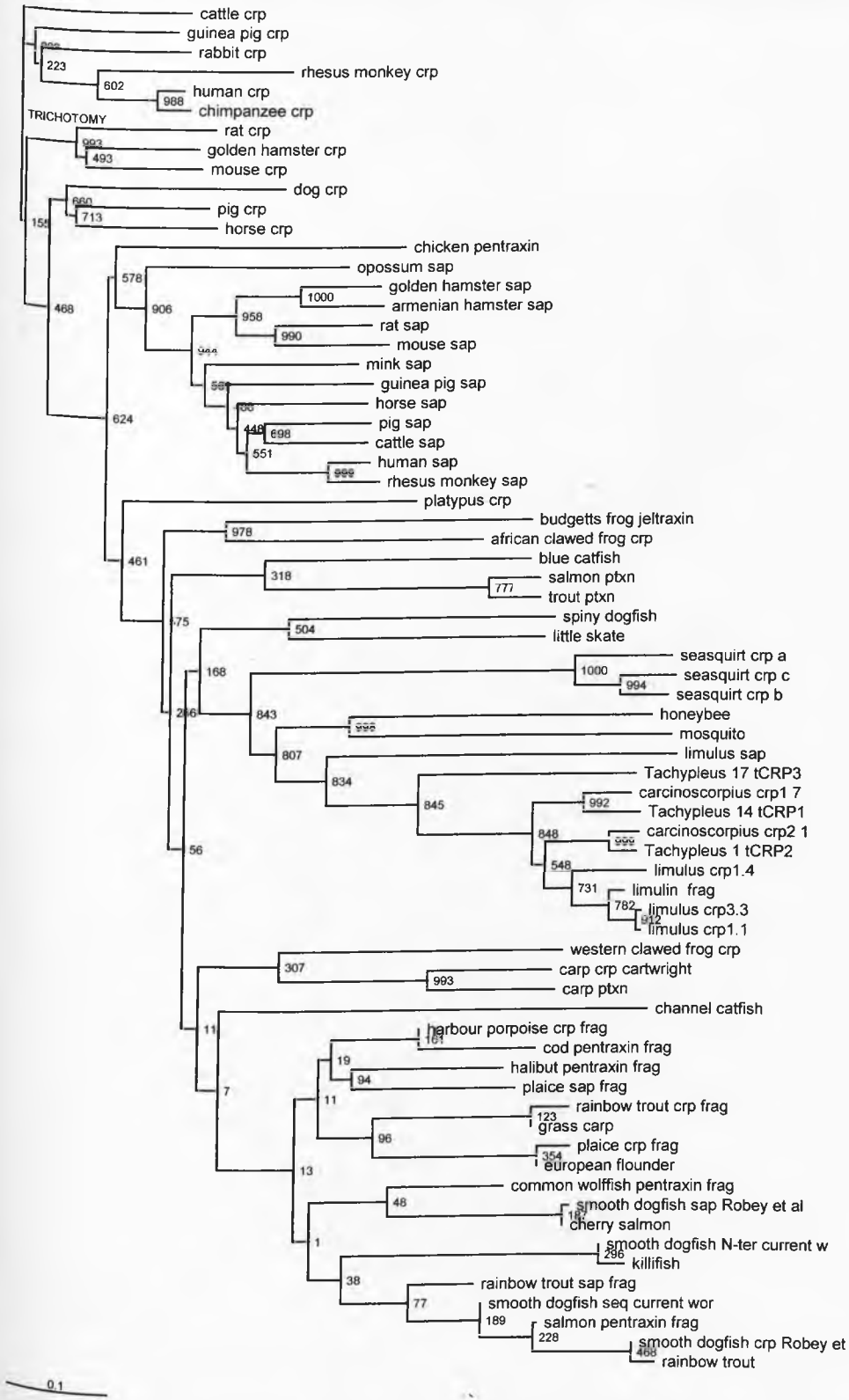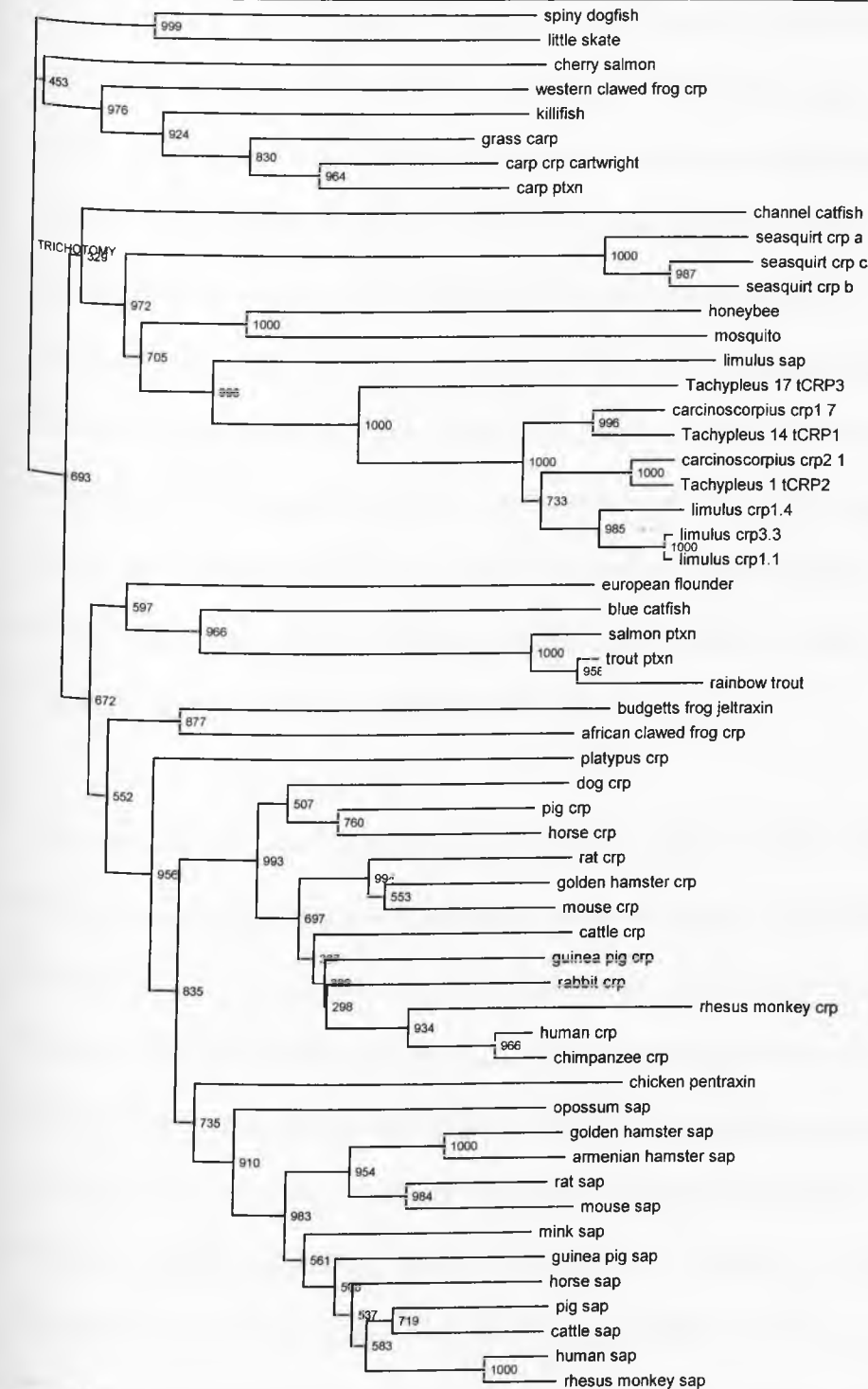
Figure 6.5 Unrooted phylogenetic tree of CRP and SAP-like whole sequences only from various organisms. Trichotomy is the point where the ancestor would be located in a rooted tree. The scale represents the number of amino acid substitutions per site. Bootstrap numbers are displayed on the nodes of the branches.

The phylogenetic trees in figures 6.4 and 6.5 show clades of mammalian CRP, mammalian SAP, fish pentraxins, and invertebrate pentraxins similar to the clades on the phylogenetic tree shown by Tharia *et al.*, (2002). The presence of the short fragments does not appear to affect the phylogenetic tree (figure 6.4) according to clade formation for mammal, fish, and invertebrate sequences with the exception of the harbour porpoise that is grouped with fish. The amphibian pentraxins do not form a clade, but are instead branched individually (Budgetts frog jeltraxin and *Xenopus laevis* CRP), or with fish (*Xenopus tropicalis* CRP), suggesting a higher sequence identity with pentraxins from other species than within the amphibians. The harbour porpoise CRP is positioned in the fish clade of the phylogenetic tree possibly because it is a fragment which shows greater identity with the pentraxin sequences of the fish than to mammals or invertebrates.

The majority of fish pentraxins are positioned within the same clade when both the whole sequences and sequence fragments are used to create a phylogenetic tree (see figure 6.4), however, a few sequences: spiny dogfish pentraxin, little skate pentraxin, blue catfish pentraxin, salmon pentraxin and the trout pentraxin, are positioned near the invertebrate clade. This is possibly due to the large number of fish pentraxin sequences that are short fragments which therefore affect the alignment. On the phylogenetic tree of pentraxin sequences without fragments (figure 6.5), the spiny dogfish pentraxin, little skate pentraxin, blue catfish pentraxin, trout pentraxin, and salmon pentraxin all cluster with the fish not the invertebrates, although the fish pentraxins split into two clades. The channel catfish pentraxin sequence however, branches with the invertebrates. The presence of fragments lowers the reliability of the phylogenetic tree as indicated by the very low bootstrap values on figure 6.4 between the fish branches which contain most of the sequence fragments. When the fragments are removed, the branching relationships between the fish sequences becomes more reliable as shown by the much higher bootstrap values on figure 6.5.

The mammalian pentraxin sequences form a clade that splits into CRP pentraxins and SAP pentraxins. Within the CRP and SAP clades are smaller species related clusters, for example the rodent SAP are clustered together as are the primate SAP, indicating that these proteins have similar structure and function and therefore are perhaps orthologs. The bootstrap values that indicate reliability of the tree branching is high between mammalian branches on both the phylogenetic tree with sequence fragments (figure 6.4) and the one without (figure 6.5), although they are generally higher on the latter suggesting a higher reliability of phylogenetic analysis without sequence fragments. The platypus CRP is on its own branch within the mammalian clade, and so is less similar to any mammalian CRP or SAP than other mammalian sequences, but not dissimilar enough that it is grouped with the non-mammals. The chicken pentraxin shows higher identity to mammalian SAP rather than CRP, indicating that although this pentraxin is not from a mammal, it exhibits sequence similarity with mammalian pentraxins, more specifically SAP.

All the invertebrate pentraxins group together and appear to form a clade. Seasquirt CRP sequences are clustered together, as are the insect pentraxins and the horseshoe crab pentraxins, showing species related branching similar to the mammalian pentraxins. Overall, the horseshoe crab pentraxins CRP and SAP show most identity with other invertebrate pentraxins than with those from fish, amphibians, or mammals. It is clear that *Limulus* SAP belongs with the invertebrates as this was not clear in a previous study by Tharia *et al.*, (2002) as there were no other invertebrate pentraxin sequences in that alignment except for those of the *Limulus* pentraxins. The bootstrap value of branching of invertebrate pentraxins from fish pentraxins on the phylogenetic tree that included fragments (figure 6.4) was 56, whilst in the phylogenetic tree that did not include fragments (figure 6.5) the value was higher at 693, indicating that the presence of fragments lowered the reliability of the branching.

In conclusion, invertebrate, fish, and mammalian CRP and SAP-like proteins form their own clades on a phylogenetic tree indicating conservation of sequence and possibly function within the clades. As CRP and SAP proteins form separate clusters on the mammalian clade, it is possible that these two proteins are paralogues (i.e. have different but related functions, but arose from a common ancestor). It is possible that invertebrate, fish, and mammalian CRP and SAP may be orthologues (i.e. have the same function but in different organisms) although not enough detail is known about the function of these proteins to draw any firm conclusions.

Phylogenetic trees with all sequences and with only whole sequences were rooted using a number of outgroups such as human CRP, human SAP, *Limulus* CRP, and *Limulus* SAP. Rooting the tree did not appear to affect its cladistics to a great extent and did not appear to alter the phylogenetic relationship of the invertebrate CRP and SAP-like proteins to those of fish and mammals.

6.3.2 Horseshoe crab CRP and SAP

To determine how all the horseshoe crab CRP and SAP-like proteins are phylogenetically related to each other and human CRP, rather than just sequences representative of *Tachypleus* and *Carcinoscorpius* CRP groups as shown in figures 6.4 and 6.5, the sequences shown in table 6.3 were used to generate a phylogenetic tree (see figure 6.6). The phylogenetic tree was rooted with human CRP (see figure 6.7) and *Limulus* SAP (see figure 6.8) to determine the relationship between the horseshoe crab CRP and SAP-like proteins, and the horseshoe crab CRPs to these proteins respectively. From the phylogenetic trees it can be observed which protein sequences are homologues and which are likely to be orthologues (homology through speciation) or paralogues (homology through gene duplication).

Table 6.3 Table showing the accession numbers of the horseshoe crab CRP and SAP sequences used for

phylogenetic analysis.

Common name	Latin name	Protein	Accession number	Database
Atlantic horseshoe crab	*Limulus polyphemus*	CRP 1.1	P06205	EMBL-EBI
Atlantic horseshoe crab	*Limulus polyphemus*	CRP 1.4	P06206	EMBL-EBI
Atlantic horseshoe crab	*Limulus polyphemus*	CRP 3.3	P06297	EMBL-EBI
Atlantic horseshoe crab	*Limulus polyphemus*	SAP	Q8WQK3	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP2 1	Q9U8Z0	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP2 2	Q9U8Z2	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP2 3	Q9U8Z1	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP2 4	Q9U8Y9	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP2 5	Q9U8Y8	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP2 6	Q9U8Y7	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP2 7	Q9U8Z9	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP1 8	Q9U8Y6	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP1 9	Q9U8900	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP1 10	Q9U8Z8	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP1 11	Q9U8Z7	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP1 12	Q9U8Z5	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP1 13	Q9U8Z6	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP1 14	Q9U8Z4	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP1 15	Q9U8Z3	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP3 16	Q9U8X9	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP3 17	Q9U8Y1	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP3 18	Q9U8Y0	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP3 19	Q9U8Y5	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP3 20	Q9U8Y4	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP3 21	Q9U8Y3	EMBL-EBI
Asian horseshoe crab	*Tachypleus tridentatus*	CRP3 22	Q9U8Y2	EMBL-EBI
Southeast Asian horseshoe crab	*Carcinoscorpius rotundicauda*	CRP1 7	Q2TS37	EMBL-EBI
Southeast Asian horseshoe crab	*Carcinoscorpius rotundicauda*	CRP1 8	Q2TS38	EMBL-EBI
Southeast Asian horseshoe crab	*Carcinoscorpius rotundicauda*	CRP1 9	Q2TS39	EMBL-EBI
Southeast Asian horseshoe crab	*Carcinoscorpius rotundicauda*	CRP2 1	Q2TS31	EMBL-EBI
Southeast Asian horseshoe crab	*Carcinoscorpius rotundicauda*	CRP2 2	Q2TS32	EMBL-EBI
Southeast Asian horseshoe crab	*Carcinoscorpius rotundicauda*	CRP2 3	Q2TS33	EMBL-EBI
Southeast Asian horseshoe crab	*Carcinoscorpius rotundicauda*	CRP2 4	Q2TS34	EMBL-EBI
Southeast Asian horseshoe crab	*Carcinoscorpius rotundicauda*	CRP2 6	Q2TS36	EMBL-EBI
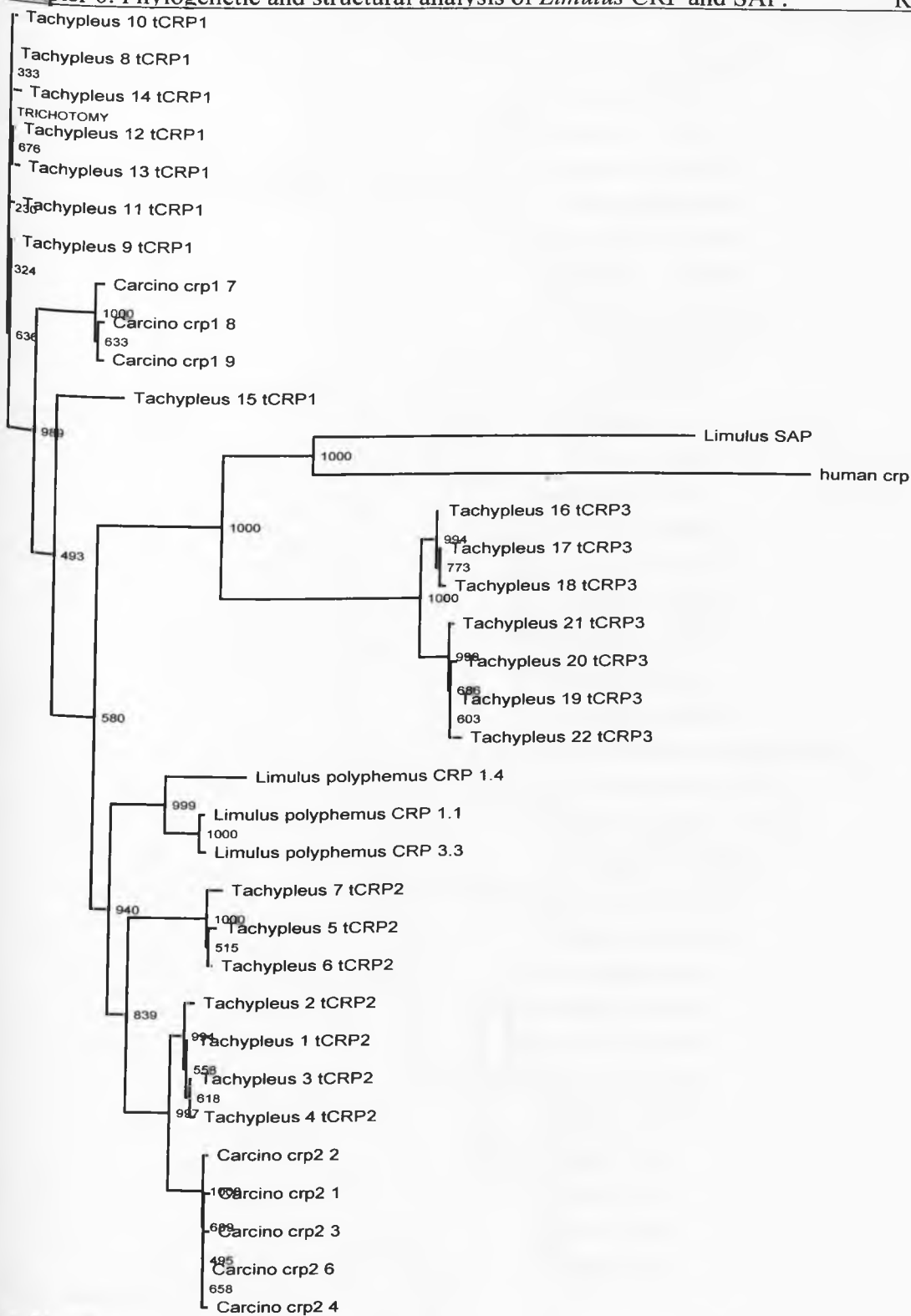
Figure 6.6 Unrooted phylogenetic tree for horseshoe crab CRP and SAP, and human CRP. Trichotomy is the point where an ancestor would be located in a rooted tree. Bootstrap numbers are displayed on the nodes of branches. The scale represents the number of amino acid substitutions per site.
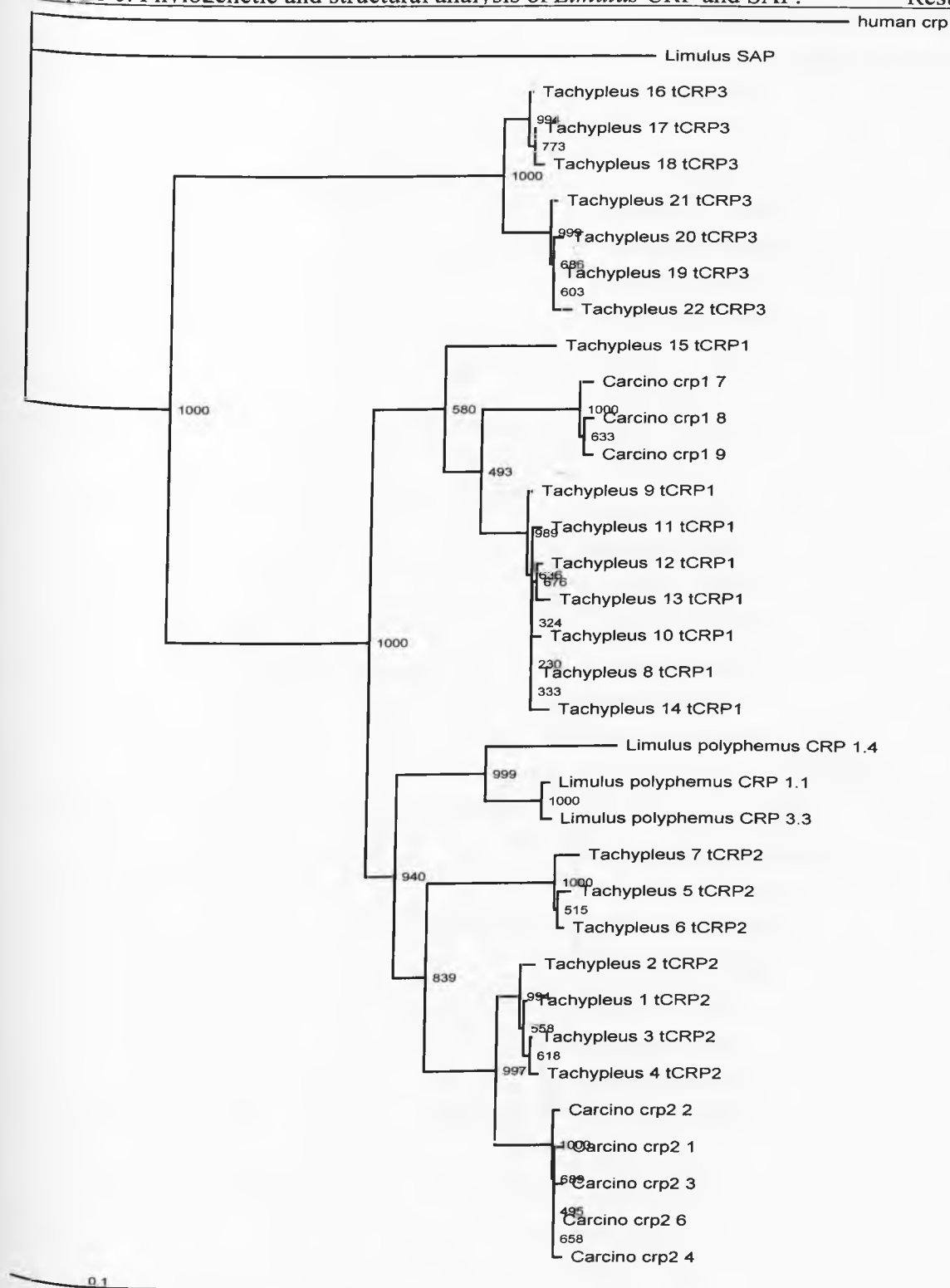
Figure 6.7 Phylogenetic tree for horseshoe crab CRP and SAP rooted with human CRP. Bootstrap numbers are displayed on the nodes of branches. The scale represents the number of amino acid substitutions per site.
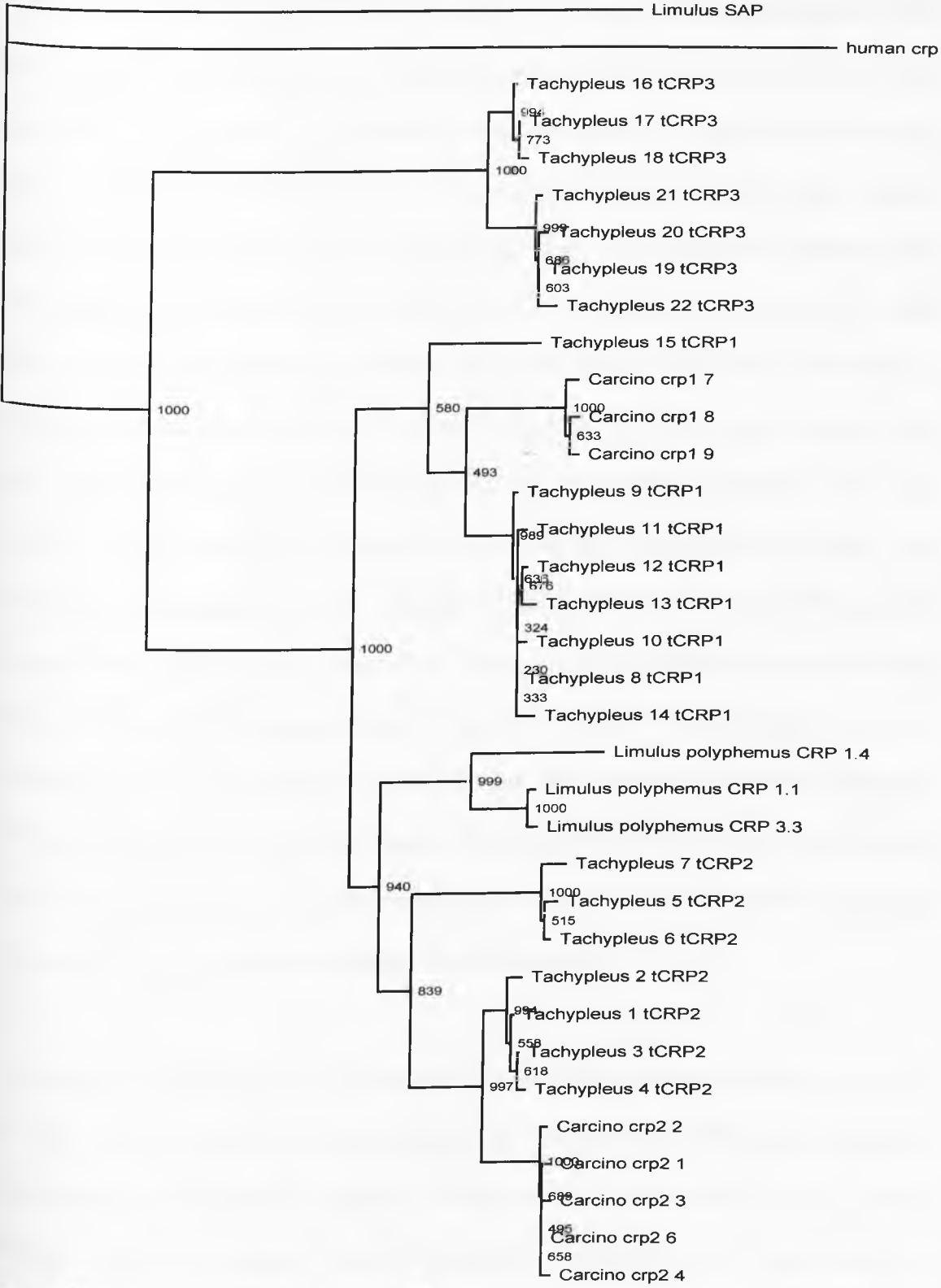
Figure 6.8 Phylogenetic tree for horseshoe crab CRP and human CRP, rooted with *Limulus* SAP. Bootstrap numbers are shown on the nodes of branches. The scale represents the number of amino acid substitutions per site.

On the unrooted tree (figure 6.6) *Limulus* SAP and human CRP branch separately from the horseshoe crab CRP sequences indicating that they have low sequence identity with them. The bootstrap value is 1000 showing complete reliability in the branching. Rooting the tree with human CRP (figure 6.7) or *Limulus* SAP (figure 6.8) leaves the respective other sequence on its own branch indicating that these two sequences are different from each other and are branched together due to their dissimilarities with the horseshoe crab CRP sequences. The dissimilarity between human CRP and *Limulus* SAP was shown in phylogenetic analysis (see figures 6.4 and 6.5) which showed clustering of *Limulus* SAP with invertebrate pentraxins, and human CRP with mammalian pentraxins. The long branch lengths for *Limulus* SAP and human CRP indicate a large number of amino acid substitutions per site indicating low sequence identity. Therefore, *Limulus* SAP and CRP sequences have less sequence conservation than *Limulus* CRP does with *Tachypleus* and *Carcinoscorpius* CRP, suggesting that *Limulus* CRP and SAP are paralogues (i.e. their homology has arisen through gene duplication, but they perhaps have different function). Even though there are differences between horseshoe crab CRP sequences and those of *Limulus* SAP and human CRP, similarities between them allow for comparison of possible structure and function which is explained later in the chapter.

Clustering of horseshoe crab CRP sequences corresponds to that observed by Iwaki *et al.*, (1999) for *Tachypleus* which corresponded to differences in CRP binding properties within the clusters of CRP like- proteins, and also with that observed by Ng *et al.*, (2004) between CRP-like proteins from *Tachypleus tridentatus* and *Carcinoscorpius rotundicauda*. The tCRP1 sequences are situated close to the cluster of *Carcinoscorpius* cCRP 1 sequences, and the tCRP2 sequences are clustered with the *Carcinoscorpius* cCRP 2 sequences. This indicates that these proteins are likely to be related and therefore may be orthologues, i.e. homologous proteins in different species with similar function. The bootstrap values between the clusters of tCRP2-cCRP2 and tCRP1-cCRP1 are high indicating reliability in the tree branching. This was to be expected because the cCRP1 and

cCRP2 groups show the greatest similarity to tCRP1 and tCRP2 groups in a previous study (Ng *et al.*, 2004), hence their similar names. The tCRP3 sequences are not clustered with any other horseshoe crab CRP sequences which may be because the tCRP3 sequences contain a unique hydrophobic sequence that the other *Tachypleus* CRPs do not possess (Iwaki *et al.*, 1999). This sequence is: Asn72-Thr-Ile-Gly-Leu-Ile-Ile-Gly-Gly80, and is completely conserved in the tCRP3 group suggesting that it is important to the haemolytic and haemagglutinating activity of the tCRP3 group (Iwaki *et al.*, 1999).

A majority of the branch lengths between CRP sequences within the same clade are short, indicating a small number of amino acid substitutions per site and that divergence of proteins from each other is relatively recent compared to the divergence of sequences between clades. The branch length of *Limulus* CRP 1.4 is noticeably longer than those of *Limulus* CRP 1.1 and 3.3 suggesting that *Limulus* CRP 1.4 evolved before *Limulus* CRP 1.1 and 3.3, as the number of amino acid substitutions between 1.4 and 1.1/3.3 is greater than those between 1.1 and 3.3. Therefore, 1.1 and 3.3 are more similar to each other than 1.4 and the high bootstrap values suggest high reliability in the branching. This is shown by directly inspecting the amino acid sequences as 1.1 and 3.3 have only 2 amino acid differences between them, whilst there are 23 amino acid differences between 1.4 and 1.1/3.3. *Limulus* CRPs are branched with tCRP2 and *Carcinoscorpius* CRP 2 with a high bootstrap value (940) suggesting higher sequence identity with these protein sequences than tCRP1, tCRP3, or *Carcinoscorpius* CRP 1, which suggests that they could be orthologues. This may suggest functional similarities with tCRP2 such as haemolytic and haemagglutinating properties as shown in table 6.1.

Therefore, from inspection of the phylogenetic tree for the horseshoe crab CRP and SAP with human CRP, it can be said that there are groups of sequences within and between species which have similar sequence and perhaps similar structure and function as mature proteins and could therefore be orthologues. *Limulus* CRP and SAP appear to be

paralogues as they have sequence similarity but different ligand binding properties so could have arisen through gene duplication. It can also be said that although the CRP sequences between horseshoe crab species show high sequence identity, and therefore appear to be orthologues, CRP within species may have slight changes in ligand binding (e.g. tCRP1-3 (Iwaki *et al.*, 1999)) leading to slight changes in function which would mean they were recent paralogues.

Further analysis of conservation between the horseshoe crab CRP and SAP sequences is presented in section 6.4 where the sequences are mapped onto a known structure to highlight known calcium and ligand binding regions.

6.4 Sequence and structure comparisons of horseshoe crab pentraxins with human CRP

The structure of human CRP has already been determined (Shrive *et al.*, 1996; Thompson *et al.*, 1999) and areas of calcium binding, phosphocholine (PC), and putative C1q binding have been identified. Using the structure and sequence of human CRP, it is possible to compare the amino acid composition of these sites with the corresponding amino acids in the horseshoe crab CRP and SAP-like pentraxins. This would indicate whether the horseshoe crab CRP and SAP could possibly have the same structure and/or function of human CRP, and also allow comparison of binding sites between the horseshoe crab CRP and SAP. The horseshoe crab CRP and SAP pentraxins are composed of *Limulus polyphemus* (Atlantic horseshoe crab) CRP 1.1, 1.4, and 3.3, and SAP; *Tachypleus tridentatus* (Japanese horseshoe crab) CRP 1-22 (1-7 tCRP2, 8-15 tCRP1, 16-22 tCRP3); and *Carcinoscorpius rotundicauda* (Southeast Asian horseshoe crab) CRP named according to accession numbers 1-4 and 6-9 in subgroups CRP 1 (7-9) and CRP 2 (1-4 and 6). Human CRP and the horseshoe crab pentraxins do not appear to have as high levels of identity between their sequences (see figure 6.3) as those seen within the horseshoe crab

CRP and SAP. Multiple sequence alignments (figure 6.9) show areas of sequence homology of horseshoe crab CRP, and SAP sequences with each other and human CRP indicating conserved structure and function between the horseshoe crab CRPs and SAP with some areas of human CRP that will be discussed later. The multiple sequence alignment also shows that the differences in phylogenetic analysis are due to single amino acid mutations not block changes in amino acids that would occur with splice variants, which would be unlikely as no introns have been found in horseshoe crab CRP and SAP genes.

```
Tachypleus_17             --------KVKFPPSTASSFPRLVMDGTLPNLHELTLCYWFKAHRLDRRLHVFSYAQNATDNEILAYVEAGNT----
Tachypleus_18             --------KVKFPPSTASSFPRLVMDGTLPNLHELTLCYWFKAHRLDRRLHVFSYAQNATDNEILAYVEAGNT----
Tachypleus_16             --------KVKFPPSTASSFPRLVMDGTLPNLHELTLCYWFKAHRLDRRLHVFSYAQNATDNEILAFVEAGNT----
Tachypleus_19             --------KVKFPPSTASSFPRLVMDGTLPNLHELTLCYWFKAHRLDRRLHVFSYAQNATDNEILAFVEAGNT----
Tachypleus_22             --------KVKFPPSTASSFPRLVMDGTLPNLHELTLCYWLKAHRLDRRLHVFSYAQNATDNEILAFVEAGNT----
Tachypleus_20             --------KVKFPPSTASSFPRLVMDGTLPNLHELTLCYWFKAHRLDRRLHVFSYAQNATDNEVLAFVEAGNT----
Tachypleus_21             --------KVKFPPSTASSFPRLVMDGTLPNLHELTLCYWFKAHRLDRRLHVFSYAQNATDNEILAFVEAGNT----
Limulus_polyphemus_CRP_1.1 LEEGEITSKVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKVNRLKGTLHMFSYATAKKDNELLTLIDEQGD----
Limulus_polyphemus_CRP_3.3 LEEGEITSKVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKVNCLKGTLHMFSYATAKKDNELLTLLDEQGD----
Limulus_polyphemus_CRP_1.4 LEEGEITSKVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKVNHLKSTLTIFSYNTAKNDNELLTSLEKQGA----
carcino_CRP2_6            --------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKIHRLKASLHMFSYATTGKDNEILTFINQQGD----
carcino_crp2_4           --------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKIHRLKASLHMFSYATTGKDNEILTFINQQGD----
carcino_crp2_3           --------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKIHRLKASLHMFSYATTGKDNEILTFINQQGD----
carcino_crp2_1           --------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKIHRLKASLHMFSYATTGKDNEILTFINQQGD----
carcino_crp2_2           --------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKIHRLKASLHMFSYATTGKDNEILTFINQQGD----
Tachypleus_3             --------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKIHRLKASLHMFSYATTEKDNEILTFLNQQGL----
Tachypleus_4             --------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKIHRLKASLHMFSYATTEKDNEILTFLNQQGD----
Tachypleus_1             --------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKIHRLKASLHMFSYATTEKDNEILTFLNQQGD----
Tachypleus_2             --------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKIHRLKASLHMFSYATTEKDNDILTFLNQQGD----
Tachypleus_5             --------KVKFPPSSSPSFPRLVMIGTLPDLHEITLCYWFKIHHLKASLTMFSYASTETDNEIATSLDEQGD----
Tachypleus_6             --------KVKFPPSSSPSFPRLVMIGTLPDLHEITLCYWFKIHHLKASLTMFSYASTETDNEIATSLDEQGD----
Tachypleus_7             --------KVKFPPSSSPSFPRLVMIGTLPDLHEITLCYWFKIHHLKASLTMFSYASTETDYEIATSLDEQGD----
Tachypleus_9             --------KVKFPPSSSPSFPRLVMVGTLPDLHEITLCYWFKLHRLKGAPHIFSYATSETDNEILTSLNENGD----
Tachypleus_8             --------KVKFPPSSSPSFPRLVMVGTLPDLHEITLCYWFKLHRLNGAPHIFSYATSETDNEILTSLNENGD----
Tachypleus_14            --------KVKFPPSSSPSFPRLVMVGTLPDLHEITLCYWFKLHRLNGAPHIFSYATSETGNEILTSLNENGD----
Tachypleus_10            --------KVKFPPSSSPSFPRLVMVGTLPDLHEITLCYWFKLHRLNGAPHIFSYATSETDNEILTSLNENGD----
Tachypleus_12            --------KVKFPPSSSPSFPRLVMVGTLPDLHEITLCYWFKLHRLNGAPHIFSYATSETDNEILTSLNENGD----
Tachypleus_13            --------KVKFPPSSSPSFPRLVMVGTLPNLHEITLCYWFKLHRLNGAPHIFSYATSETDNEILTSLNENGD----
Tachypleus_11            --------KVKFPPSSSPSFPRLVMVGTLPDLHEITLCYWFKLHRLNGAPHIFSYATSETDNEILTSLNENGD----
Tachypleus_15            --------KVKFPPSSSPSFPRLVMVGTLPDLHEITLCYWFKLHRLNGAPHIFSYATSETDNEILTSLNENGD----
carcino_crp1_8           --------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKLHRLKGTPHIFSYANSETDNEILTSLNEQND----
carcino_crp1_9           --------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKLHRLKGTPHIFSYANSETDNEILTSLNEQND----
carcino_crp1_7           --------KVKFPPSSSPSFPRLVMVGTLPDLQEITLCYWFKLHRLKGTPHIFSYATSETDNEILTSLNEQND----
Limulus_SAP              ---AVDIRKISFPGTQNPKFPHLRFMQTLPAVRQLTVCQRIKPFHRN-TGYIFSCATSNQDNQFITSMYVKSDGTLN
human_crp                -QTDMSRKAFVFPKESDTSYVSLKAPLTKPLKAFTVCLHFYTELSSTRGYSIFSYATKRQDNEILIFWSKDIG----
human_CRP numbering      .1.......10........20........30........40........50........60........70......
```

Figure 6.9 Sequence alignment of horseshoe crab CRP and SAP, and human CRP sequences. Green background denotes identical residues, yellow background denotes similar residues and white background denotes different residues.

```
Tachypleus_17           IGLIIGGKTQLHIACPSDIEVGKWHHVCYLWSSWEGSSSLMMNGYRIYGTTGVTKGKTVQAEGTVVLGQDQDTVGGGF
Tachypleus_18           IGLIIGGKTQLHIACPSDIEVGKWHHVCYLWSSWEGSSSLMMNGYRCYGNTTGVTKGKTVQAEGTVVLGQDQDTVGGGF
Tachypleus_16           IGLIGGGKTQLHIACPSDIEVGKWHHVCYLWSSWEGSSSLMMNGYRCYGTTGVTKGKTVQAEGTVVLGQDQDTVGGGF
Tachypleus_19           IGLIIGGKTQLHIACPSDIEVGKWHHVCYLWSSWEGSSSLMMNGYRIYGNTTGVTKGKTVQAGGTVVLGQDQDTVGGGF
Tachypleus_22           IGLIIGGKTQLHIACPSDIEVGKWHHVSYLWSSWEGSSSLMMNGYRGYGNTTGVTKGKTVQAGGTVVLGQDQDTVGGGF
Tachypleus_20           IGLIIGGKTQLHIACPSDIEVLKWHHVCYLWSSWEGSSSLMMNGYRCYGTTGVTKGKIVQAGGTFVLGQDQDTVGGGF
Tachypleus_21           IGLIIGGKTQLHIACPSDIEVGKWHHVCYLWSSWEGSSSLMMNGYCCYGNTTGVTKGKVQAGTTFVLGQDQDTVGGGF
Limulus_polyphemus_CRP_1.1   FLFNVHGAPQLKVQCPNKIHIGKWHHVCHTWSSWEGEATIAVDGFHCKGNATGIAVGRGLSQGGLVVLGQDQDSVGGK
Limulus_polyphemus_CRP_3.3   FLFNVHGAPQEKVQCPNKIHIGKWHHVCHTWSSWEGEATIAVDGFHCKGNATGIAVGRTLSQGGLVVLGQDQDSVGGKE
Limulus_polyphemus_CRP_1.4   THMEVHGAPQLKVQCPNKIHIGKWHHVCHTWSSWEGEATIGVDGFHIKGNATGIAMGVTLSQGGLVVLGQEQDSVGGEY
carcino_CRP2_6          FLFNVHGSPMLKVQCPNKIHIGRWHHALHTWSSWKEEATTNVDGFHCVGNATGIATGATLRQGGLVVLGQDQDTVGGGF
carcino_crp2_4          FLFNVHGSPMLKVQCPNKIHIGRWHHACHTWSSWKEEATTNVDGFHCVGNATGIATGAELRQGGLVILGQDQDTVGGGF
carcino_crp2_3          FLFNVHGSPMLKVQCPNKIHIGRWHHACHTWSSWKEEATTNVDGFHCVGNATGIATGATLRQGGSVVLGQDQDTVGGGF
carcino_crp2_1          FLFNVHGSPMLKVQCPNKIHIGRWHHACHTWSSWKGEATTNVDGFHCVGNATGIRTGATLRQGGLVVLGQDQDTVGGGF
carcino_crp2_2          FLFNVHGSPMLKVQCPNKIHIGRWHHACHTWSSWKEEATTNVDGFHCVGNATGIATGATLRQGGLVVLGQDQDTVGGGF
Tachypleus_3            FLFNVHGTPMLKVQCPNKIHIGRWHHACHTWSSWKEEATTAVDGFHCEGNATGIAMGATLRQGGLVVLGQDQDTVGGGF
Tachypleus_4            FLFNVHGTPMLKVQCPNKIHIGRWHHACHTWSSWKEEATTAVDGFYCEGNATGIAMGATLRQGGLVVLGQDQDTVGGGF
Tachypleus_1            FLFNVHGTPMLKVQCPNKIHIGRWHHACHTWSSWKEEATTAVDGFHLKNATGIRMEATLRQGGLVVLGQDQDTVGGGF
Tachypleus_2            FLFNVLTPMLKVQCPNKIHIGRWHHACHTWSSWKGEATTAVDGFHLENATGIRMEATLRQGGLVVLGQDQDTVGGGF
Tachypleus_5            IFLLVHGNPQLKIQCKNKIHIGRWYHVCHTWSSWEGDTTIAVDGFHCEGNATGIAMGATLRQGGLVVLGQDQDTVGGGF
Tachypleus_6            EFLLVHGNPQLKTQCKNKIHIGRWYHVCHTWSSWEGETGIAVDGFHLEGNATGIAMGATLRQGGLVVLGQDQDTVGGGF
Tachypleus_7            EFLLVHGNPQLKTQCKNKIHIGRWYHVCHTWSSWEGETFIAVDGFHCEENATGIATGATLRQGGLVVLGQDQDTVGGGF
Tachypleus_9            FLFNIHGNTQLNVQCNNKILAGRWHHVCHTWSSWEGEATIAVDGFHGKGNATGKATGVIFRQGGLVVIGQDQDSVGGGF
Tachypleus_8            FLFNIHGNTQLNVQCNNKILAGRWHHVCHTWSSWEGEATIAVDGFHCKGNATLKATGVTFRQGGLVVIGQDQDSVGGGF
Tachypleus_14           FLFNIHGNTQLNVQCNNKILAGRWHHVCHTWSSWEGEATIAVDGFHGKGNATEKATGVTFRQGGLVVIGQDQDSVGGGF
Tachypleus_10           FLFNIHGNTQLNMRCNNKILAGRWHHVCHTWSSWEGEATIAVDGFHCKGNATGKATGVTFRQGGLVVIGQDQDSVGGGF
Tachypleus_12           FLFNIHGNTQLNHVQCNNKILAGRWHHVCLLWSSWEGEATIAVDGFHCKENATGKATGVTFRQGGLVVIGQDQDSVGGGF
Tachypleus_13           FLFNIHGNTQLNVQCNNKILAGRWHHVCRTWSSWEGEATIAVDGFHCKENATGKATGVIFRQGGLVVIGQDQDSVGGGF
Tachypleus_11           FLFNIHGNTQLHVQCNNKILAGRWHHVCHTWSSWEGEATIAVDGFHCKENATGKATGVIFQGGLVVIGQDQDSVGGGF
Tachypleus_15           FLFNIHGNTQLNVQCNNKILAGRWHHVCHTWSSWEGEATIAVDGFHCKENATGKATGVIFRQGGLVVIGQDQDSVGGGF
carcino_crp1_8          FLFNIHGKTQLNVQCNNKIHAGRWHHVCHTWSSWEGEATTAVDGFRCKGNATGKAMGVIFIQGGLVVLGQDQDSVGGGF
carcino_crp1_9          FLFNIHGKEQLNVQCNNKIHAGRWHHVCHTWSSWEGEATTAVDGFRCKENATGKAMGVIFRQGGLVVLGQDQDSVGGGF
carcino_crp1_7          FLFNIHGKTQLNVQCNNKIHAGRWHHVCHTWSSWEGEATTVVDGFRCKENATGKAMGVIFRQGGLVVLGQDQDSVGGGF
Limulus_SAP             LGLQVNASSNKYISGEIEIELGQIYHVCIVISGVDGRMAVYANGSPEGLTMENVGKIHQISAIGTVVIGQEQDKIGGGF
human_crp               YSFTVGGSEIIFEVPEVTVAP---VHICTSIEGASIIVEFWGEKPR--VRKSLKKGYTVGAEASIILGQEQDSFGGNF
human CRP numbering     ........80.........90..........100.......110..........120.......130.......140.....
```

```
Tachypleus_17            V SESMEGELSELNM NSVLNSNQILHLSNC DVSERHLYGNI QEEKTSFMFYDG------------
Tachypleus_18            V SE MEGELSEMNMWNSVL SNQILHLSNC DVSERHLYGNI QEEKTSFMFYDG-----------
Tachypleus_16            V SE MEGELSELNMWNSVL SNQILHLSNC DVSERHLYGNI IQ EKTSFMFYDG-----------
Tachypleus_19            IVSESMEGELSELNM NS D K ILHLSKCADV ERYLYGNI QE KTSFMFYDG-----------
Tachypleus_22            IVSE MEGELSELNMWNSA DSK ILHLSKCADV ERYLYGNI Q K  SFMFYDG-----------
Tachypleus_20            IVSE MEGELSELNM NSA DSK IL LSKCADV ERYLYGNI Q K  S MFYDG-----------
Tachypleus_21            IVSE MEGELSELNMWNSA DSKQ LHLSKC DV ERYLYGNI Q K  S MF YDG-----------
Limulus_polyphemus_CRP_1.1  DATQSLEGELSELNLW  VLNHEQIK LSKCA PSERHIYGNI IQWDKTQFKAYDGVVLSPNEICA-
Limulus_polyphemus_CRP_3.3  DATQSLEGELSELNLWMTVLNHEQIK LSKCA PSERHIYGNI IQ DKTQFKA DGVVLSPNEICA-
Limulus_polyphemus_CRP_1.4  D EQSLEGELSELNLW TVLNHEQIK LSKCAHPSERHIH NIIQ DKTQ QA DGVVLSPNEICA-
carcino_CRP2_6           DANQSLEGELSELNLW AVLNHEQIKHLSKC HSERHIYGNIIRNDKTQF AYDGVVLSPNEICA-
carcino_crp2_4           DANQSLEGELSELNLW A LNHEQIKHLSKC HSERHIYGNIIR DKTQFAA DGVVLSPNEICA-
carcino_crp2_3           DANQSLEGELSELNLW A LNHEQIKHLSKC HSERHIYGNIIRNDKTQFRAYDGVVLSPNEICA-
carcino_crp2_1           GANQSLEGELSELNLWDAVLNHEQIKHLSK V HSERHIYGNIIRNDKTQFRAYDGVVLSPNEICA-
carcino_crp2_2           DANQSLEGELSELNLWDA LNHEQIKHLSKCVHHSERHIYGNIIRNDKTQFRL DGVVLSPNEICA-
Tachypleus_3             DAKQSLEGELSELNLWDTVLNHEQIKHLSKCVHHSERHIYGNIIH NKTQF AYDG-----------
Tachypleus_4             DAKQSLEGELSELNLWDTVLNHEQIKHLSKCVHHSERHIYGNIIH NKTQF A DG-----------
Tachypleus_1            E  SLEGELSELNLWDTVLNHEQIKHLSKC VHHSERHIYGNIIH NKTQF A DG-----------
Tachypleus_2             DAKQSLEGELSELNLWDTVLNHEQIKHLSKCVHHSERHIYGNI IH DKTQFRAYDG-----------
Tachypleus_5             GDG WEGELSELNLWDTVLNHEQIKHLSKCVHHSERHIYGNIIHWDKTQFRA DG-----------
Tachypleus_6            DGDG WEGELSELNLWDTVLNHEQIKHLSKCVHHSERHIYGNIIH WDKTQFRVYDG-----------
Tachypleus_7            DGDQSWEGELSELNLWDTVLNHEQIKHLSKCVHHSERHIYGNIIH GKTQFRV DG-----------
Tachypleus_9            DEKQS VGELSELNLW MVLNHEQIKHLSKCAHPSERHIYGNIIHWDKTQFRV DG-----------
Tachypleus_8            DEKQSLVGELSELNLWDMVLNHEQIKHLSKCAHPSERHIYGNIIH DKTQFRVYDG-----------
Tachypleus_14          EEKQS GELSELNLWDMVLNHEQIKHLSKCAHPSERHIYGNIIH DKTQFRVYDG-----------
Tachypleus_10          DEKQSLVGELSELNLW MVLNHEQIKHLSKCAHPSERHIYGNIIH WDKTQFRVYDG-----------
Tachypleus_12          DEKQSLVGELSELNLW MVLNHEQIKHLSKCAHPSERHIYGNIIH DKTQFRVYDG-----------
Tachypleus_13          DEKQSLVGELSELNLWDMVLNHEQIKHLSKCAHPSERHIYGNIIH DKTQFRVYDG-----------
Tachypleus_11          DEKQSLVGELSELNLW MVLNHGQIKHLSKCAHPSERHIYGNI HWDKTQFRVYDG-----------
Tachypleus_15          V SE MEGELSELNM NSVLNSNQIL LEN  DVSERHLYGNI QWEKTSFMFYDG-----------
carcino_crp1_8         DAKQSLVGELSELNLWDMVLNHEQIKHLSE V PSERHIYGNVIH DKTQ QAYDGVVLSPNEICT-
carcino_crp1_9         DAKQSLVGELSELNLWDMVLNHEQIKHLSE V PSERHIYGNVIHWDKTQHQAYDGVALSPNEICA-
carcino_crp1_7         DAKQSLVGELSELNLWDMVLNHEQIKHL E V PSERHIYGNVIHWDKTQ QAYDGVVLSPNEICA-
Limulus_SAP            EEQE WSGELSDLQVWDEALTTH VSTVAS NGIRP ---GNVISWMEDS VADDGVIVGISHMCSL
human_crp             EGSQSLVGDIGNVMWDFV SPDE NTIYLGGPF PNVLNWRALKYEVQGEVFTKPQLWP-------
human CRP numbering    ...150.......160.......170.......180.......190.......200............
```

6.4.1 Overview of areas of conservation between horseshoe crab CRP and SAP pentraxins using human CRP as reference structure

To gain an overall view of horseshoe crab CRP and SAP sequence conservation, the horseshoe crab CRP and SAP sequences were mapped onto structures of both human CRP and *Limulus* SAP subunit A using Protein Explorer and ConSurf. This identified conserved parts of the horseshoe crab sequences and allowed comparison with the calcium, phosphocholine (PC), and putative C1q binding areas on human CRP, and where the calcium binding site is on *Limulus* SAP.

When the horseshoe crab CRP and SAP sequences were mapped onto human CRP using ConSurf (see figure 6.10), most of the conserved region was around the calcium binding area, with a more variable area seen at the ligand binding pocket adjacent to the calcium binding site. There were areas of "no information" which were coloured yellow due to ConSurf being unable to calculate the conservation between residues in sequences that are very closely related. The ConSurf view of the putative human CRP C1q cleft (figure 6.11) showed some amino acids with high conservation and a few that are variable, although most of the residues showed "no information" at all.

Figure 6.10. Horseshoe crab CRP and SAP sequences mapped onto the structure of human CRP chain A using ConSurf showing the calcium and PC binding site. Conservation between sequences is shown on the structure as the ConSurf colours below. The box highlights the human calcium ion binding sites and the ligand binding site where phosphocholine is shown as sticks and calcium ions shown in green. The most conserved residues are labelled.
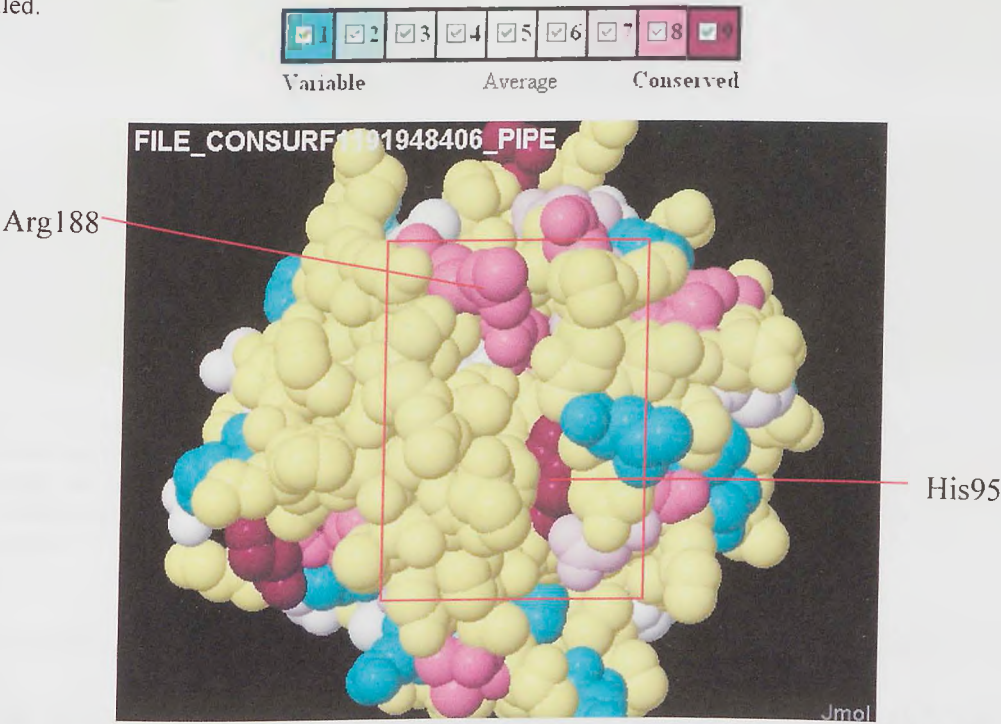
Figure 6.11. Horseshoe crab CRP and SAP sequences mapped onto the structure of human CRP chain A using ConSurf showing the putative human C1q binding cleft. Conservation between sequences is shown on the structure as the ConSurf colours below. The box highlights the human C1q cleft. The most conserved residues are labelled.

Using Protein Explorer it is possible to look at the conservation of all the residues even though the sequences are very similar. However, Protein Explorer only identifies residues that are identical, similar, or different, so does not allow for grading of the sequence

310

conservation. The consensus percentage of identity in Protein Explorer can be changed, for example if the consensus percentage was 100% all the amino acids at a certain position would have to be identical for the residue to be identified as identical; at 90% consensus percentage only 90% of them would have to be identical, and so forth. Changing the consensus percentage allowed for identification of residues that were only partly conserved. When the horseshoe crab CRP and SAP sequences were mapped onto human CRP chain A structure using Protein Explorer, only a few residues were completely conserved at the calcium and PC binding site (see figure 6.12).
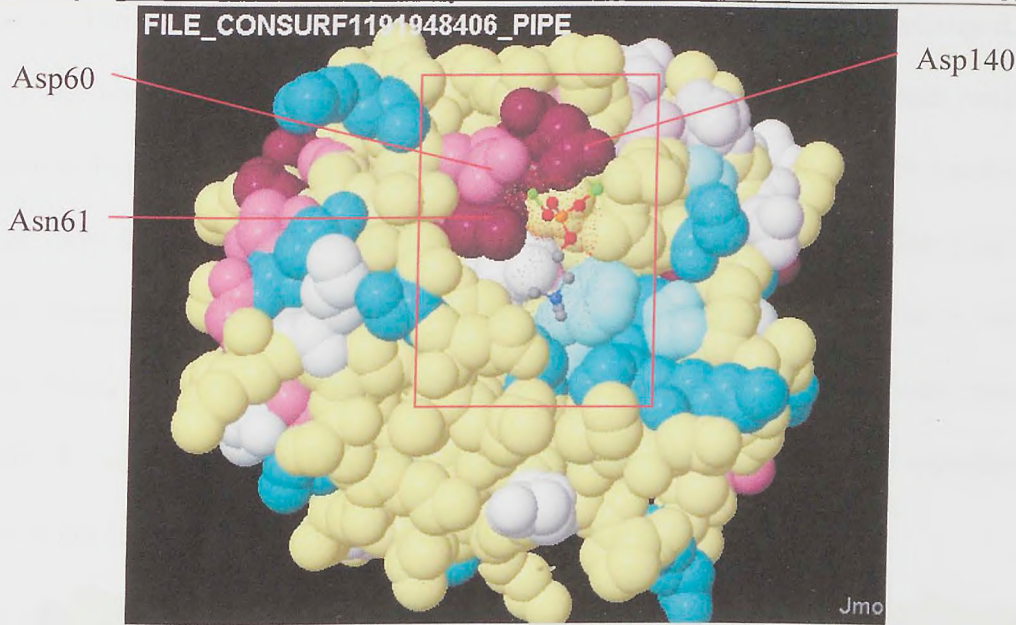

Figure 6.12. Horseshoe crab CRP and SAP sequences mapped onto human CRP chain A using Protein Explorer showing the calcium and PC binding site. The box highlights the human calcium and PC binding site, showing the calcium ions (dark grey spheres) and PC. Green indicates identical residues, yellow indicates similar residues, and red indicates different residues. The consensus percentage is shown below each protomer.

As the consensus percentage was lowered, more residues became classed as identical in the calcium and PC binding site, although more so at the calcium ion binding site than the PC binding site. The PC binding site seems to be less conserved, indicating that there is much variation within the residues of the horseshoe crab CRP and SAP sequences corresponding to those in the PC binding site of human CRP, which may perhaps be related to structure and/or function. At 60% consensus percentage all of the calcium binding residues were identified as identical, indicating that this area is important for structure and/or function of calcium binding throughout the horseshoe crab CRP and SAP. At 30% consensus percentage, all of the residues on the calcium and PC binding face of

the protomer are identical which means that at least 30% of all the horseshoe crab CRP and

SAP sequences are identical at these positions.

Protein Explorer showed that only a few horseshoe crab CRP and SAP residues

corresponding to the putative C1q cleft on human CRP are 100% conserved between

human CRP and the horseshoe crab CRP and SAP sequences, but a majority are conserved
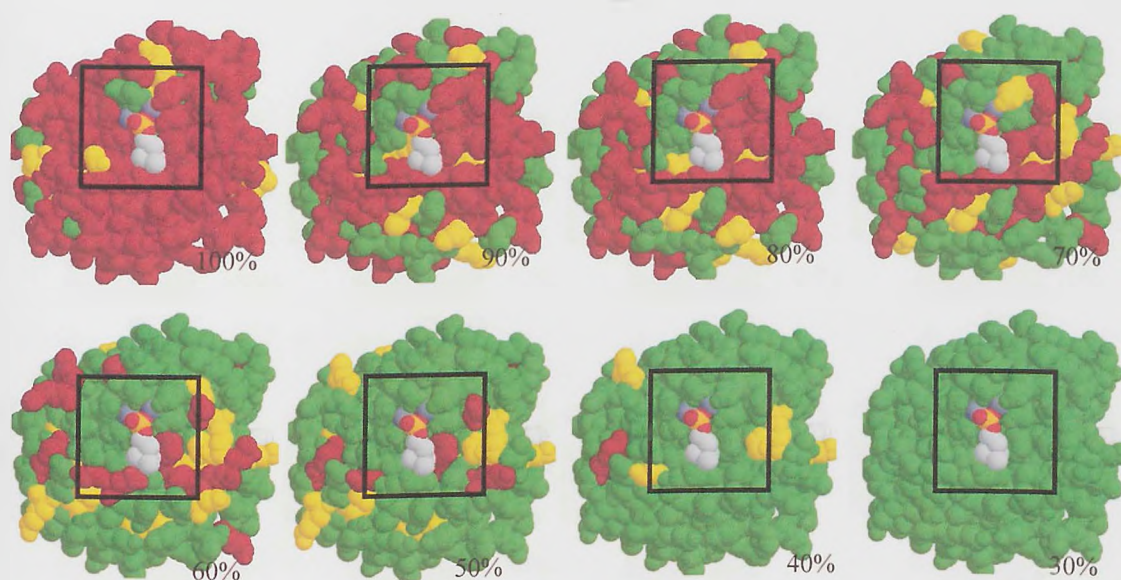
at 90% consensus percentage (see figure 6.13).
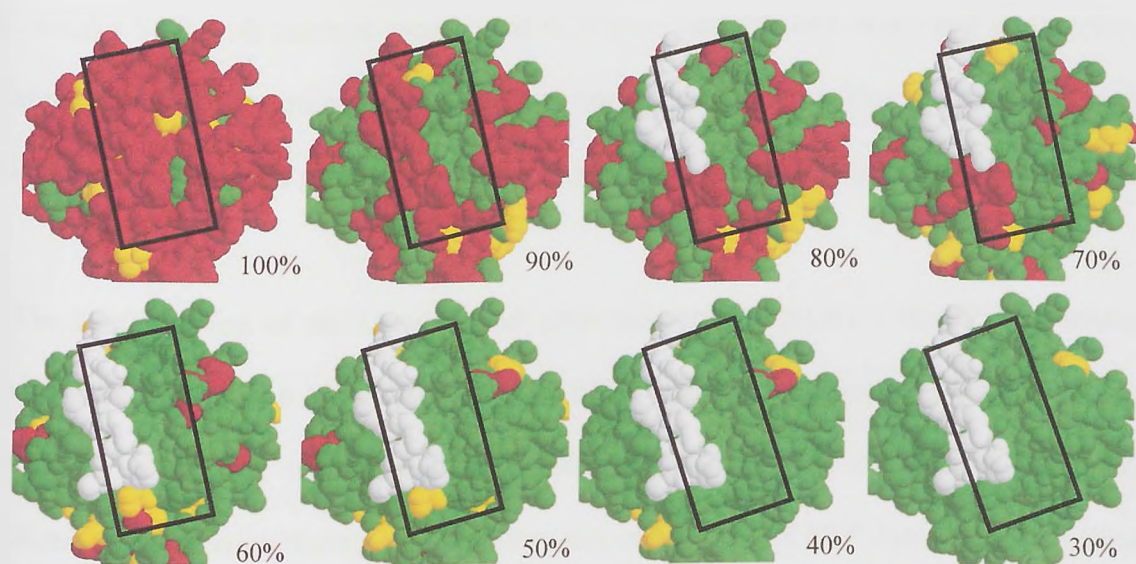
Figure 6.13. Horseshoe crab CRP and SAP sequences mapped onto human CRP chain A using Protein Explorer showing the putative human C1q binding cleft. The box highlights the putative C1q binding cleft. Green indicates identical residues, yellow indicates similar residues, red indicates different residues, and white indicates residues with "no-information". The consensus percentage is shown below each protomer.

At 80% some of the residues that form part of one wall of the putative C1q cleft are

classed as having "no-information" meaning that these residues do not align with any other

residues in the multiple sequence alignment. More residues with "no-information" appear

at 60% consensus percentage making up a large part of one wall of the putative C1q cleft.

These residues are the first seven residues of the human CRP sequence and are found as

having "no-information" at 80% consensus percentage because the sequence extends

beyond a majority of the multiple sequence alignment, and conservation cannot be

calculated. As most of the putative C1q cleft was conserved at 90% consensus percentage,

and all of it was conserved at 60% consensus percentage, except those parts with "no-

information", the residues involved are fairly highly conserved amongst the horseshoe crab

CRP and SAP perhaps indicating that they have similar structure and/or function.

6.4.2 Overview of areas of conservation between horseshoe crab CRP and SAP pentraxins using *Limulus* SAP as reference structure

The structure of *Limulus* SAP was also used in ConSurf and Protein Explorer to explore possible conservation between the horseshoe crab CRP and SAP sequences. The structure of *Limulus* SAP with calcium ions bound is known (unpublished data), and so therefore analysis of the area surrounding the calcium ions should give an idea of conservation of calcium ion binding and perhaps ligand binding too in the adjacent ligand binding pocket.

The ConSurf view of the *Limulus* SAP protomer with horseshoe crab CRP sequences mapped onto it (figure 6.14) shows that most of the residues surrounding the calcium ions are highly conserved, with only a few residues that show "no-information" (yellow) indicating conservation throughout the horseshoe crab CRP and SAP. However, the ligand binding pocket on the surface of the *Limulus* SAP near the calcium binding site seems to be more variable, indicating a possibility of different ligand binding properties throughout horseshoe crab CRP and SAP.
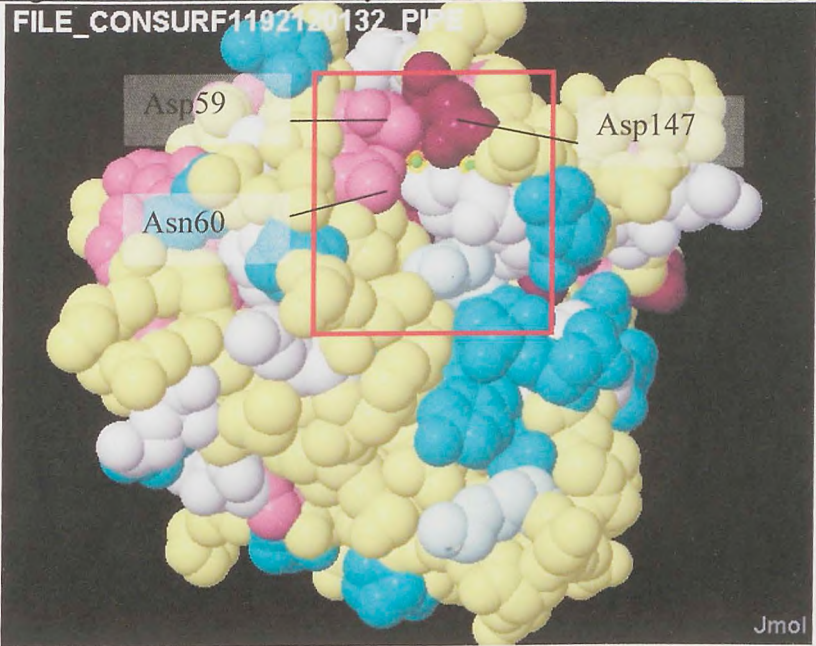
Figure 6.14. Horseshoe crab CRP sequences mapped onto *Limulus* SAP chain A using ConSurf and coloured with Consurf colours as below. The box highlights the *Limulus* SAP calcium binding area, and the calcium ions are shown as green spheres. The most conserved residues are labelled.

A view of the *Limulus* SAP protomer in Protein Explorer with horseshoe crab CRP sequences mapped onto it (see figure 6.15) showed that a few residues involved in calcium binding are totally conserved, and most residues are conserved at 90% consensus percentage indicating that the corresponding region in the horseshoe crab CRP sequences is conserved in calcium binding site structure and function. None of the residues in the pocket next to the calcium ions are totally conserved, although most are conserved at 60% consensus percentage. The variability in the ligand binding pocket shows that the corresponding sequence areas in the horseshoe crab CRP sequences are dissimilar indicating variability in ligand interaction.

Therefore, mapping the horseshoe crab CRP and SAP sequences onto the structure of human CRP and *Limulus* SAP suggested that the calcium binding site, ligand binding site and putative C1q binding site (on human CRP structure only) are areas of conservation, with the calcium binding site showing the highest conservation. However, small areas with "no information" on both the human CRP and *Limulus* SAP structure indicate dissimilar

314

structures to the horseshoe crab CRPs, which may themselves have slight structural variation between species. More information is known about human CRP calcium, PC and putative C1q binding interactions than *Limulus* SAP interactions, so it is more useful to use human CRP as a model for displaying differences in horseshoe crab CRP and SAP at calcium, PC, and putative C1q binding areas.
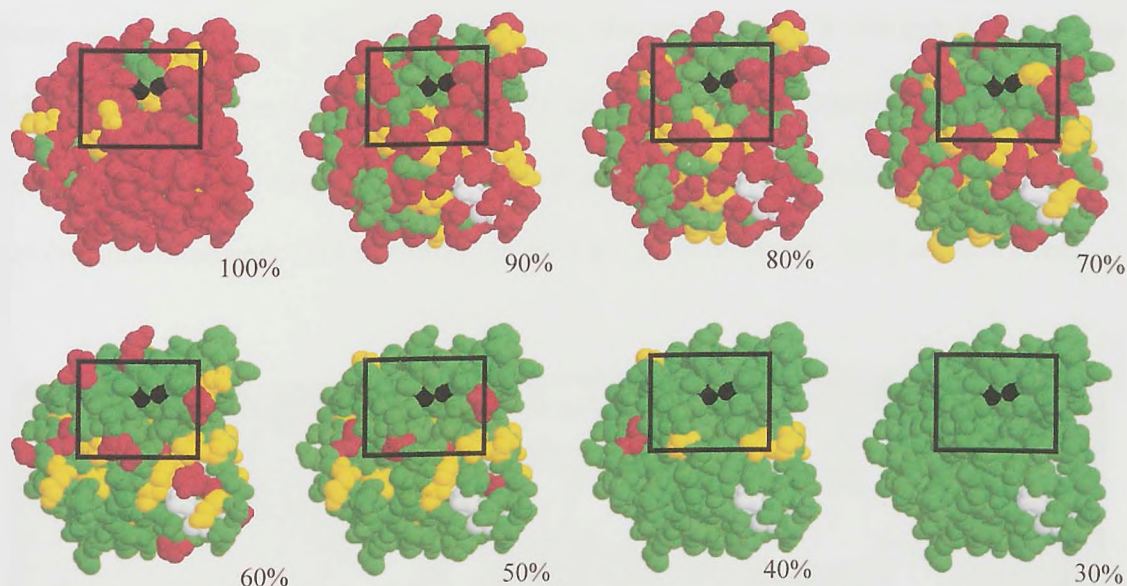
Figure 6.15. Horseshoe crab CRP sequences mapped onto *Limulus* SAP chain A using Protein Explorer showing the calcium binding site and ligand binding pocket. The box highlights the *Limulus* SAP calcium binding area where the calcium ions are black. Residues are coloured according to conservation: green (identical), yellow (similar), red (different), and white ("no-information"). The consensus percentage is shown below each protomer.

Sections 6.4.3 to 6.4.5 use Protein Explorer to map the horseshoe crab CRP and SAP sequences onto human CRP subunit A structure, and the conservation of individual amino acids were analysed in detail at areas corresponding to calcium ion binding, PC/ligand binding and putative C1q binding by reducing the consensus percentage of the alignment in 10% steps. The residue numbers given are those for human CRP.

6.4.3 Calcium ion binding site conservation throughout horseshoe crab CRP and SAP

The first calcium ion in human CRP is co-ordinated by the residues Asp60, Asn61, Glu138, Asp140 and the main chain carbonyl of Glu139; and the second calcium ion is co-ordinated by Glu147, Glu138, Asp140, and Gln150 (Shrive *et al.*, 1996; Thompson *et al.*,

1999) (see figures 6.16 and 6.17). Glu147 extends to close the calcium binding pocket

when binding to the second calcium ion. When the calcium ions are not present, instead of

the calcium binding loop 140-150 staying close to the surface of the protein, it moves away

from the body leaving an area exposed to proteolysis (Ramadan *et al.*, 2002).

The calcium binding site of human CRP can essentially be split up into three parts:

calcium ion one binding site, calcium ion two binding site, and a calcium binding loop

which is defined here as the residues Glu138-Gln150 between the two calcium ion binding

sites without including those residues involved directly in binding. Attempts were made to

group the horseshoe crab CRP according to their amino acid conservation at these sites.
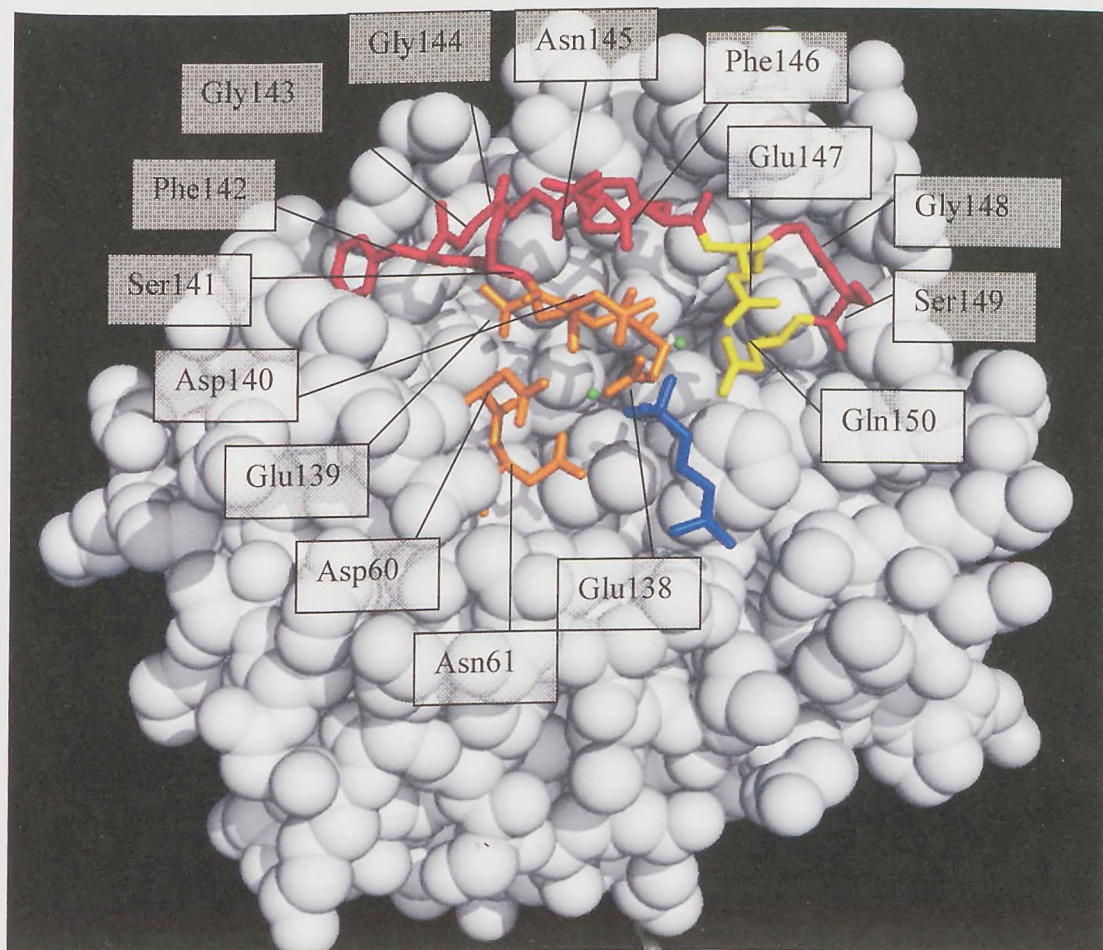
Figure 6.16 Human CRP chain A showing the calcium binding area. Calcium ions are shown as green spheres and phosphocholine is shown in blue. Residues that interact with the calcium ion one are shown in orange, and residues that interact with calcium ion two are shown in yellow. Glu138 and Asp140 that interact with both calcium ion one and two are shown in orange, and the residues that define the calcium binding loop are shown in red.
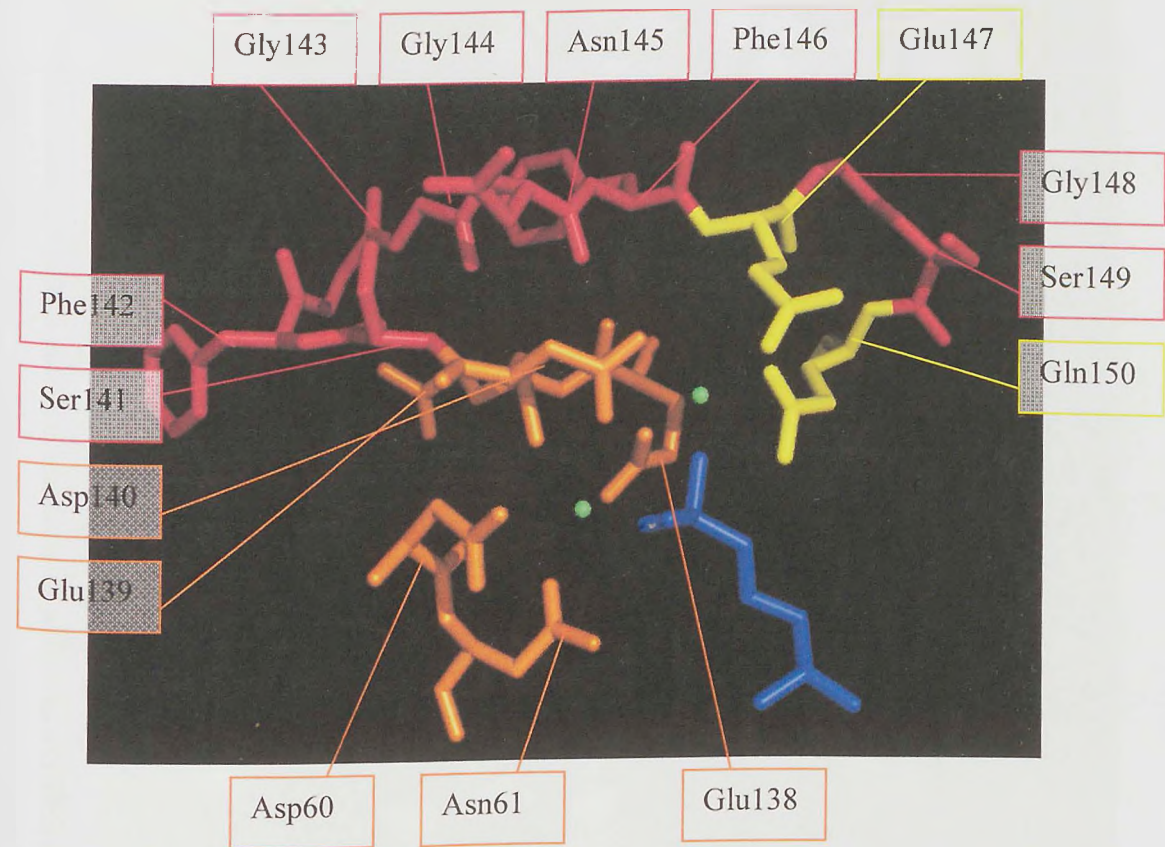
Figure 6.17 Human CRP protomer chain A showing the calcium binding area. Calcium ions are shown as green spheres and phosphocholine is shown in blue. Residues that interact with the calcium ion one are shown in orange, and residues that interact with calcium ion two are shown in yellow. Glu138 and Asp140 that interact with both calcium ion one and two are shown in orange, and the residues that define the calcium binding loop are shown in red.

The calcium ion one binding site shows high conservation (see table 6.4) amongst the

five residues that compose it (see figure 6.18).

Table 6.4 Conservation of residues corresponding to the human calcium ion one binding site between horseshoe crab pentraxins and human CRP. Conservation is shown as percentage of residues that are identical to the most common residue.

hCRP residue number	hCRP residue	Most common residue in horseshoe crab pentraxin and human CRP sequences	Conservation
139	Glu	Glu	100%
140	Asp	Asp	100%
60	Asp	Asp	90%
61	Asn	Asn	90%
138	Glu	Asp	90%

Figure 6.18. Human CRP protomer chain A showing the calcium ion one co-ordinating residues in orange. Calcium ions are shown as green spheres, phosphocholine is shown in blue.

Glu139 and Asp140 are totally conserved, and the Asp60, Asn61, and Glu138 residues have 90% conservation, i.e. only one sequence has to be different for the conservation to drop from 100% to 90%. This means that only individual sequences have a different amino acids from the majority of sequences, rather than groups of sequences with identical amino acids.

At residue 60, only *Tachypleus* CRP14 (tCRP1) has a different residue (glycine) from all the horseshoe crab CRP and SAP sequences, and human CRP that have aspartic acid. At residue 61, only *Tachypleus* CRP7 (tCRP2) has a different residue (tyrosine) from the horseshoe crab and human CRP sequences that have asparagine, which may alter calcium binding through a change in the side chain size even though they are both polar. Residue 138 and the corresponding residues in the horseshoe crab CRP and SAP sequences show two different amino acids at this position: aspartic acid in the horseshoe crab CRPs with

the exception of glutamic acid in human CRP, *Limulus* SAP, and *Limulus* CRP 1.4. Both

aspartic acid and glutamic acid have very similar polar and negatively charged side chains.

Overall, the calcium ion one binding site is highly conserved amongst all the horseshoe

crab pentraxins and human CRP. Where there are different amino acids at residues

corresponding to the calcium ion one binding residues, they do not signify sequence

differences within groups of species, but instead often just a single sequence. Therefore,

despite there being groups of CRP sequences both between the horseshoe crab CRPs (such

as *Limulus* CRP and *Carcinoscorpius* CRP) and within the horseshoe crab species specific

CRPs (such as *Tachypleus* tCRP1, tCRP2, and tCRP3), these do not signify groups of

sequences that have different amino acids for the calcium ion one binding site residues.

The calcium ion two binding site (see figure 6.19) is overall less conserved than the

calcium ion one binding site (see table 6.5).

Table 6.5 Conservation of residues corresponding to the human calcium ion two binding site between horseshoe crab pentraxins and human CRP. Conservation is shown as percentage of residues identical to the most common residue.

hCRP residue number	hCRP residue	Most common residue in horseshoe crab pentraxin and human CRP sequences	Conservation
140	Asp	Asp	100%
138	Glu	Asp	90%
150	Gln	Gln	70%
147	Glu	Asp	60%

Asp140 and Glu138 are involved in both the calcium ion one and two binding sites and

their conservation has previously been described. The residues specific to calcium ion two

binding are Gln150 and Glu147. Gln150 shows 70% conservation, with most sequences

having glutamine at this position. However, all of tCRP3 and one sequence (15) from

tCRP1 group have glutamic acid at this position. *Limulus* SAP also has glutamic acid at

this position which indicates that the tCRP3 group and one of the tCRP1 group (15), have

more similar calcium ion two binding to *Limulus* SAP than they do human CRP or the

other horseshoe crab CRPs. Glutamine has an uncharged side chain whilst glutamic acid

has a negatively charged oxygen which could affect calcium ion binding, however they
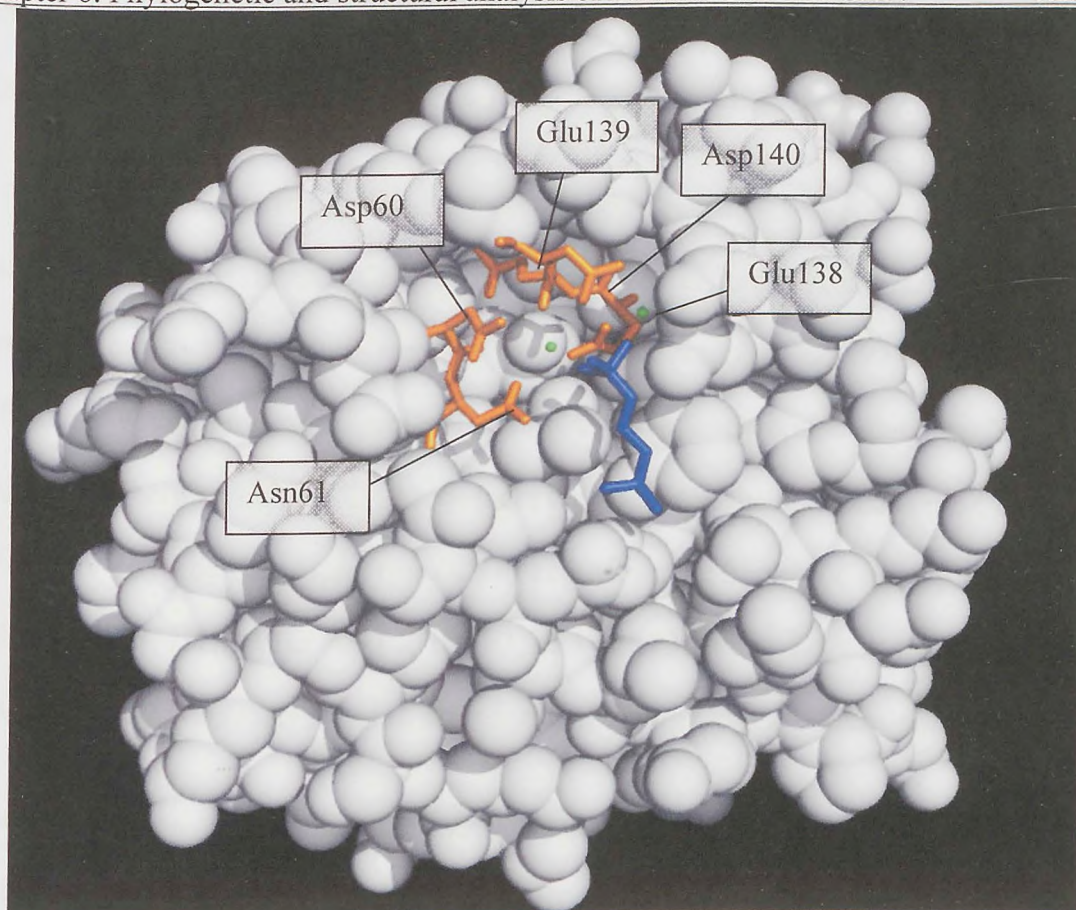
both have polar side chains.
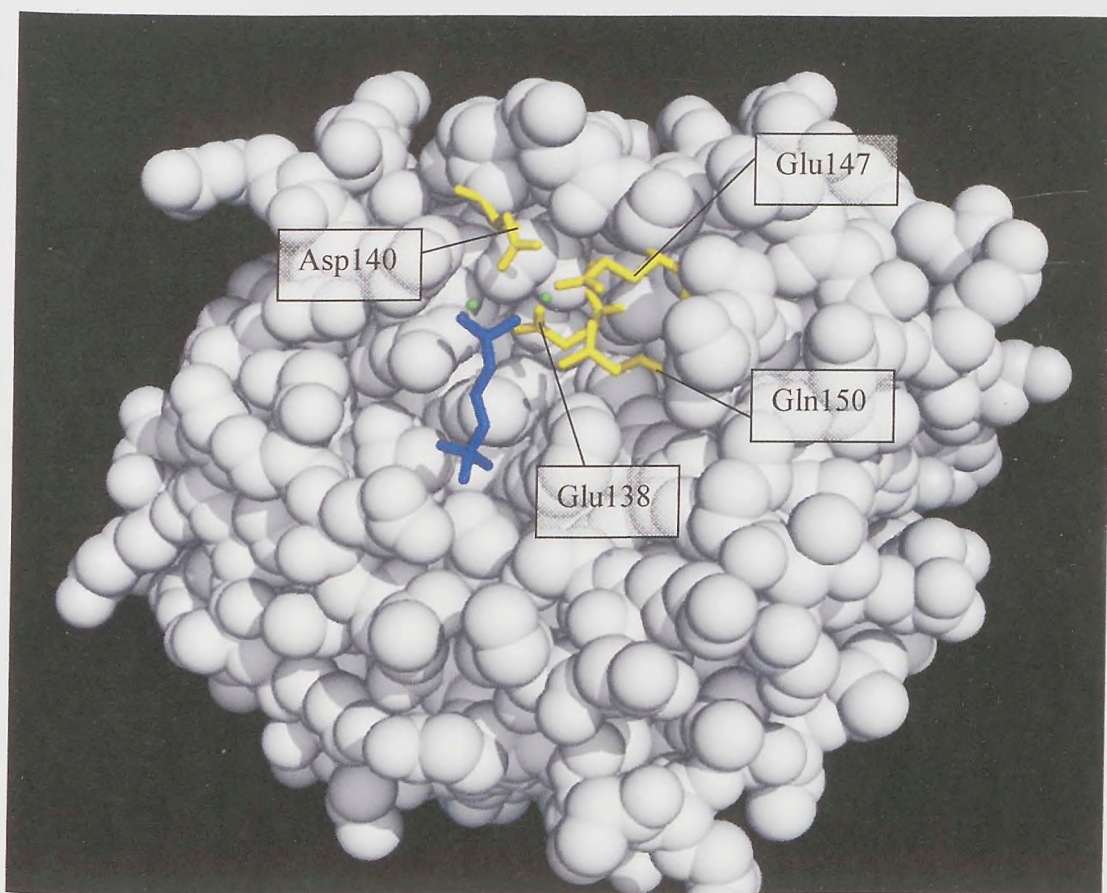
Figure 6.19. Human CRP protomer chain A showing the calcium ion two co-ordinating residues in yellow. Calcium ions are shown as green spheres, phosphocholine is shown in blue

The tCRP3 group and the single tCRP1 sequence (15) also show a different amino acid at residue 147 compared to the other horseshoe crab CRP and SAP. Most of the horseshoe crab CRPs have aspartic acid at this residue, which is similar in charge and side chain polarity to glutamic acid which both human CRP and *Limulus* SAP have at this residue position. In human SAP aspartic acid at this position does not extend into the calcium binding site as it is shorter in length, and therefore the horseshoe crab CRPs may have similar binding properties to human SAP. However, the tCRP3 group and a single tCRP1 sequence (15) have either valine or isoleucine (both with non-polar and hydrophobic side chains) at this residue, indicating that they have similar calcium ion two binding properties within their group but not with the human CRP, *Limulus* SAP, or the other horseshoe crab CRPs. Phylogenetic analysis (section 6.3) showed the tCRP3 group split into two: sequences 16-18, and 19-22, which agrees with the sequences that have valine (16-18) and those which have isoleucine (19-22) at residue 147. *Carcinoscorpius* CRP2 sequence 1

also has a different amino acid at 147, but this is glycine which is not similar to either the amino acids for the horseshoe crab CRPs, SAP, or human CRP. This could mean that *Carcinoscorpius* CRP2 sequence 1 has a unique calcium ion two binding site.

The calcium ion binding loop incorporates residues 130-150 in human CRP, but as some of these residues act directly with the calcium ions and have been previously mentioned, only the remainder of the calcium binding loop was analysed. As none of the residues in the remainder of the loop come into direct contact with the calcium ions (see figure 6.20), it was not expected that there would be a high level of conservation. However, residues 143 and 144 show total conservation, and residues 142 and 146 show 90% conservation (see table 6.6). In human CRP, the glycine at residues 143 and 144 allow for flexibility in the calcium binding loop allowing residues that interact with calcium ions to be positioned correctly (Shrive *et al.*, 1996). This is also likely to be the case in the horseshoe crab pentraxins.

Table 6.6 Conservation of residues corresponding to the human calcium binding loop site between horseshoe crab pentraxins and human CRP. Conservation is shown as percentage of residues identical to the most common residue.

hCRP residue number	hCRP residue	Most common residue in horseshoe crab pentraxin and human CRP sequences	Conservation
143	Gly	Gly	100%
144	Gly	Gly	100%
142	Phe	Val	90%
146	Phe	Phe	90%
145	Asn	Gly	80%
141	Ser	Ser	50%
148	Gly	Ala	50%
149	Ser	Lys	40%
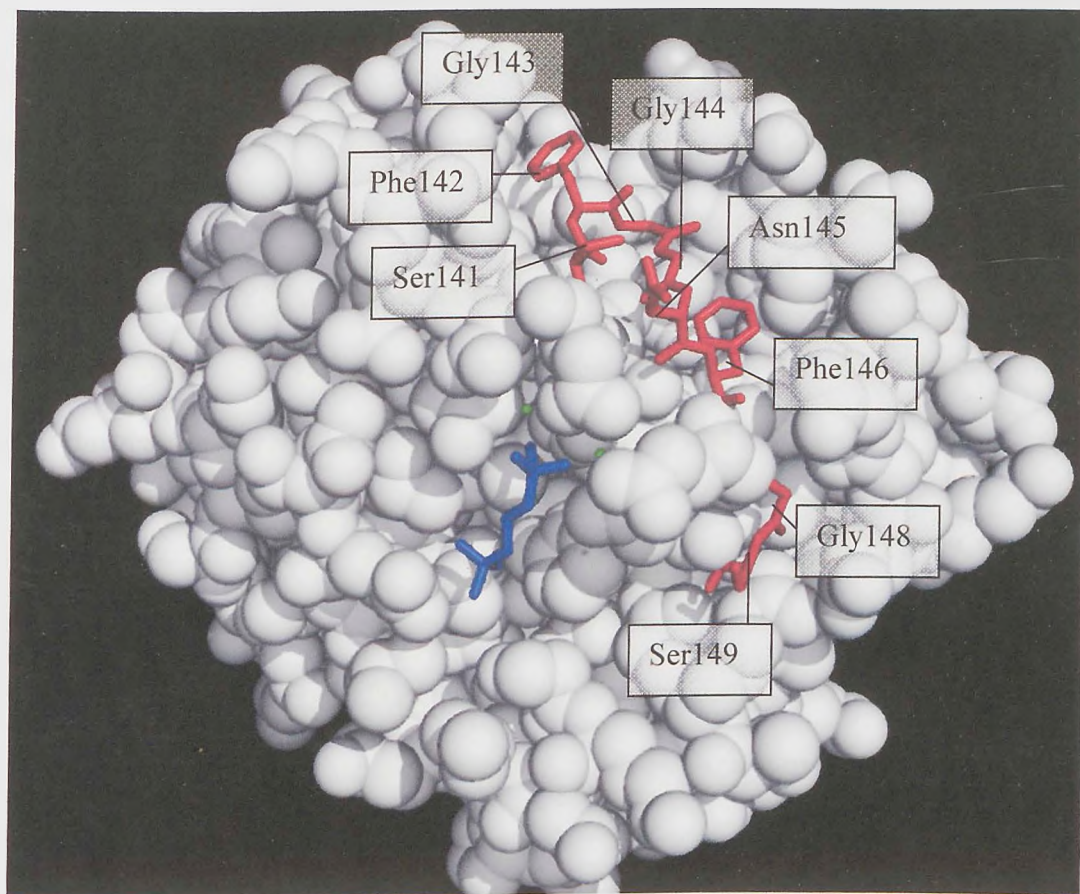
Figure 6.20 Human CRP protomer chain A showing the calcium binding loop in red. Calcium ions are shown as green spheres. Phosphocholine is shown in blue.

At both residue 142 and 146, the amino acids are all hydrophobic and different amino acids are seen in only single sequences. Residue Asn145, in all the horseshoe crab CRPs and SAP except the *Limulus* CRPs, is glycine. In *Limulus* CRP 1.4 the residue corresponding to 145 is glutamic acid which has an oppositely charged but still polar side chain to the lysine amino acid present in *Limulus* CRP 1.1 and 3.3. This indicates a possible difference of calcium loop organisation perhaps corresponding to grouping of *Limulus* CRP's as shown during phylogenetic analysis (see section 6.3).

Remaining residues in the calcium binding loop have relatively low conservation compared to calcium interacting residues, and there seems to be differences in amino acids that are indicative of the groups of horseshoe crab CRP sequences, for example, the tCRP3 group has both valine and alanine at 148. These have similar side chains indicating that function of 148 is similar within the tCRP3 group and with most of the horseshoe crab

CRPs that also have alanine. Some of the tCRP2 group 5-7 (a subgroup of tCRP2 as indicated by phylogenetic analysis (see section 6.3)), share an identical amino acid (glycine) with human CRP at position 148 which indicates more similar function of this residue in tCRP2 (5-7) with human CRP than the remainder of the tCRP2 group or the other horseshoe crab CRPs that have alanine. All of the tCRP1 group with the exception of sequence 15, have an identical amino acid (glutamic acid) with *Limulus* SAP which indicates this residue has similar function in *Limulus* SAP and tCRP1.

Subgroups also appear at position 141 with the *Limulus* CRPs, tCRP1 and *Carcinoscorpius* CRP 1 group, which have the same amino acid (serine) as human CRP. The remaining horseshoe crab CRPs have threonine at position 141 which is similar to serines side chain in size and nature. *Limulus* SAP is an exception as it has lysine that is not similar in side chain size or charge to either serine or threonine, but is still polar. Therefore, even though there are two main groups of sequences according to amino acid differences at position 141, the two amino acids are similar indicating similar residue function.

The residues corresponding to 149 are many and different as shown by a 40% conservation (see table 6.6). Sequences that have polar uncharged side chains similar or identical to serine in human CRP are tCRP3 group (serine), tCRP1 sequence 15 (serine), *Limulus* SAP (glutamine), *Limulus* CRP 1.1 and 3.3 (threonine), and *Carcinoscorpius* CRP 2 (asparagine). tCRP2 have either aspartic acid or lysine, and remaining tCRP1 and *Carcinoscorpius* CRP 1 have lysine at position 149 whilst *Limulus* CRP 1.4 has glutamic acid. Therefore, within and between horseshoe crab species CRP groups there are differences in amino acids at residue 149, which may indicate little side chain importance of residue 149 in calcium ion binding in horseshoe crab CRPs and SAP. However, as Ser149 is directly next to the calcium ion two interacting residue 150 in human CRP, the large differences in side chain charge shown by horseshoe crab pentraxins may affect calcium ion two binding.

Overall, the calcium ion binding loop shows fairly high conservation with only a few residues that have low conservation. Residues 143 and 144 are totally conserved as glycine which, in human CRP, give the calcium binding loop flexibility allowing the residues that interact with calcium to be positioned correctly (Shrive *et al.*, 1996). With the exception of residues 138, 139, 140, 147, and 150, the loop is not involved in direct calcium ion interaction and therefore differences in amino acids in this area do not give any indication of differences in direct calcium binding, but may affect it indirectly by altering the position of the calcium interacting residues within the loop.

6.4.4 Ligand binding site conservation throughout horseshoe crab pentraxins

Phosphocholine (PC) ligand binding by human CRP occurs in a hydrophobic pocket situated adjacent to the calcium binding site formed by Phe66, Leu64, Gly79, Asn61 and Thr76, whilst side chains of Ser68, Ser74, and Glu81 are at the opposite end of the pocket to the calcium binding site (Shrive *et al.*, 1996; Thompson *et al.*, 1999) (see figures 6.21 and 6.22). The phosphate groups of PC bind to the calcium ions, whilst the three choline methyl groups run along the surface of CRP against Phe66 and towards the side chain of Glu81 (Thompson *et al.*, 1999). Mutational studies indicated that Thr76 is important in PC binding (Agrawal *et al.*, 1997), and the crystal structure of CRP with PC showed that Thr76 formed a hydrophobic cavity near PC but did not show any direct interactions with PC (Thompson *et al.*, 1999). The major interactions occur between the two phosphate oxygens of PC to the calcium ions, and between the positively quaternary charged nitrogen of PC and the negative side chain of Glu81 (Thompson *et al.*, 1999).
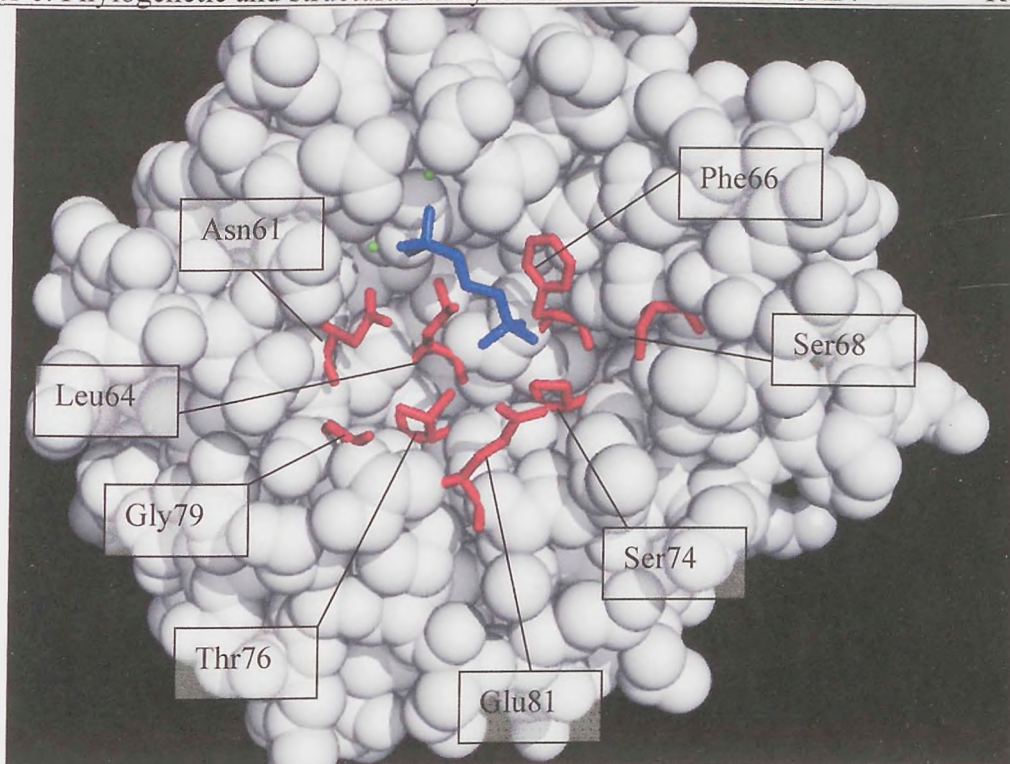
Figure 6.21 Human CRP protomer chain A showing the residues that surround the ligand pocket in red. Calcium ions are shown as green spheres, and phosphocholine is in blue.
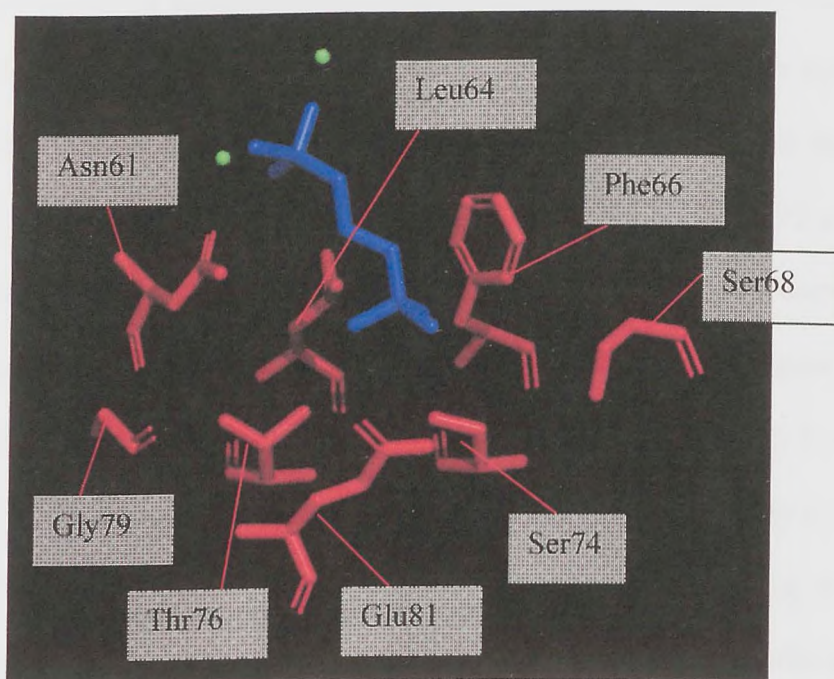
Figure 6.22 Enlarged view of human CRP protomer chain A showing the residues that surround the ligand pocket in red. Calcium ions are shown as green spheres, and phosphocholine is in blue.

Sequences corresponding to the PC binding site have low conservation compared to the calcium ion binding sites. The most important residues in PC binding in human CRP are Phe66 which has least conservation amongst the horseshoe crab CRP and SAP sequences, and Glu81 which only has 50% conservation (see table 6.7).

Table 6.7 Conservation of residues corresponding to the human PC binding site between horseshoe crab pentraxins and human CRP sequences. Conservation shown as percentage of residues identical to the most common amino acid.

hCRP residue number	hCRP residue	Most common residue in horseshoe crab pentraxins and human CRP	Conservation
66	Phe	Ser	40%
68	Ser	Asn	50%
81	Glu	Thr	50%
74	Ser	Leu	60%
76	Thr	Asn	70%
64	Leu	Leu	80%
61	Asn	Asn	90%
79	Gly	Gly	90%

At position 66, the sequences that have phenylalanine the same as human CRP are *Carcinoscorpius* CRP 2 proteins, tCRP2 sequence 14, and tCRP3 sequences 16, and 18-22 indicating that all these sequences may have similar ligand binding properties to human CRP. tCRP1 have serine at residue 66 which has a much smaller and polar side chain than phenylalanine that most of the tCRP2 and tCRP3 groups have. The tCRP3 and tCRP2 groups have sialic acid binding and haemolytic activity although tCRP1 does not (Iwaki *et al.*, 1999). Although some tCRP2 sequences (5-7) have serine at residue 66 the same as tCRP1, perhaps the difference in amino acid at this residue in tCRP3, tCRP2 and tCRP1 contributes to the difference in sialic acid binding and haemolytic properties between these groups. Differences in binding properties may also be present for the *Carcinoscorpius* CRP group 2 which has a phenylalanine at 66, and *Carcinoscorpius* CRP group 1 which has serine. The tCRP3 sequences 17 and 18 are unlike the rest of the tCRP3 group as they have tyrosine and not phenylalanine at position 66. Both amino acids are similar in size and have hydrophobic side chains, but tyrosine is slightly larger and polar which may be enough to affect tCRP3 17 and 18 ligand interaction compared to the rest of tCRP3. Tyrosine is also present in human, mouse, rat and hamster SAP at position 66 (Shrive *et al.*, 1996) suggesting that tCRP3 17 and 18 may have similar binding function at this residue to SAP rather than CRP. The *Limulus* CRP group is also split into two at this residue position as *Limulus* CRP 1.1 and 3.3 have a leucine, while *Limulus* CRP 1.4 has a serine. As leucine has a hydrophobic and non-polar side chain, like phenylalanine, it seems

that *Limulus* CRPs 1.1 and 3.3 have similar ligand binding to human CRP, whilst CRP 1.4 has similar ligand binding similar to *Limulus* SAP that has an amino acid (serine) with an uncharged polar side chain. As *Limulus* SAP does not bind PC (Shrive *et al.*, 1999), it is possible that proteins with serine at residue 66 also have altered ligand specificity, i.e. tCRP1, tCRP2 (5-7), *Limulus* CRP 1.4, and *Carcinoscorpius* CRP 1 group. This may be testable by mutagenesis studies using these proteins and their PC and PE binding.

Thr76 forms part of the hydrophobic pocket that PC sits in and shows higher conservation than Phe66 and Glu81 which are important PC interacting residues in human CRP (Thompson *et al.*, 1999). In human CRP, residue 76 is threonine, which is similar in side chain polarity to asparagine that is present in most of the horseshoe crab CRPs, and to glutamine in *Limulus* SAP. The tCRP3 group however, has an isoleucine at this residue position which has a non-polar hydrophobic side chain compared to the uncharged polar side chains of threonine, asparagine, and glutamine found in the other CRP and SAP sequences. It is possible that the isoleucine in tCRP3 at position 76 forms different ligand interactions than human CRP does.

Residue Ser74 is at the edge of the hydrophobic pocket furthest away from the calcium binding area and only has 60% conservation (see table 6.7) throughout the horseshoe crab CRP and SAP sequences, with no identical residues to serine possessed by human CRP. Most of the horseshoe crab CRPs have a leucine in this position, which has a hydrophobic side chain similar to phenylalanine which is present in tCRP2 sequences 5-7. Substitutions of serine (polar) with these hydrophobic amino acids may affect the PC binding site hydrophobicity and so affect ligand interactions. Contrary to the hydrophobic side chains of phenylalanine and leucine found in this position, tCRP3 and *Limulus* SAP have a glycine which has no side chain which may also affect ligand interaction compared to human CRP and other horseshoe crab CRPs. In human CRP, only the carbonyl of the main chain of serine is present in the PC binding site and does not form an interaction with PC

which may explain why this residue is not as conserved as others, such as Leu64, that have

more of their surface exposed in the PC binding site.

Two residues of the PC binding site only have 50% conservation throughout the

horseshoe crab pentraxins: 68 and 81 (see table 6.7). Glu81 in human CRP is a large part

of the pocket and forms an important interaction with the choline group of PC (Thompson

et al., 1999). Only in human CRP is residue 81 glutamic acid, which has a large polar

negatively charged side chain. The majority of sequences (*Carcinoscorpius* CRP group 1,

tCRP1 and tCRP3) have a threonine at position 81 which has a similar polar uncharged

side chain to serine that is present in *Limulus* SAP, suggesting a conservation in ligand

interaction function between these pentraxins. *Carcinoscorpius* CRP group 2, tCRP2 and

Limulus CRP 1.1, 1.4, and 3.3 all have proline at residue position 81 which is dissimilar to

threonine, serine, and glutamic acid suggesting that these pentraxins have different ligand

binding interactions by this residue to the other horseshoe crab CRPs and human CRP.

Ser68 is situated at the edge of the hydrophobic pocket and its side chain extends into

the ligand binding pocket in human CRP. Most horseshoe crab CRPs sequences have

asparagine which is similar in side chain charge and polarity to serine indicating similar

ligand binding interactions. tCRP3 and *Limulus* CRP 1.4 have glutamic acid at position 68,

tCRP2 sequences 5-7 and *Limulus* CRP 1.1 and 3.3 have an aspartic acid at this position

and *Limulus* SAP has a tyrosine. All these amino acids have different side chain charge and

size but similar polarity to serine found in human CRP which may suggest different ligand

binding pocket size, shape, or hydrophobicity which could alter ligand interactions.

Only in residues Asn61, Gly79, and Leu64 are the amino acids the same in human CRP

as they are in most of the horseshoe crab pentraxins (see table 6.7) suggesting that function

of these residues is conserved between horseshoe crab CRPs and SAP and human CRP.

Only tCRP2 sequence 7 has a different amino acid (tyrosine) to the other CRP and SAP

sequences at position 61 (asparagine), which could mean altered function at this residue, although both amino acids have polar side chains. Residue 79 is a glycine in all sequences apart from *Limulus* SAP where it is alanine, which is only slightly larger than glycine, and may affect the ligand binding site so that *Limulus* SAP does not bind PC. At residue 64, *Limulus* SAP has isoleucine similar to leucine in human CRP suggesting a similarity in binding site. tCRP2 sequences 5-7 also have a similar amino acid (alanine) to human CRP at position 64, suggesting that they may also have similar ligand binding to human CRP. Alanine has a smaller side chain to leucine which may protrude less into the ligand binding hydrophobic pocket leaving more space for a larger PC-like ligand side chain.

Horseshoe crab sequence groups according to conservation of sequence corresponding to the ligand binding site.

The tCRP3 group seems to be the most individual group in terms of sequence differences compared to the other horseshoe crab pentraxins. As previously mentioned, the tCRP3 group can show homology with both *Limulus* SAP and human CRP, indicating that the group can show similarity to both PC and PE ligand binding sites, although it has been shown that tCRP3 does not bind PE (Iwaki *et al.*, 1999). tCRP3 showed conservation with human CRP at residue 66 which was a phenylalanine, which is different to serine found in *Limulus* SAP. Similarity of the tCRP3 group with *Limulus* SAP is shown with the residue that corresponds to Ser74 in human CRP. This residue is glycine in tCRP3 and *Limulus* SAP which is different in side chain size and polarity to serine that is present in human CRP. However, in residues that correspond to positions 81 and 68, tCRP3 is similar to other horseshoe crab sequences such as tCRP1, *Carcinoscorpius* CRP 1 and *Limulus* SAP for position 81, and *Limulus* CRP 1.4 for position 68. tCRP1 and tCRP3 have no similarity in binding for PE or sialic acid, haemolytic activity, or haemagglutinating (Iwaki *et al.*, 1999). This suggests that residue 81 is not critical for these differences in activity, assuming activity proceeds via the PC binding site. Some sequences (5-7) of the tCRP2

group also seem to differ from the other horseshoe crab CRPs at the ligand binding site, although not to the extent that tCRP3 does. In residues corresponding to 64 and 74, tCRP2 (5-7) have an amino acid which is not shared by any of the other sequences, although some similarities in side chain hydrophobicity indicate similar residue function with those from the other horseshoe crab CRPs.

In the *Limulus* CRP sequences, *Limulus* CRP 1.4 is different from *Limulus* CRP 1.1 and 3.3 in three residues forming the ligand binding site. Previously mentioned was Phe66 where *Limulus* CRP 1.1 and 3.3 have a leucine and *Limulus* CRP 1.4 has a serine. Such a difference in side chain size, polarity and hydrophobicity compared to each other and phenylalanine in human CRP, could alter the ligand interaction in 1.4, whilst 1.1 and 3.3 would have reduced affects due to the similarity of hydrophobicity between leucine and phenylalanine. Another difference between the *Limulus* CRP 1.4 and 1.1/3.3 occurs at the residue at position 74 where *Limulus* CRP 1.4 has a histidine, whilst *Limulus* CRP 1.1 and 3.3 have leucine. Both these amino acids are different in side chain size, charge, and polar nature, indicating that ligand interaction at this site is different between them, and is also different to human CRP. The third position where they are different is residue 68, although the amino acids that the *Limulus* CRPs possess (glutamic acid and aspartic acid) have polar negatively charged side chains at this residue, and so perhaps ligand interaction would not vary too much within *Limulus*. Overall, it seems *Limulus* CRP 1.4 could possibly bind different ligands, or show variable specificity than *Limulus* CRP 1.1 and 3.3 due to differences in the PC binding site.

6.4.5 Conservation of putative human CRP C1q binding site residues throughout horseshoe crab pentraxins

The putative C1q binding site of human CRP is a cleft located on the opposite side of the protomer to the calcium and ligand binding sites (Shrive *et al.*, 1996). The sides of the cleft are composed of residues 5, 6, 203, 206, 187, 188, 160, 177, 175, 176, 95, and 112, whilst the bottom of the cleft is composed of residues 158, 38, 37, 94, and 112 (Thompson *et al.*, 1999) (see figures 6.23 and 6.24). Residues 114, 173, and 186 around the cleft have also been implicated in FcγR binding (Bang *et al.*, 2005).

Mutational studies indicated that Asp112 and Tyr175 are major determinants for CRP-C1q binding, as mutations at these residues from aspartic acid to alanine or asparagine in 112, and from tyrosine to alanine in 175 show C1q binding 5-10 fold less efficiently than non-mutated CRP (Agrawal *et al.*, 2001). This also led to a reduction in complement activation of at least 100 fold. Mutation at residue 38 from histidine to arginine, and at 158 from asparagine to alanine also showed reduction in C1q binding and complement activation, although it was thought that His38 and Asn158 contribute to the geometry of the C1q binding site rather than be major determinants for C1q binding. Mutation of Glu88 to glutamine and alanine slightly reduced C1q binding by CRP and hence reduced complement activation, and it was thought that Glu88 contacts C1q during binding and influences its conformational change. Mutation of Lys114 to alanine improves CRP-C1q binding by ~4.5 fold and increases complement activating efficiency because the positively charged side chain of lysine hinders C1q binding, but substitution with amino acids with negatively charged, neutral or non-polar side chains does not. Results from these C1q binding and C3 activation studies indicate that the pocket at the shallow end of the CRP cleft is the site of C1q binding.

Later mutational studies indicated that Asp169, Thr173, Tyr175 and Leu176 which are

all situated within the α-helix (169-176) at the C-terminus of the protein, influenced C1q binding to CRP as mutations at these sites reduced C1q binding. This study also confirmed that mutation of Lys114 to alanine increases C1q binding and complement activation (Bang *et al.*, 2005).

In human CRP the putative C1q cleft has an Aspartic acid (112) at the end situated closest to the pore of the pentamer, leading to a ring of negative charges (i.e. one from each protomer) lining the pore. It is proposed that CRP binds to disturbed cell membranes which show phospholipid heads or bacterial ligands such as LPS, and that due to the multiple ligands available, many CRP molecules bind close to each other with their calcium binding face next to the ligand, therefore with the putative C1q cleft exposed (Thompson *et al.*, 1999). It is possible that each C1q head group binds to a CRP pentamer which are adjacent to each other, via the clefts and/or central pore (Thompson *et al.*, 1999). The structure of the globular head of C1q has been found, and modelling of the structure with that of human CRP showed that the C1q head, which is predominantly basic in nature, can be accommodated by the negatively charged central pore of CRP (Gaboriaud *et al.*, 2003). The model shows that residue Asp112 from CRP subunit A, and Tyr175 from CRP subunit E are within hydrogen bond distance of Tyr175 from C1q subunit B, and Lys200 from C1q subunit A respectively. The model also indicates that the observed increase in C1q binding exhibited by mutation of Lys114 to Ala114 (Agrawal *et al.*, 2001), could be due to interference with a C1q lysine residue that points toward the negatively charged residues of the pore.

A comparison of the residues corresponding to the putative C1q cleft in human CRP (see figure 6.23 and 6.24) with the CRP and SAP like proteins from horseshoe crabs has been performed. Four residues in human CRP: 5, 6, 203, and 206, are part of the N-terminal and C-terminal end sequences which do not align with the horseshoe crab pentraxin sequences and therefore cannot be compared against anything to identify conservation. Therefore,

when comparing the structure of human CRP with the mapped sequences of horseshoe crab pentraxins, it cannot be assumed that the structure of the putative human CRP C1q binding site is the same. The C1q cleft is not visible in either human SAP (Emsley *et al.*, 1994) or *Limulus* SAP (Shrive *et al.*, 1999) suggesting that these proteins do not have similar equivalent interaction to human CRP involving this site. Also no C1q or C1q homologous proteins have been found as yet in the horseshoe crabs, although there are homologues to other complement proteins in *Carcinoscorpius* (Zhu *et al.*, 2005), *Tachypleus* (Ozaki *et al.*, 2005) and *Limulus* (Armstrong *et al.*, 1998).
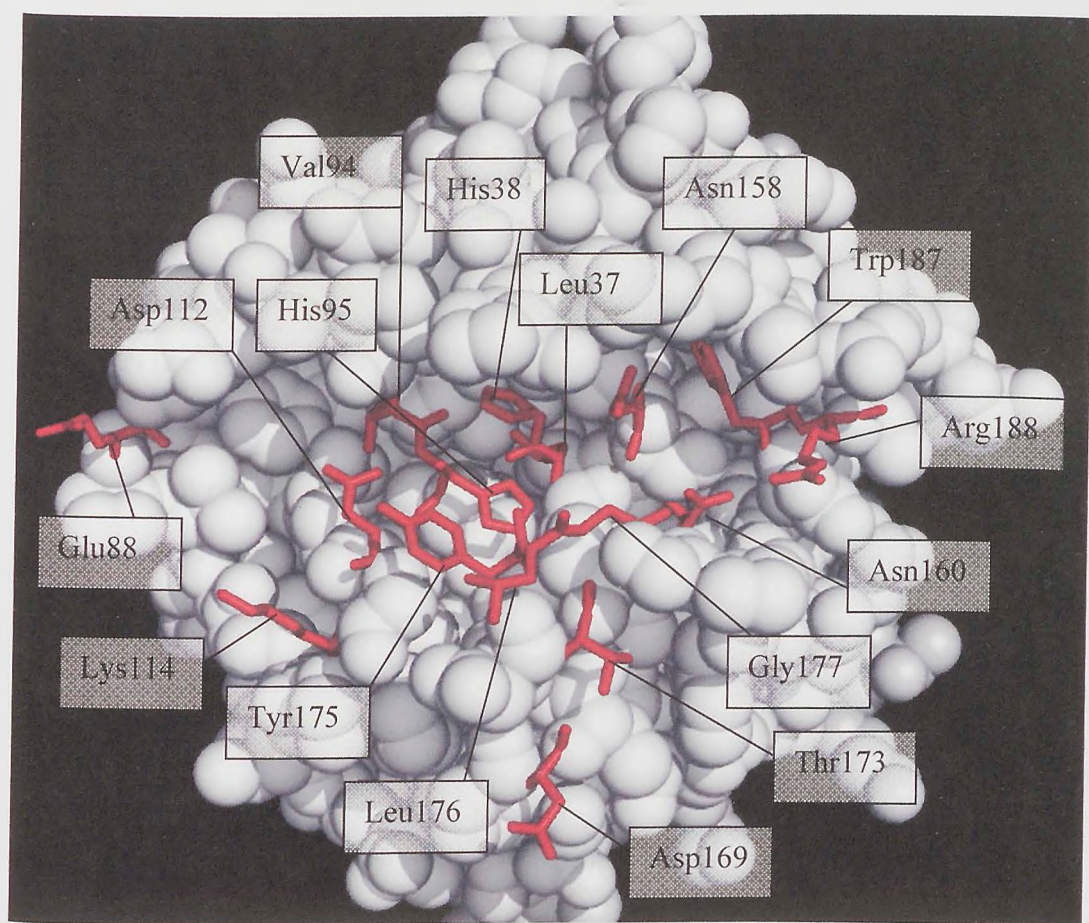
Figure 6.23 Human CRP protomer chain A showing putative human C1q binding cleft residues in red. The C1q cleft is running from left (central pore) to right (outer rim).
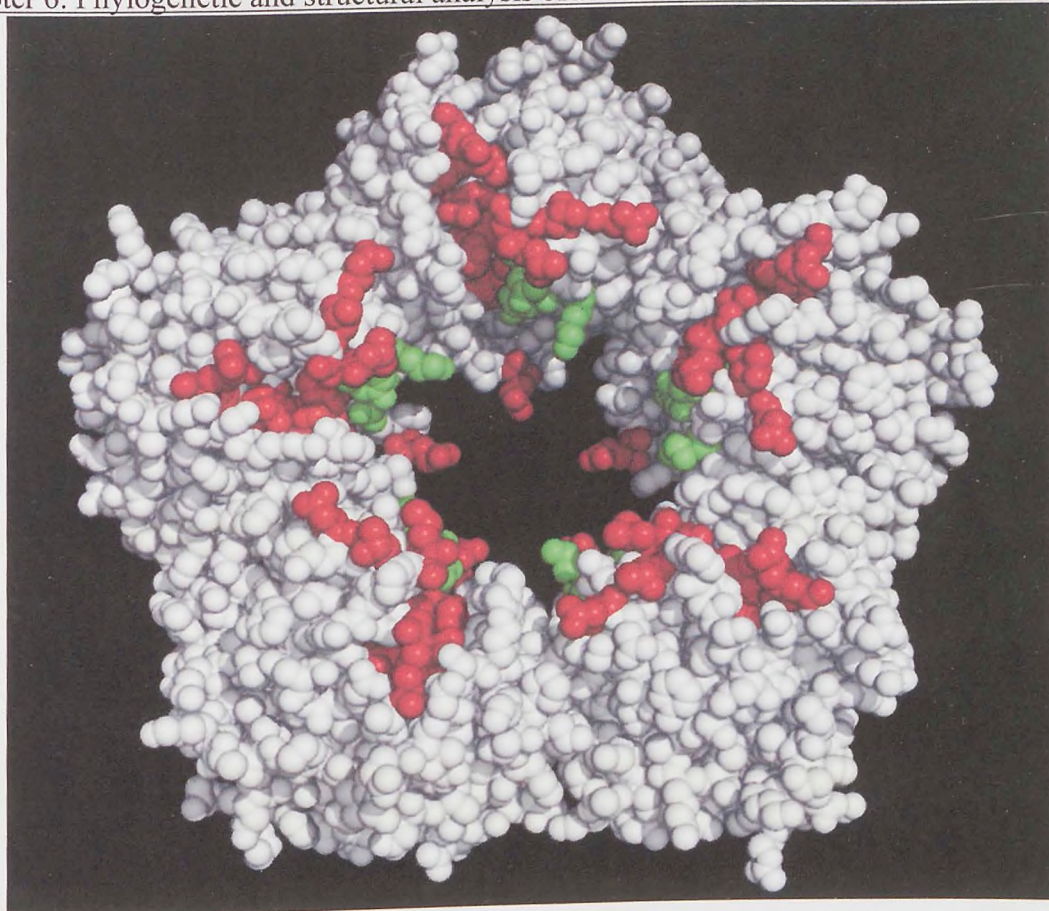
Figure 6.24 Human CRP pentamer chain A showing the residues of the putative human C1q binding cleft in red, and the most important residues 112, 114, and 175 are highlighted in green.

Residues that correspond to the putative C1q binding site residues in human CRP show high conservation (see table 6.8). His95 is totally conserved throughout the horseshoe crab pentraxins and human CRP, as it is part of the pentraxin signature sequence which may or may not be of functional significance. A majority of the other residues have 90% conservation (37, 38, 158, 175, 177, 187, 188, 160) with the main difference being between the human CRP and the horseshoe crab pentraxins, rather than between horseshoe crab pentraxins, although *Limulus* SAP also shows some differences. As the majority of amino acid differences are between human CRP and the horseshoe crab pentraxins for most of the putative C1q binding site, and there have been no C1q found in horseshoe crabs, there must be a different interaction of this area in the horseshoe crab pentraxins compared to human CRP but conservation within horseshoe crab pentraxins. As horseshoe crabs lack C1q, the conservation of residues at this site may infer similar or identical binding to protein(s) found only in horseshoe crabs. There may also be a difference in

structure and/or function between *Limulus* SAP and the horseshoe crab CRPs at this

putative binding site.

Table 6.8 Conservation of residues corresponding to the human CRP putative C1q binding site between horseshoe crab pentraxins and human CRP. Conservation is shown as the percentage of residues identical to the most common amino acid. The most important residues in C1q interaction are highlighted.

hCRP number	hCRP residue	Most common residue in horseshoe crab pentraxins and human CRP	Conservation
95	His	His	100%
37	Leu	Cys	90%
38	His	Tyr	90%
158	Asn	Glu	90%
175	Tyr	Ser	90%
177	Gly	Cys	90%
187	Trp	Gly	90%
188	Arg	Asn	90%
160	Asn	Asn	90%
94	Val	His	80%
173	Thr	His	80%
112	Asp	Asp	70%
114	Lys	Phe	70%
176	Leu	Lys	70%
169	Asp	Glu	60%
88	Glu	Pro	50%

Residues that show lower conservation are 94, 173, 176, and the important C1q interacting residues Asp112 and Lys114 (see table 6.8). Again, in most of these residues, human CRP does not have the same amino acid as the horseshoe crab pentraxins. Residues at position 94 varied not only from human CRP, which has a valine, but also within the horseshoe crab pentraxins as most possess a histidine, but tCRP2 sequences 5-7 and *Limulus* SAP possess a tyrosine. Valine in human CRP, and tyrosine in tCRP2 5-7 and *Limulus* SAP have some similarity in side chain charge but different polarities, indicating possible conservation of equivalent interaction to human CRP C1q binding at this residue. However, histidine varies in side chain charge and polarity from valine in human CRP, indicating no conservation of function of this residue.

A majority of the horseshoe crab CRPs at position 176 possess a lysine, which has a positively charged polar side chain compared to leucine (non-polar, neutral, hydrophobic side chain) found in human CRP, asparagine (polar, neutral side chain) found in tCRP1

sequence 15 and tCRP3 sequences 16-18, glutamic acid (negatively charged polar side chain) found in *Carcinoscorpius* CRP 1, and serine (polar side chain) found in *Limulus* SAP. As these residues vary in side chain polarity, charge and hydrophobicity, it is likely that they do not possess the same function within the horseshoe crab pentraxins. Position 173 shows conservation at 80% and in human CRP and *Limulus* SAP is threonine residue that has a polar uncharged side chain, but in a majority of the horseshoe crab CRP sequences is a histidine, which has a positively charged polar side chain. *Limulus* CRP 1.1 and 3.3 have a tyrosine at position 173 that has a large hydrophobic polar side chain dissimilar to histidine in *Limulus* CRP 1.1, suggesting a difference in function of this residue within species and also between species.

In human CRP, Asp112 has been shown to be important in C1q binding by mutational studies which indicate that its function is through its negative charge (Agrawal *et al.*, 2001). As most of the horseshoe crab CRP also have aspartic acid at this residue, it is likely to be of similar equivalent interaction in horseshoe crab CRPs. However, *Limulus* SAP and tCRP3 possess an asparagine at position 112 which has an uncharged side chain, suggesting a similarity in interactions between *Limulus* SAP and the tCRP3 group at this residue. Lys114 in human CRP is thought to hinder C1q binding due to its large positive side chain as neutral or non-polar side chains give enhanced C1q binding and complement activation (Agrawal *et al.*, 2001). Position 114 shows 70% conservation, with most horseshoe crab pentraxin sequences having phenylalanine which has a non-polar hydrophobic side chain. tCRP3 possess tyrosine at this residue position which also has a hydrophobic side chain but is polar, whilst *Limulus* SAP has serine at this position which has an uncharged polar side chain.

Position 169 has only 60% conservation. Human CRP has aspartic acid at position 169 and *Limulus* SAP has histidine, which both have negative and positive side chains respectively but are both polar. A majority of the horseshoe crab CRP sequences have

glutamic acid at position 169, which also has a polar negative side chain, and so conservation of equivalent interactions to the putative human CRP C1q interaction could occur at this site. However, tCRP3 sequences 19-22 have lysine at position 169 which show similarities to *Limulus* SAP which has histidine as they both have large polar positive side chains. Other horseshoe crab pentraxins tCRP1 sequence 15 and tCRP3 sequences 16-18, and tCRP1 sequence 11 have asparagine and glycine respectively. Neither of these residues are alike to those in human CRP or *Limulus* SAP and so it is likely that function of these residues is different. It also indicates difference of residue function within the tCRP1 and 3 groups.

Glu88 is thought to influence conformational change in C1q as mutations in this amino acid affect complement activation (Agrawal *et al.*, 2001). Position 88 has only 50% conservation indicating its function is not that important in the horseshoe crab pentraxins. At position 88 the tCRP2 group is split into two: tCRP2 sequences 1-4 have proline, and tCRP2 sequences 5-7 have lysine. tCRP2 sequences 1-4 show amino acid conservation with tCRP3, all *Limulus* pentraxins, and *Carcinoscorpius* CRP 2, indicating similar binding, but tCRP2 (5-7) are not similar to any other pentraxins at this residue. Another *Tachypleus* group tCRP1, has the same amino acid (asparagine) at position 88 as the *Carcinoscorpius* CRP 1 group. Together, these differences in amino acids indicate that tCRP3 and some of tCRP2 have similar function at residue 88, but that the other tCRP2 proteins and tCRP1 have dissimilar interactions. This could be an indication of different structure and/or functions within horseshoe crab CRP at this site. No horseshoe crab pentraxins had a similar amino acid to human CRP (glutamic acid) at position 88.

Overall, the horseshoe crab CRPs may not have a structure similar to that of the putative C1q binding cleft in human CRP because of the absence of residues corresponding to one side of the cleft, and also the absence of C1q in horseshoe crabs. The high conservation of horseshoe crab pentraxins at residues corresponding to the human CRP putative C1q

binding site, indicates that there may be a binding site there, although whether this is for a protein or proteins in horseshoe crabs with a similar interaction to human CRP-C1q is unknown. Small variations in the horseshoe crab CRP and SAP sequences could change the specificity at this possible binding site within the individual horseshoe crab groups such as tCRP2, but also between groups such as *Carcinoscorpius* CRP 1 and 2 groups, and between species such as *Limulus* and *Tachypleus*. Interesting is the high conservation of the horseshoe crab pentraxins with human CRP at position 112, as this residue has been determined to be important in human CRP-C1q binding. Although it is unlikely the horseshoe crab CRPs have a C1q binding cleft as they do not possess C1q, this does not necessarily mean that they do not have a binding site which performs an equivalent function of interacting with other immune proteins to clear pathogens from the horseshoe crab plasma. Perhaps differences at this possible binding site in the *Tachypleus* pentraxins cause some of the different properties in the CRP groups seen by Iwaki *et al.*, (1999).

6.4.6 Other areas of conservation between the horseshoe crab CRP and SAP

ConSurf and Protein Explorer were used to determine whether there were any other areas of sequence conservation between the horseshoe crab CRP and SAP sequences that were not in the areas of sequence corresponding to the calcium, ligand and C1q binding sites in human CRP.

Using ConSurf, a few areas of conservation mapped onto the human CRP chain A structure can be identified. The first area of high conservation appears away from both the calcium binding area and the C1q cleft (see figure 6.25A). The residues of high conservation are Thr27, Pro29, Glu101, Ser102, Ala129, Gly124, Thr126, Glu130, and Ile133. The second area of high conservation is situated adjacent to where the putative C1q cleft is (see figure 6.25B) and is composed of the residues Ala8, Phe11, Pro12, Ser18, Thr41, Val91, Gly154, Asp155 and Lys201.
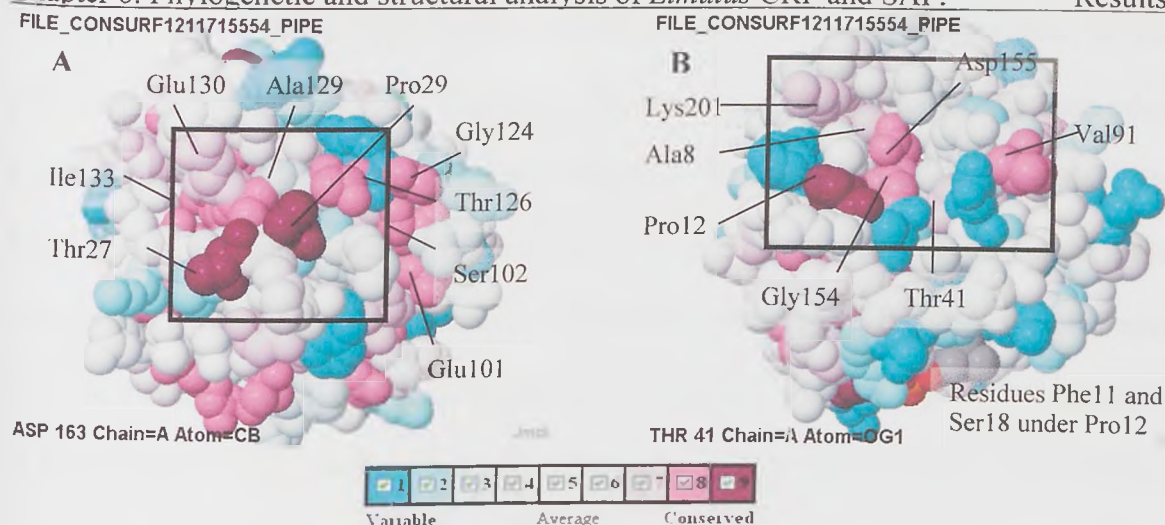
Figure 6.25. Horseshoe crab CRP and SAP sequences mapped onto human CRP chain A using ConSurf. **A** shows the first identified area of conservation highlighted by the box. **B** shows the second identified area of conservation highlighted by the box. Colour guide of homology between sequences is shown below the figures as ConSurf colours of shades 1-9.

Some of the amino acids surrounding these residues are variable suggesting that if these amino acids were part of a conserved binding site in the horseshoe crab CRP and SAP, the structure may be different to that of human CRP.

The residues in these areas are also shown as having conservation in Protein Explorer (see figure 6.26 A-D. At area 1 designated by ConSurf analysis (see figure 6.26 A and B) residues Pro29, Thr27, Thr126, Gly124, and Ser102 were 100% conserved between the horseshoe crab CRP and SAP sequences and human CRP. The remaining residues corresponding to this area (Ala129, Glu101, Glu130, Ile133) showed 90% conservation between the horseshoe crab CRP and SAP sequences and human CRP. At area 2 designated by ConSurf analysis (see figure 6.26 C and D) residues Pro12, Gly154, Asp155, Val91 and Phe11 were 100% conserved between the horseshoe crab pentraxin sequences and human CRP. The remaining residues (Ser18, Lys201, Ala8, Thr41) were 90% conserved between the horseshoe crab CRP and SAP sequences. This high level of conservation agrees with that shown by ConSurf and indicates that these areas may be important in horseshoe crab CRP and SAP function or stability.

Observation of the position of the two highlighted areas of conservation in the pentameric structure of human CRP, shows that these areas are situated on the outer edge of the protomer, and between protomers (see figure 6.27). Should the horseshoe crab CRPs have a similar protomer arrangement to human CRP, it is possible that the residues of these areas interact with other CRP pentamers or ligands on the edge of the pentamer. It is also possible that they play a part in interprotomer interactions, which is especially noticeable in figure 6.27 B where residues from conserved area one appear to interact with residues from conserved area two. These interactions depend on the protomer arrangement of the horseshoe crab pentraxins, and should they have a different cyclic arrangement or protomer topology, then these areas of homology may be positioned elsewhere, and may form different interactions to those suggested.
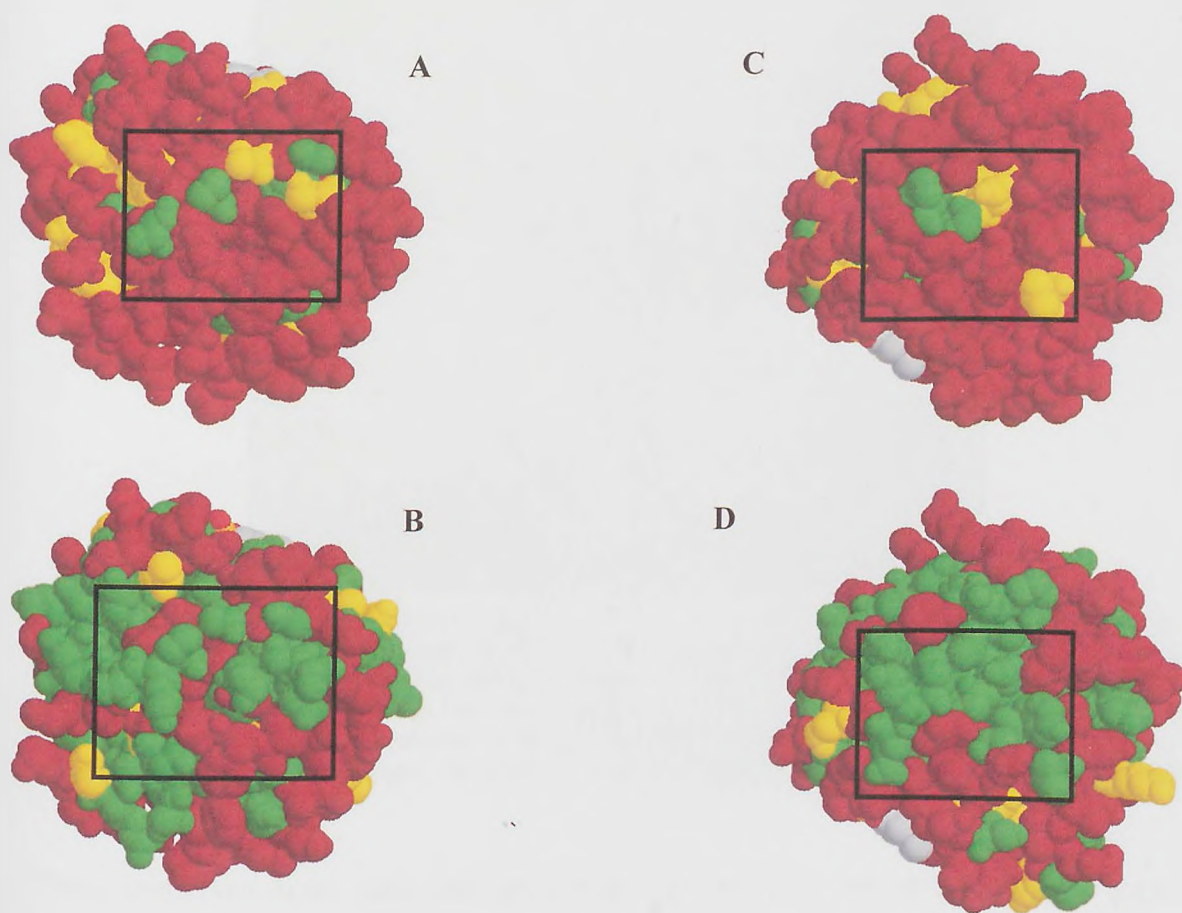
Figure 6.26. Human CRP protomer chain A showing areas of conservation identified initially using ConSurf between horseshoe crab CRP sequences at areas that do not correspond to the calcium, ligand or C1q cleft, as highlighted by the box. **A** shows the first conserved area identified initially in ConSurf at 100% consensus percentage. **B** shows the first conserved area identified initially in ConSurf at 90% consensus percentage. **C** shows the second conserved area identified initially in ConSurf at 100% consensus percentage. **D** shows the second conserved area identified in ConSurf at 90% consensus percentage. Green represents identical residues, yellow represents similar residues, and red represents different residues.

A
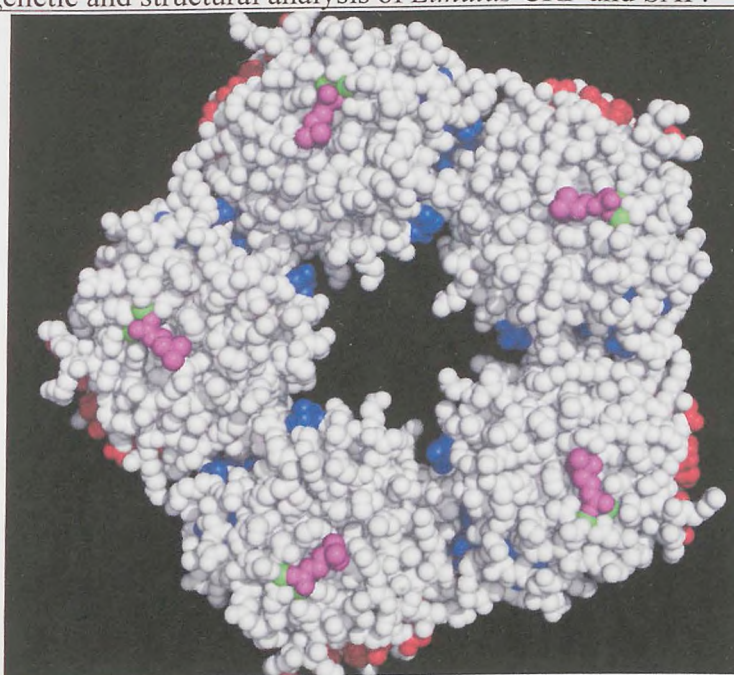
B

Figure 6.27. Areas of new sequence homology between the horseshoe crab CRP and SAP sequences that do not correspond to the calcium, ligand or C1q binding areas, mapped onto the pentameric structure of human CRP. The red amino acids correspond to those identified as conserved using ConSurf for area 1, blue amino acids correspond to those identified as conserved using ConSurf for area 2. Calcium ions are shown in green and phosphocholine in magenta. **A** shows the human CRP calcium binding face. **B** shows the human CRP α-helix face. The yellow box highlights the apparent interaction between residues from conserved area 1 and 2.

These conserved areas were also projected onto the structure of the *Limulus* SAP hexadecamer (unpublished) as shown in figure 6.28. The two areas corresponding to the areas of high conservation between the horseshoe crab CRPs and SAP, that appear on the human CRP protomer when the horseshoe crab sequences are modelled onto it, are mainly on the top surface area of the *Limulus* SAP hexadecamer and within the pore. However,

there appear to be some conserved amino acids on the outside surface of the hexadecamer near to the PE binding site, and also buried beneath the surface. It is possible that the conserved areas on the surface have some function in hexadecamer-hexadecamer interaction, agglutination, or binding ligands or other proteins such as *Limulus* SAP, *Limulus* CRP or other proteins similar to CrOctin binding in *Carcinoscorpius rotundicauda* (Li *et al.*, 2007) which participate in the immune response. This may be conserved in CRPs should they have a similar protomer arrangement to *Limulus* SAP.
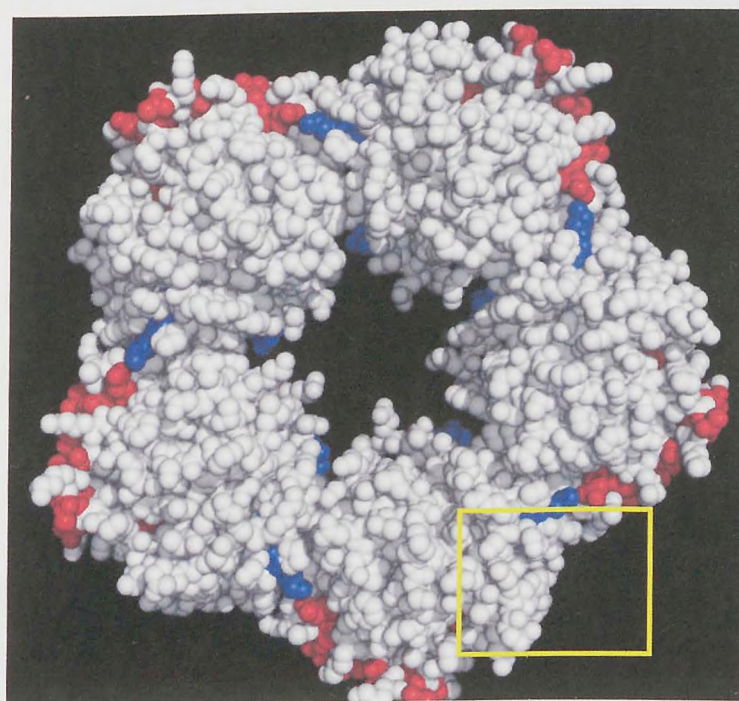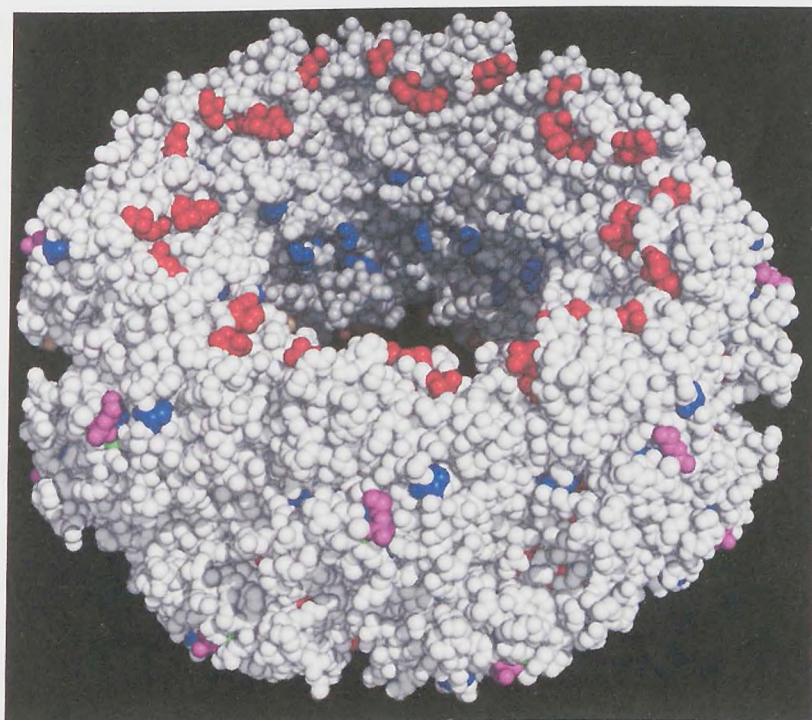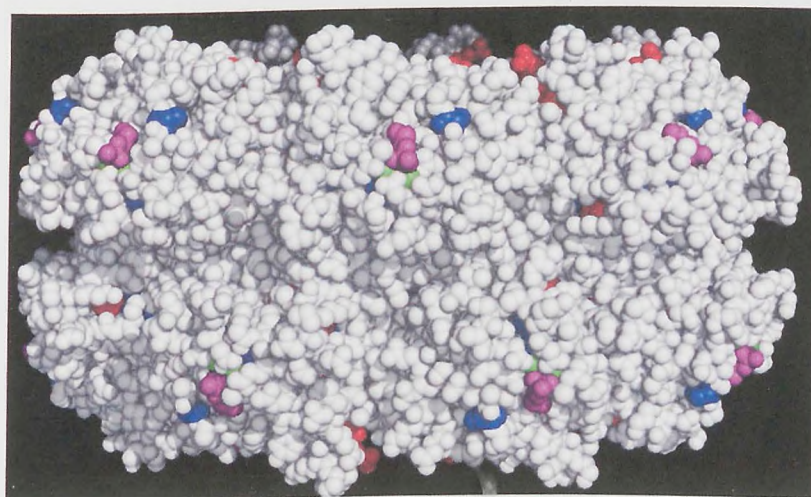

Figure 6.28. Areas of new sequence conservation between the horseshoe crab CRP and SAP sequences that do not correspond to the human calcium, ligand of C1q binding areas, mapped onto the hexadecameric structure of *Limulus* SAP. The red amino acids correspond to those identified as conserved using ConSurf for area 1, blue amino acids correspond to those identified as conserved using ConSurf for area 2. Calcium ions are shown in green and PE in magenta. **A** shows the *Limulus* SAP pore. **B** shows the *Limulus* SAP PE binding site.

6.5 Summary and conclusions

Phylogenetic analysis indicates that horseshoe crab CRPs and SAP are ancient paralogues of each other as they have sequence similarity but perhaps different functions. CRPs appear as orthologues as they have very similar sequence and possible function and appear in different organisms, but could also be recent paralogues within species as they could have slightly different binding properties. The small amount of diversity between horseshoe crab CRP sequences was shown not to be due to splice variants because of the presence of single amino acid mutations rather than changes in stretches of sequences.

Preliminary analysis of horseshoe crab CRPs and SAP using a simple phylogenetic tree indicate that there are subgroups of like sequences within the *Tachypleus* and *Carcinoscorpius* species, but not within the *Limulus* species as there is too few to group. Such subgroups in the *Tachypleus* species agree with the groups identified by Iwaki *et al.*, (1999) determined by properties such as sialic acid and PE binding. The *Limulus* CRP pentraxins do not appear to have great variation between them compared to the CRP pentraxins in *Carcinoscorpius* and *Tachypleus*. However, there is no reason not to assume that *Limulus* has just as many, if not more CRP variants as *Tachypleus* does, and that if such variants were found, they would also show sequence similarity in groups according to ligand binding. Indeed, if such a number of sequences were to be present within all three horseshoe crabs, it is possible that they would show sequence similarities for certain subgroups which may be indicative of function. An example would be a third CRP group for the *Carcinoscorpius* species which is most homologous to *Tachypleus* tCRP3 group, as *Carcinoscorpius* CRP 1 group is most homologous to *Tachypleus* tCRP1, and *Carcinoscorpius* CRP 2 group is most homologous to *Tachypleus* tCRP2 and the *Limulus* CRP pentraxins. Likewise, new CRP sequences found in *Limulus* might be similar to the *Tachypleus* tCRP1 and tCRP3 groups.

The comparison of calcium and PC binding site residues between human CRP and horseshoe crab CRPs and SAP shows which core residues have or have not been conserved between these proteins, and what this might imply for function. As sequence variation within the *Limulus* CRP pentraxins is fairly low compared to the variation in *Tachypleus* CRP pentraxins, it is difficult to suggest areas of possible ligand binding that would vary within the *Limulus* CRP pentraxins. What can be seen though is the high conservation at the possible calcium binding sites determined by alignment of human CRP sequence against the horseshoe crab CRP and SAP sequences, and also the high conservation at the putative C1q binding site. However, this area is unlikely to have C1q binding function in the horseshoe crab CRP and SAP sequences due to lack of C1q-like proteins, lack of C1q binding cleft in *Limulus* SAP (Shrive *et al.*, 1999), and also function of Limulin as a cell-lysis protein (Armstrong *et al.*, 1996) which is the role of C1q complement cascade. High conservation in this area however, suggests similar function between the horseshoe crab pentraxins, and visualisation of this area on *Limulus* SAP structure shows no cleft but clustering of the amino acids which may suggest a possible binding site. Interestingly, the ligand binding area is the least conserved of all three areas analysed (i.e. calcium, C1q, ligand), suggesting variation in ligand binding properties and perhaps function. Most importantly, phenylalanine 66 and glutamic acid 81 which play major roles in PC binding in human CRP (Thompson *et al.*, 1999), and are two of the most variable amino acids in the horseshoe crab CRP and SAP ligand binding sites, indicating that they have different ligand binding properties to human CRP.

When other areas of high conservation that do not correspond to the calcium binding site, PC binding site, and the C1q binding site are found from mapping of the horseshoe crab CRPs and SAP onto a human CRP protomer, and subsequently mapped onto the pentameric structure of human CRP and the hexadecameric structure of *Limulus* SAP, they show different results. When mapped onto the human CRP pentamer, the areas of high conservation are on the outer edge of the CRP pentamer and between protomers. If

horseshoe crab CRPs have a similar protomer arrangement to human CRP this suggests

conservation of protomer-protomer interaction and also perhaps conservation of

interactions on outside edge of pentamer such as pentamer-pentamer interaction or

ligand/other protein binding. However, when mapped onto the *Limulus* SAP hexadecamer,

the areas of high conservation correspond mainly to areas of surface on the top face of the

octamer, i.e. the face not interacting with the other octamer of the hexadecamer, and also

areas within the pore. Therefore, if *Limulus* CRPs have protomer arrangement similar to

human CRP, these areas of high conservation will be in a different location than if they

have protomer arrangement similar to *Limulus* SAP.

Therefore, CRP and SAP pentraxins within the horseshoe crabs show such sequence

similarity that it can be said that they have similar calcium binding properties allowing

them to bind ligands through calcium like human CRP binds phosphocholine. Variation

within the ligand binding site suggests that either they can bind a variety of ligands but

have a similar overall function, or bind similar ligands but with different overall function.

It is possible that once the ligands are bound, the horseshoe crab CRPs and SAP function

in a similar way, perhaps through the highly conserved area corresponding to C1q-binding

in human CRP, although different to C1q binding because of absence of C1q in horseshoe

crabs.

To further analyse conservation of key calcium and ligand binding residues throughout

the horseshoe crab CRPs and SAP, it would be possible to compare them throughout all the

CRP and SAP-like pentraxins, especially in the mammals where the functions of CRP and

SAP-like proteins are more characterised than those in invertebrates. It may be possible to

identify residues in horseshoe crab CRPs and SAP that have higher similarity with all

mammalian CRP or SAP-like proteins respectively, therefore allowing a more detailed

analysis.

Further work could also include the analysis of conservation of residues throughout the horseshoe crab CRPs and SAP that correspond to the ligand binding residues in *Limulus* SAP (unpublished data), and those from human SAP that interact with dAMP (Hohenester *et al.*, 1997) now that it is known that *Limulus* CRP and *Limulus* SAP interact with AMP and ribose-5-phosphate (see chapters 3 and 4). These further studies may shed more light on *Limulus* CRP and SAP interaction with ribose-5-phosphate and AMP, and may also provide data to allow suggestion of other ligands to *Limulus* CRP and SAP.

Chapter 7. Conclusions and future work.

7.1 *Limulus* CRP

Limulus CRP has previously been shown to comprise a family of heterogeneous proteins, of which there are at least three genes (Nguyen *et al.*, 1986a), and three subunit forms (Nguyen *et al.*, 1986b; Tennent *et al.*, 1993). Isolation and characterisation of CRP-like proteins from *Tachypleus tridentatus* and *Carcinoscorpius rotundicauda* showed the presence of more than three CRP genes (Iwaki *et al.*, 1999; Ng *et al.*, 2004). Phylogenetic analysis showed that *Limulus* CRPs have homology with *Tachypleus* CRP2 and *Carcinoscorpius* CRP2, so there is a possibility of other *Limulus* CRP sequences which show homology to *Tachypleus* CRP1 / *Carcinoscorpius* CRP1, and *Tachypleus* CRP3. Therefore, it is likely that *Limulus polyphemus* possesses more than three CRP genes, which may result in several molecular aggregate forms. The CRP-like proteins in *Tachypleus tridentatus* were shown to have different binding properties and activities (Iwaki *et al.*, 1999), and so it is also possible that *Limulus* CRP have different binding properties.

This study verifies that whole *Limulus* CRP isolated from *Limulus* plasma is composed of three different subunits as determined by SDS-PAGE, but also shows novel data that there are at least four CRP molecular aggregates that differ in surface hydrophobicity. The changes in surface hydrophobicity indicate either variable post-translational modification, or changes in amino acid sequence composition between *Limulus* CRP molecular aggregate forms, or a mixture of both. If CRP molecular aggregates are composed entirely of like-subunits, then there would be expected at least three molecular aggregate forms that vary in surface hydrophobicity as there are at least three genes (Nguyen *et al.*, 1986a) and three subunit forms (Nguyen *et al.*, 1986b; Tennent *et al.*, 1993). The presence of at least four CRP molecular aggregates that vary in surface hydrophobicity indicates possibly more

347

than three CRP genes, which would not be unexpected due to the presence of many CRP genes in horseshoe crabs *Tachypleus tridentatus* and *Carcinoscorpius rotundicauda* (Iwaki *et al.*, 1999; Ng *et al.*, 2004). It is also possible that *Limulus* CRP molecular aggregate forms are composed of different subunit forms. This may explain the presence of more than three molecular aggregate forms that vary in surface hydrophobicity as being due to multiple subunit form combinations.

The presence of minor variants of surface hydrophobicity being attributed to trace contamination from other plasma proteins such as haemocyanin, α_2-macroglobulin, or *Limulus* SAP, cannot be ruled out. A reduction in the trace contamination could be achieved by the use of PEG-cut plasma which would reduce the amount of haemocyanin present in the CRP fraction, and also by multiple re-application and purification of CRP during chromatography. A determination of the molecular aggregate forms that vary in surface hydrophobicity is currently underway by collection of fractions during FPLC, and analysis using SDS-PAGE and N-terminal sequencing.

Human CRP interaction with PC had previously been shown to be inhibited by PE, 2-deoxy-ribose-5-phosphate, galactose-6-phosphate, ribose-5-phosphate, glucose-6-phosphate, adenosine-5-monophosphate (AMP), and mannose-6-phosphate among other phosphorylated sugars (Lee *et al.*, 2002). As *Limulus* CRP interacts with PC and PE (Robey and Liu 1981; Quigley *et al.*, 1994) in a calcium dependent manner, it was thought that it may interact with other human CRP ligands, and that this may isolate different molecular aggregate forms.

Whole *Limulus* CRP that had been isolated from *Limulus* plasma using a 10mM PC elution, was re-eluted from a PE-agarose column using ribose-5-phosphate and AMP. This indicates that these molecules interact with *Limulus* CRP at the PE-binding site or near to it, thereby disrupting the interaction of CRP with the PE-agarose column. It is likely that

348

these ligands interact with *Limulus* CRP at the PE-binding site through the calcium ion-phosphate interaction, as this is how human CRP interacts with its phosphorylated ligands (Gotschlich and Edelman 1967; Thompson *et al.*, 1999). This is the first time that *Limulus* CRP has been shown to interact with these ligands, and suggests a conservation of binding properties and perhaps functions of CRP involving these ligands, from invertebrates to humans.

The proportion of *Limulus* CRP eluted with AMP was slightly higher than the proportion eluted with ribose-5-phosphate, but both eluted the majority of CRP at ligand concentrations of 5mM, 10mM and 30mM. Isocratic and gradient elutions using these ligands failed to isolate an individual molecular aggregate form of CRP during affinity chromatography, but did show some separation of molecular aggregate forms when the surface hydrophobicity was analysed. *Limulus* CRP that was eluted using ribose-5-phosphate showed the presence of two molecular aggregate forms that vary in surface hydrophobicity, compared to three molecular aggregate forms in CRP eluted with AMP, and the four molecular aggregate forms in whole CRP. This shows some specificity of *Limulus* CRP molecular aggregate forms for ribose-5-phosphate and AMP. Characterisation of the molecular aggregate CRP forms separated by ribose-5-phosphate and AMP is underway.

It is possible that the molecular aggregate CRP forms vary through glycosylation, and deglycosylation of the CRP before surface hydrophobicity analysis would indicate the extent to which this occurs. As *Limulus* CRP has shown similarity in affinity to ribose-5-phosphate and AMP which have been shown to inhibit human CRP-PC interaction (Lee *et al.*, 2002), there is no reason to not assume that other possible human CRP ligands such as mannose-6-phosphate, would also interact with *Limulus* CRP. Such ligands could be used in further attempts to isolate molecular aggregate forms of *Limulus* CRP that have different binding properties.

7.2. *Limulus* SAP

Unlike *Limulus* CRP, there has only been reported a single gene for *Limulus* SAP (Tharia *et al.*, 2002). However, there appears to be multiple subunit forms, and possible multiple sequences found in peptide fragments (Shrive *et al.*, 1999; Tharia *et al.*, 2002). CrOctin, a protein in *Carcinoscorpius rotundicauda* with some homology to *Limulus* SAP, has several different gene sequences (Li *et al.*, 2007). It is possible that *Limulus* SAP also comprises a heterogenous family of proteins, like horseshoe crab CRP.

This study verified that whole *Limulus* SAP which was isolated from *Limulus* plasma is composed of multiple subunits, in particular three were identified by SDS-PAGE. This study also shows the novel finding that whole *Limulus* SAP is composed of four molecular aggregate forms that vary in surface hydrophobicity. As only one gene has been reported for *Limulus* SAP (Tharia *et al.*, 2002) it was possible that the differences in surface hydrophobicity were due to different post-translational modifications such as glycosylation which occurs at Asn81 and possibly Asn118 (Tharia *et al.*, 2002). Deglycosylation of *Limulus* SAP did not reduce the number of subunits as identified by SDS-PAGE, but did reduce the size of two subunits. Deglycosylation of *Limulus* SAP also reduced the number of molecular aggregate forms that vary in surface hydrophobicity to three, as compared to four in whole SAP. This indicated that one of the four molecular aggregate forms that vary in surface hydrophobicity was glycosylated, and that this probably accounted for its different surface hydrophobicity to other molecular aggregate forms. The presence of at least three molecular aggregate forms that vary in surface hydrophobicity after deglycosylation, suggests that variations in surface hydrophobicity occurs in other ways perhaps through variation in sequences which would indicate a number of SAP genes, or perhaps through other post-translational modifications. There is also a possibility of trace contamination within the SAP samples used for surface hydrophobicity analysis, such as haemocyanin, *Limulus* CRP, α_2-macroglobulin, or other plasma proteins. SDS-PAGE and

N-terminal sequencing of isolated molecular aggregate forms of *Limulus* SAP that vary in surface hydrophobicity to determine protein composition is underway.

Previous analysis of *Limulus* CRP has shown that it interacts with ribose-5-phosphate and AMP. Human SAP has been shown to bind not only phosphorylated sugars (Loveless *et al.*, 1992), but also dAMP (Hohenester *et al.*, 1997). Therefore, it was thought that perhaps *Limulus* SAP interacted with ribose-5-phosphate and AMP in a similar way to *Limulus* CRP and human SAP. Affinity chromatography elutions of *Limulus* SAP from a PE-agarose column, showed that both ribose-5-phosphate and AMP bind at or near to the PE-binding site on *Limulus* SAP, thereby displacing it from the PE-agarose column. The structure of *Limulus* SAP with ligand bound (unpublished data) suggests interaction of phosphate groups on the ligand with the calcium ions on SAP. It is likely that this is the mechanism by which *Limulus* SAP interacts with ribose-5-phosphate and AMP. This is the first time that *Limulus* SAP has been shown to interact with these ligands, and suggests a conservation of binding properties and perhaps function with *Limulus* CRP, human CRP, and human SAP.

The concentration of ribose-5-phosphate and AMP needed to elute the majority of *Limulus* SAP from the PE-agarose column was higher than that needed to elute the majority of *Limulus* CRP. This indicates that either *Limulus* SAP has a lower binding affinity for ribose-5-phosphate and AMP than *Limulus* CRP does, or that *Limulus* SAP has a higher affinity for the PE-agarose column than *Limulus* CRP does, or both. At 10mM concentration, both ribose-5-phosphate and AMP only elute a small amount of *Limulus* SAP during affinity chromatography, but at 30mM concentration they eluted the majority of SAP. This was unexpected for such a small change in ligand concentration, as the equivalent change of ligand concentration for elution of *Limulus* CRP did not give such a large difference in elution proportion.

Limulus SAP that was eluted with ribose-5-phosphate from a PE-agarose column was composed of the four molecular aggregate forms that vary in surface hydrophobicity, as seen in whole *Limulus* SAP. However, *Limulus* SAP that was eluted with AMP was composed of three molecular aggregate forms that vary in surface hydrophobicity, not four. This indicated that elution of *Limulus* SAP with AMP achieved some separation of molecular aggregate forms whilst elution with ribose-5-phosphate did not. This would indicate that there were molecular aggregate forms of *Limulus* SAP with different binding affinities for ribose-5-phosphate and AMP. The characterisation of the molecular aggregate forms of *Limulus* SAP that vary in surface hydrophobicity is currently underway.

As *Limulus* SAP interacts with ribose-5-phosphate and AMP that are possible human CRP ligands (Lee *et al.*, 2002), further attempts at isolating *Limulus* SAP molecular aggregate forms through different binding affinities may be achieved by using other human CRP ligands such as 6' phosphorylated sugars (Kottgen *et al.*, 1992; Culley *et al.*, 2000; Lee *et al.*, 2002; Lee and Lee 2003). Human SAP has been shown to bind with sugars phosphorylated at position 1 and 6, and also those sulphated at position 3 (Loveless *et al.*, 1992). Using ligands with these properties during affinity chromatography of *Limulus* SAP may also achieve some separation of molecular aggregate forms.

7.3 *Limulus* CRP and SAP interaction with ribose-5-phosphate and AMP

Human SAP has been shown to interact with dAMP, and the crystal structure has been solved (Hohenster *et al.*, 1997). It was shown that the phosphate group of dAMP bound to the calcium ions on human SAP, and there were also interactions of SAP with the deoxyribose ring and purine ring. It is possible that these interactions are conserved in *Limulus* SAP and also *Limulus* CRP, not only with AMP but also with ribose-5-phosphate except the purine ring interactions. As the affinity of *Limulus* CRP and SAP are similar for

both ribose-5-phosphate and AMP, binding is likely to occur through the ribose-5-phosphate moiety which both molecules possess (see figure 7.1). This would indicate that the purine ring on AMP is not essential for binding.

Ribose-5-phosphate is a constituent of DNA and RNA, and it is possible that *Limulus* CRP and SAP bind to nuclear debris in the horseshoe crab and elicit their removal. It is also possible that binding of *Limulus* CRP and SAP to ribose-5-phosphate and AMP is coincidental due to their affinity for phosphorylated molecules (Robey and Liu 1981; Quigley *et al.*, 1994).

Figure 7.1. Structures of ligands ribose-5-phosphate and AMP (adapted from www.bmrb.wisc.edu).

Ribose-5-phosphate   Adenosine-5-monophosphate
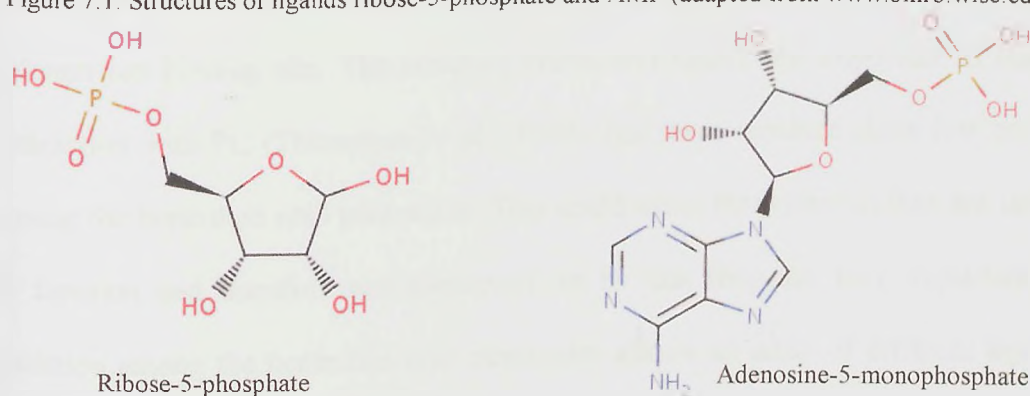
7.4 Analysis of horseshoe crab CRP and SAP sequences

Phylogenetic analysis indicated that CRP from *Tachypleus tridentatus*, *Limulus polyphemus*, and *Carcinoscorpius rotundicauda* form groups of homologous sequences. The clustering of *Carcinoscorpius* CRP 1 and 2 with *Tachypleus* CRP 1 and 2 on a phylogenetic tree agreed with results shown by Ng *et al.*, (2004). As there are only three sequences for *Limulus* CRP, they do not form individual groups but do cluster with *Tachypleus* CRP2 and *Carcinoscorpius* CRP2 indicating sequence homology.

Analysis of the CRP and SAP sequences from *Limulus polyphemus*, *Tachypleus tridentatus*, and *Carcinoscorpius rotundicauda* at areas corresponding to the calcium

binding site, PC binding site and putative C1q cleft of human CRP showed that the calcium ion binding regions were most conserved. This was to be expected as the proteins exhibit calcium dependent ligand binding. The areas of sequence corresponding to the calcium binding loop (minus the residues that interact directly with the calcium ions) showed less conservation between the horseshoe crab CRP sequences, which may be because these residues do not interact directly with the calcium ions and do not need to be conserved to retain function.

At the areas of sequence of horseshoe crab CRPs and SAP that corresponds to the PC binding site in human CRP, there is lower conservation than at areas corresponding to the calcium ion binding site. The residues Phe66 and Glu81 are important in human CRP interaction with PC (Thompson *et al.*, 1999), but these residues show low conservation among the horseshoe crab pentraxins. This could mean that either a) they are unimportant in function and therefore not conserved, or b) that they are very important and that variation among the horseshoe crab pentraxins allows an array of different ligands to be bound.

No C1q-like proteins have been discovered in the horseshoe crabs so it is very unlikely that the horseshoe crab pentraxins have a C1q cleft. However, comparison of areas of sequence from the horseshoe crab CRP and SAP pentraxins corresponding to the putative C1q binding cleft in human CRP, showed a high conservation but only within the horseshoe crab sequences themselves and not with human CRP. It is possible that the high conservation is because of a conserved site in horseshoe crab pentraxins that has a similar function within the horseshoe crabs, but dissimilar to that of human CRP due to the lack of C1q. Such an example may be where CRP interacts with C2/bf (a complement homologue), factor C (a clotting cascade protein), or CrOctin in *Carcinoscorpius rotundicauda* (Li *et al.*, 2007; Le Saux *et al.*, 2008). On the structure of *Limulus* SAP, the area on the protomer that is compared to the putative C1q cleft in human CRP is in the

354

pore of the octamer, and therefore is very unlikely to function as a binding site. This would also be the case if the protomer arrangement in horseshoe crab CRP was similar to that in *Limulus* SAP. Identification of other areas of conservation within the horseshoe crab CRP corresponded to an area of *Limulus* SAP that was on the exposed face of the hexadecamer which may be a candidate for a binding area.

Comparison of horseshoe crab CRP and SAP sequences with more mammalian CRP and SAP sequences rather than just human CRP, could highlight other areas of conservation at the calcium and ligand binding sites, and also any other areas of conservation that may be important in protein function and interactions. As the human SAP structure with dAMP has been solved (Hohenester *et al.*, 1997) and the novel ligands ribose-5-phosphate and AMP have been discovered for *Limulus* CRP and SAP, a sequence comparison between the horseshoe crab CRP and SAP at the residues corresponding to those which interact with dAMP in human SAP, may shed some light on how these ligands interact with *Limulus* CRP and SAP.

7.5 Isolation of genes for dogfish CRP and SAP

The genes for dogfish (*Mustelus canis*) CRP and SAP were not found, despite several attempts to discover them. This was thought to be primarily due to the degeneracy of the primers that were used, which could not have been avoided due to the use of only limited peptide sequences for primer design. The amino acid sequences used to design the primers were only short at around 7 amino acids, leaving only around 20 base pairs for each primer. The small length of primer probably contributed to the observed non-specific annealing of the primers to dogfish DNA.

Further attempts to isolate the genes for dogfish CRP and SAP could be performed using the same techniques as used in this study, but with an increased chance in amplifying

the correct DNA by using reverse transcription to create more template DNA for PCR, and by using longer less degenerate primers. The latter could be achieved by further peptide cleavage experiments which may give longer peptide sequences from which to design primers, and by comparison of the nucleotide sequences for CRP and SAP between species to identify conserved stretches of nucleotides. It would also be possible to create a cDNA library of dogfish DNA which would be screened with multiple oligonucleotide probes.

References

Abe, N., Osanai, T., Fujiwara, T., Kameda, K., Matsunaga, T., Okumura, K. (2005). C-reactive protein-induced upregulation of extracellular matrix metalloproteinase inducer in macrophages; inhibitory effect of fluvastatin. Life Sciences. **78**. 9. 1021-1028

Abernethy, T. J. (1936). Studies on the somatic C-polysaccharide of pneumococcus. II. Journal of Experimental Medicine. **65**. 75-89.

Abernethy, T. J. and Avery, O. T. (1941). The occurrence during acute infections of a protein not normally present in the blood. I. Journal of Experimental Medicine. **73**. 173-182.

Ablij, H. C., Meinders, A. E. (2002). C-reactive protein: history and revival. European Journal of Internal Medicine. **13**. 412-422

Agrawal, A., Xu, Y., Ansardi, D., Macon, K. J., Volanakis, J.E. (1992). Probing the phosphocholine-binding site of human C-reactive protein by site-directed mutagenesis. The Journal of Biological Chemistry. **267**. 35. 25352-25358.

Agrawal, A. and Volanakis, J. E. (1994). Probing the C1q-binding site on human C-reactive protein by site-directed mutagenesis. The Journal of Immunology. **152**. 5404-5410.

Agrawal, A., Lee, S., Carson, M., Narayana, S. V. L., Greenhough, T. J., Volanakis, J. E. (1997). Site-directed mutagenesis of the phosphocholine-binding site of human C-reactive protein. The Journal of Immunology. **158**. 345-350.

Agrawal, A., Shrive, A. K., Greenhough, T. J., Volanakis, J. E. (2001). Topology and structure of the C1q-binding site on C-reactive protein. The Journal of Immunology. **166**. 3998-4004.

Agrawal, A., Simpson, M. J., Black, S., Carey, M. P., Samols, D. (2002). A C-reactive protein mutant that does not bind to phosphocholine and Pneumococcal C-polysaccharide. The Journal of Immunology. **169**. 3217-3222.

Aitkens, J. A., Geisow, M. J., Findlay, J. B. C., Holmes, C., Yarwood, A. (1989). Protein Sequencing. A Practical Approach. IRL Press. Oxford.

Aketagawa, J., Miyata, T., Ohtsubo, S., Nakamura, T., Morita, T., Hayashida, H., Miyata, T., Iwanaga, S. (1986). Primary structure of *Limulus* anticoagulant anti-lipopolysaccharide factor. The Journal of Biological Chemistry. **261**. 16. 7357-7365.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research. **25**. 17. 3389-3402.

Amatayakul-Chantler, S., Dwek, R. A., Yennent, G. A., Pepys, M. B., Rademacher, T. W. (1993). Molecular characterization of *Limulus polyphemus* C-reactive protein. II. Asparagine-linked oligosaccharides. European Journal of Biochemistry. **214**. 1. 99-110

Anderson, H. C. and McCarty, M. (1950). The occurrence in the rabbit of an acute phase protein analogous to human C-reactive protein. Journal of Experimental Medicine. **93**. 25-36.

Anderson, J. K., Mole, J. E. (1982). Large scale isolation and partial primary structure of human plasma amyloid P component. Annals of the New York Academy of Sciences. **389.** 216-238

Andrews, A. T. (1986). Electrophoresis theory, techniques, and biochemical and clinical applications. Second Edition. Oxford Clarendon Press.

Armstrong, P. B. (1977). Interaction of the motile blood cells of the horseshoe crab, *Limulus*. Studies on contact paralysis of pseudopodial activity and cellular overlapping in vitro. Experimental Cell Research. **107.** 1. 127-138

Armstrong, P. B., Levin, J. (1979). In vitro phagocytosis by *Limulus* blood cells. Journal of Invertebrate Pathology. **34.** 2. 145-151

Armstrong, P. B. (1980). Adhesion and spreading of *Limulus* blood cells on artificial surfaces. J. Cell. Sci. **44.** 243-262.

Armstrong, P. B., Rickles, F. R. (1982). Endotoxin-induced degranulation of the *Limulus* amebocyte. Experimental Cell Research. **140.** 15-24.

Armstrong, P. B., Quigley, J. P. (1985). Proteinase inhibitory activity released from the horseshoe crab blood cell during exocytosis. Biochimica et Biophysica Acta. **827.** 453-459

Armstrong, P. B. and Quigley, J. P. (1987). *Limulus* α2-macroglobulin. First evidence in an invertebrate for a protein containing an internal thiol ester bond. Biochem. J. **248.** 703-707.

Armstrong, P. B., Misquith, S., Srimal, S., Melchior, R., Quigley, J. P. (1994). Identification of Limulin as a major cytolytic protein in the plasma of the American horseshoe crab, *Limulus polyphemus*. Biol. Bull. **187.** 227-230.

Armstrong, P. B., Melchior, R., Quigley, J. P. (1996). Humoral imunity in long-lived arthropods. J. Insect. Physiol. **42.** 1. 53-64.

Armstrong, P. B., Swarnakar, S., Srimal, S., Misquith, S., Hahn, E. A., Aimes, R. T., Quigley, J. P. (1996). A cytolytic function for a sialic acid-binding lectin that is member of the pentraxin family of proteins. The Journal of Biological Chemistry. **271.** 25. 14717-14721

Armstrong, P. B., Melchior, R., Swarnakar, S., Quigley, J. P. (1998). α2-macroglobulin does not function as a C3 homologue in the plasma haemolytic system of the American horseshoe crab, *Limulus*. Molecular Immunology. **35.** 47-53.

Armstrong, P. B. and Quigley, J. P. (1999). α2-macroglobulin: an evolutionarily conserved arm of the innate immune system. Developmental and comparative immunology. **23.** 375-390.

Armstrong, P. B. (2001). The contribution of proteinase inhibitors to immune defense. TRENDS in Immunology. **22.** 1. 47-52.

Ashton, A W., Boehm, M. K., Gallimore, J. R., Pepys, M. B., Perkins, S. J. (1997). Pentameric and decameric structures in solution of serum amyloid P component by X-ray and neutron scattering and molecular modelling. J. Mol. Biol. **272.** 408-422.

Assimeh, S. N., Painter, R. H. (1975a). The macromolecular structure of the first component of complement. J. Immunol. **115**. 2. 488-494

Assimeh, S. N., Painter, R. H. (1975b). The identification of a previously unrecognized subcomponent of the first component of complement. J. Immunol. **115**. 2. 482-487

Bailon, P., Ehrlich, G. K., Fung, W.-J., Berthold, W. (2000). Affinity Chromatography: Methods and Protocols. Humana Press.

Baldo, B. A., Fletcher, T. C., Pepys, J. (1977). Isolation of a peptide-polysaccharide from the dermatophyte *Epidermophyton floccosum* and a study of its reaction with human C-reactive protein and a mouse anti-phosphorylcholine myeloma serum. Immunology. **32**. 831-842.

Ballou, S. P., Buniel, J., MacIntyre, S. S. (1989). Specific binding of human C-reactive protein to human monocytes *in vitro*. J. Immunol. **142**. 8. 2708-2713

Ballou, S. P., MacIntyre, S. S. (1990). Absence of a binding reactivity of human C-reactive protein for immunoglobulin or immune complexes. J. Lab. Clin. Med. **115**. 3. 332-338

Ballou, S. P. and Cleveland, R. P. (1991). Binding of human C-reactive protein to monocytes: analysis by flow cytometry. Clin. Exp. Immunol. **84**. 329-335.

Baltz, M. L., De Beer, F. C., Feinstein, A., Pepys, M. B. (1982). Calcium-dependent aggregation of human serum amyloid P component. Biochim. Biophys. Acta. **701**. 2. 229-236

Bang, F. B. (1956). A bacterial disease of *Limulus polyphemus*. Bulletin of the Johns Hopkins Hospital. **98**. 5. 325-351

Bang, R., Marnell, L., Mold, C., Stein, M-P., Du Clos, K. T., Chivington-Buck, C., Du Clos, T. W. (2005). Analysis of binding sites in human C-reactive protein for FcγRI, FcγRIIA, and C1q by site-directed mutagenesis. The Journal of Biological Chemistry. **280**. 26. 25095-25102.

Barnum, S., Narkates, A. J., Suddath, F. L., Volanakis, J. E. (1982). Comparative studies on the binding specificities of C-reactive protein (CRP) and HOPC8. Annals of the New York Academy of Sciences. **389**. 431-434

Baum, L. L., James, K. K., Glaviano, R. R., Gewurz, H. (1983). Possible role for C-reactive protein in the human natural killer cell response. J. Exp. Med. **157**. 301-311.

Bharadwaj, D., Stein, M-P., Volzer, M., Mold, C., Du Clos, T. W. (1999). The major receptor for C-reactive protein on leukocytes is Fcγ receptor II. The Journal of Experimental Medicine. **190**. 4. 585-590.

Bharadwaj, D., Mold, C., Markham, E., Du Clos, T. W. (2001). Serum amyloid P component binds to Fcγ receptors and opsonizes particles for phagocytosis. The Journal of Immunology. **166**. 6735-6741.

Bickerstaff, M. C. M., Botto, M., Hutchinson, W. L., Herbert, J., Tennent, G. A., Bybee, A., Mitchell, D. A., Cook, H. T., Butler, P. J. G., Walport, M. J., Pepys, M. B. (1999). Serum amyloid P component controls chromatin degradation and prevents antinuclear autoimmunity. Nature Medicine. **5**. 6. 694-697.

Binette, P., Binette, M., Calkins, E. (1974). The isolation and identification of the P-component of normal human plasma proteins. Biochem. J. **143**. 253-254.

Bishayee, S., Dorai, D. T. (1980). Isolation and characterisation of a sialic acid-binding lectin (Carcinoscorpin) from Indian horseshoe crab *Carcinoscorpius rotundicauda*. Biochemica et Biophysica Acta. **623**. 89-97

Black, S., Agrawal, A., Samols, D. (2003). The phosphocholine and the polycation-binding sites on rabbit C-reactive protein are structurally and functionally distinct. Molecular Immunology. **39**. 16. 1045-1054

Bladen, H. A., Nylen, M. U., Glenner, G. G. (1966). The ultrastructure of human amyloid as revealed by the negative staining technique. J. Ultrastructure Research. **14**. 449-459.

Bodman-Smith, K. B., Melendez, A. J., Campbell, I., Harrison, P. T., Allen, J. M., Raynes, J. G. (2002). C-reactive protein-mediated phagocytosis and phospholipase D signalling through the high-affinity receptor for immunoglobulin G (FcγRI). Immunology. **107**. 252-260.

Bodman-Smith, K. B., Gregory, R. E., Harrison, P. T., Raynes, J. G. (2004). FcγRIIA expression with FcγRI results in C-reactive protein and IgG-mediated phagocytosis. Journal of Leukocyte Biology. **75**. 1029-1035.

Bodmer, B., Siboo, R. (1977). Isolation of mouse C-reactive protein from liver and serum. J. Immunol. **118**. 3. 1086-1089

Bottazzi, B., Vouret-Craviari, V., Bastone, A., De Gioia, L., Matteucci, C., Peri, G., Spreafico, F., Pausa, M., D'Ettorre, C., Gianazza, E., Tagliabue, A., Salmona, M., Tedesco, F., Introna, M., Mantovani, A. (1997). Multimer formation and ligand recognition by the long pentraxin PTX3. The Journal of Biological Chemistry. **272**. 52. 32817-32823.

Botto, M., Hawkins, P. N., Bickerstaff, M. C. M., Herbert, J., Bygrave, A. E., McBride, A., Hutchinson, W. L., Tennent, G. A., Malport, M. J., Pepys, M. B. (1997). Amyloid deposition is delayed in mice with targeted deletion of the serum amyloid P component gene. Nature Medicine. **3**. 8. 855-859

Bout, D., Joseph, M., Pontet, M., Vorng, H., Deslee, D., Capron, A. (1986). Rat resistance to Schistosomiasis: platelet-mediated cytotoxicity induced by C-reactive protein. Science. **231**. 4734. 153-156

Bray, R. A., Samberg, N. L., Gewurz, H., Potempa, L. A., Landay, A. L. (1988). C-reactive protein antigenicity on the surface of human peripheral blood lymphocytes. Characterization of lymphocytes reactive with anti-neo-CRP. J. Immunol. **140**. 12. 4271-4278

Breathnach, S. M., Melrose, S. M., Bhogal, B., de Beer, F. C., Dyck, R. F., Tennent, G., Black, M. M., Pepys, M. B. (1981). Amyloid P component is located on elastic fibre microfibrils in normal human tissue. Nature. **293**. 5834. 652-654

Breathnach, S. M., Kofler, H., Sepp, N., Ashworth, J., Woodrow, D., Pepys, M. B., Hintner, H. (1989). Serum amyloid P component binds to cell nuclei *in vitro* and to *in vivo* deposits of extracellular chromatin in systemic lupus erythematosus. J. Exp. Med. **170**. 1433-1438.

Briles, D. E., Nahm, M., Schroer, K., Davie, J., Baker, P., Kearney, J., Barletta, R. (1981). Antiphosphocholine antibodies found in normal mouse serum are protective against intravenous infection with type 3 *Streptococcus pneumoniae*. J. Exp. Med. **153**. 694-705.

Brundish, D. E. and Baddiley, J. (1968). Pneumococcal C-substance, a ribitol teichoic acid containing choline phosphate. Biochem. J. **110**. 573-582.

Buchta, R., Fridkin, M., Pontet, M., Romeo, D. (1986). Synthetic peptides from C-reactive protein containing tuftsin-related sequences. Peptides. **7**. 6. 961-968

Buchta, R., Fridkin, M., Pontet, M., Contessi, E., Scaggiante, B., Romeo, D. (1987). Modulation of human neutrophil function by C-reactive protein. Eur. J. Biochem. **163**. 141-146.

Buchta, R., Pontet, M., Fridkin, M. (1987). Binding of C-reactive protein to human neutrophils. FEBS letters. **211**. 2. 165-168.

Butler, P. J. G., Tennent, G. A., Pepys, M. B. (1990). Pentraxin-chromatin interactions: serum amyloid P component specifically displaces H1-type histones and solubilizes native long chromatin. J. Exp. Med. **172**. 13-18.

Cabana, V. G., Gewurz, H., Siegel, J. N. (1982). Interaction of very low density lipoprotin (VLDL) with rabbit C-reactive protein. J. Immunol. **128**. 5. 2342-2348

Cardin, A. D., Weintraub, H. J. (1989). Molecular modelling of protein-glycosaminoglycan interactions. Arteriosclerosis. **9**. 1. 21-32

Cartwright, J. R., Tharia, H. A., Burns, I., Shrive, A. K., Hoole, D., Greenhough, T. J. (2004). Isolation and characterisation of pentraxin-like serum proteins from the common carp *Cyprinus carpio*. Developmental and Comparative Immunology. **28**. 113-125

Castell, J. V., Gómez-Lechón, M. J., David, M., Fabra, R., Trullenque, R., Heinrich, P. C. (1990). Acute-phase response of human hepatocytes: regulation of acute-phase protein synthesis by interleukin-6. Hepatology. **12**. 5. 1179-1186

Cermak, J., Key, N. S., Bach, R. R., Balla, J., Jacob, H. S., Vercellotti, G. M. (1993). C-reactive protein induces human peripheral blood monocytes to synthesize tissue factor. Blood. **82**. 2. 513-520.

Cha-Molstad, H., Agrawal, A., Zhang, D., Samols, D., Kushner, I. (2000). The Rel family member p50 mediates cytokine-induced C-reactive protein expression by a novel mechanism. The Journal of Immunology. **165**. 4592-4597.

Chang, M-K., Binder, C. J., Torzewski, M., Witzum, J. L. (2002). C-reactive protein binds to both oxidized LDL and apoptotic cells through recognition of a common ligand: phosphorylcholine of oxidized phospholipids. PNAS. **99**. 20. 13043-13048.

Chen, S-C., Yen, C-H., Yeh, M-S., Huang, C-J., Liu, T-Y. (2001). Biochemical properties and cDNA cloning of two new lectins from the plasma of *Tachypleus tridentatus*. The Journal of Biological Chemistry. **276**. 13. 9631-9639.

Chiou, S-T., Chen, Y-W., Chen, S-C., Chao, C-F., Liu, T-Y. (2000). Isolation and characterization of proteins that bind to galactose, lipopolysaccharide of *Escherichia coli*, and protein A of *Staphylococcus aureus* from the haemolymph of *Tachypleus tridentatus*. The Journal of Biology Chemistry. **275**. 3. 1630-1634.

Christner, R. B., Mortensen, R. F. (1994). Specificity of the binding interaction between human serum amyloid P component and immobilized human C-reactive protein. The Journal of Biological Chemistry. **269**. 13. 9760-9766.

Ciliberto, G., Arcone, R., Wagner, E. F., Ruther, U. (1987). Inducible and tissue-specific expression of human C-reactive protein in transgenic mice. The EMBO Journal. **6**. 13. 4017-4022.

Cirillo, P., Golino, P., Calabro, P., Cali, G., Ragni, M., De Rosa, S., Cimino, G., Pacileo, M., De Palma, R., Forte, L., Gargiulo, A., Corigliano, F. G., Angri, V., Spagnuolo, R., Nitsch, L., Chiariello, M. (2005). C-reactive protein induces tissue factor expression and promotes smooth muscle and endothelial cell proliferation. Cardiovascular Research. **68**. 47-55.

Coe, J. E. (1977). A sex-linked serum protein of Syrian hamsters; definition of female protein and regulation by testosterone. Proc. Natl. Acad. Sci. USA. **74**. 2. 730-733.

Coe, J. E., Margossian, S. S., Slayter, H. S., Sogn, J. A. (1981). Hamster female protein. A new pentraxin structurally and functionally similar to C-reactive protein and amyloid P component. The Journal of Experimental Medicine. **153**. 977-991

Coe, J. E., Ross, M. J. (1983). Hamster female protein. A divergent acute phase protein in male and female Syrian hamsters. Journal of Experimental Medicine. **157**. 1421-1433.

Coe, J. E., Ross, M. J. (1985). Hamster female protein, a sex-limited pentraxin is a constituent of Syrian hamster amyloid. The Journal of Clinical Investigation. **76**. 66-74

Coe, J. E., Ross, M. J. (1990). Amyloidosis and female protein in the Syrian hamster. The Journal of Experimental Medicine. **171**. 1257-1267.

Conrath, C. L., Musick, J. A. (2002). Reproductive biology of the smooth dogfish, *Mustelus canis*, in the northwest Atlantic Ocean. Environmental Biology of Fishes. **64**. 367-377

Conrath, C. L., Gelsleichter, J., Musick, J. A. (2002). Age and growth of the smooth dogfish (*Mustelus canis*) in the northwest Atlantic Ocean. Fishery Bulletin. **11**. 674-682

Cook, M. T., Hayball, P. J., Birdseye, L., Bagley, C., Nowak, B. F., Hayball, J. D. (2003). Isolation and partial characterization of a pentraxin-like protein with complement-fixing activity from Snapper (*Pagrus auratus*, Sparidae) serum. Developmental and Comparative Immunology. **27**. 579-588.

Cook, M. T., Haybull, P. J., Nowak, B. F., Haybull, J. D. (2005). The opsonising activity of a pentraxin-like protein isolated from snapper (*Pagrus auratus*, Sparidae) serum. Developmental and Comparative Immunology. **29**. 703-712.

Crowell, R. E., Du Clos, T. W., Montoya, G., Heaphy, E., Mold, C. (1991). C-reactive protein receptors on the human monocytic cell line U-937. Evidence for additional binding to Fc gamma RI. J. Immunol. **147**. 10. 3445-3451

Cuatrecasas, P., Wilchek, M., Anfinsen, C. B. (1968). Selective enzyme purification by affinity chromatography. Proc. Natl. Acad. Sci. USA. **61**. 2. 636-643

Culley, F. J., Harris, R. A., Kaye, P. M., McAdam, K. P., Raynes, J. G. (1996). C-reactive protein binds to a novel lignd on *Leishmania donovani* and increases uptake into human macrophages. J. Immunol. **156**. 12. 4691-4696

Culley, F. J., Bodman-Smith, K. B., Ferguson, M. A. J., Nikolaev, A. V., Shantilal, N., Raynes, J. G. (2000). C-reactive protein binds to phosphorylated carbohydrates. Micobiology. **10**. 1. 59-65.

Danenberg, H. D., Kantak, N., Grad, E., Swaminathan, R. V., Lotan, C., Edelman, E. R. (2007). C-reactive protein promotes monocyte-platelet aggregation: an additional link to the inflammatory-thrombotic intricacy. Eur. J. Haematol. **78**. 3. 246-252

Danesh, J., Collins, R., Appleby, P., Peto, R. (1998). Association of fibrinogen, C-reactive protein, albumin, or leukocyte count with coronary heart disease: meta-analyses of prospective studies. The Journal of the American Medical Association. **279**. 18. 1477-1482

Danielson, B., Sørensen, I. J., Nybo, M., Holm Nielsen E., Kaplan, B., Svehag, S-E. (1997). Calcium –dependent and independent binding of the pentraxin serum amyloid P component to glycosaminoglycans and amyloid proteins: enhanced binding at slightly acid pH. Biochimica et Biophysica Acta. **1339**. 73-78

Das, T., Sen, A., Kempf, T., Pramanik, R. R., Mandal, C., Mandal, C. (2003). Induction of glycosylation in human C-reactive protein under different pathological conditions. Biochem. J. **373**. 345-355.

De Beer, F. C., Baltz, M. L., Holford, S., Feinstein, A., Pepys, M. B. (1981). Fibronectin and C4-binding protein are selectively bound by aggregated amyloid P component. J. Exp. Med. **154**. 1134-1149.

De Beer, F. C., Baltz, M. L., Munn, E. A., Feinstein, A., Taylor, J., Bruton, C., Clamp, J. R., Pepys, M. B. (1982). Isolation and characterization of C-reactive protein and serum amyloid P component in the rat. Immunology. **45**. 55-70

De Frutos, P. G., Dahlback, B. (1994). Interaction between serum amyloid P component and C4b-binding protein associated with inhibition of factor I mediated C4b degradation. Journal of Immunology. **152**. 2430-2437.

Devaraj, S., Venugopal, S., Jialal, I. (2006). Native pentameric C-reactive protein displays more potent pro-atherogenic activities in human aortic endothelial cells than modified C-reactive protein. Atherosclerosis. **184**. 48-52.

Devaraj, S., Dasu, M. R., Singh, U., Rao, L. V., Jiala, I. (2008). C-reactive protein stimulates superoxide anion release and tissue factor activity in vivo. Atherosclerosis. In print.

Di Camelli, R., Potempa, L. A., Siegel, J., Suyehira, L., Petras, K., Gewurz, H. (1980). Binding reactivity of C-reactive protein for polycations. The Journal of Immunology. **125**. 5. 1933-1937.

Diehl, E. E., Haines, G. K., Radosevich, J. A., Potempa, L. A. (2000). Immunohistochemical localization of modified C-reactive protein antigen in normal vascular tissue. Am. J. Med. Sci. **319**. 2. 79-83

Ding, J. L., Navas, M. A. A., Ho, B. (1993). Two forms of factor C from the amoebocytes of *Carcinoscorpius rotundicauda*: purification and characterisation. Biochimica et Biophysica Acta. **1202**. 149-156.

Dobrinich, R., Spagnuolo, P. J. (1991). Binding of C-reactive protein to human neutrophils. Inhibition of respiratory burst activity. Arthritis and Rheumatism. **34**. 8. 1031-1037

Dong, A., Caughey, W. S., Du Clos, T. W. (1994). Effects of calcium, magnesium, and phosphorylcholine on secondary structures of human C-reactive protein and serum amyloid P component observed by infrared spectroscopy. The Journal of Biological Chemistry. **269**. 9. 6424-6430.

Dorai, D. T., Bachhawat, B. K., Bishayee, S., Kannan, K., Rao, D. R. (1981). Further characterization of the sialic acid-binding lectin from the horseshoe crab *Carcinoscorpius rotunda cauda*. Arch Biochem. Biophys. **209**. 1. 325-333

Dowton, S. B., Woods, D. E., Mantzouranis, E. C., Colten, H. R. (1985). Syrian hamster female protein: analysis of female protein primary structure and gene expression. Science. **228**. 4704. 1206-1208

Dowton, S. B., McGrew, S. D. (1990). Rat serum amyloid P component. Biochemistry Journal. **270**. 553-556.

Dowton, S. B., Holden, S. N. (1991). C-reactive (CRP) of the Syrian hamster. Biochemistry. **30**. 9531-9538

Du Clos, T. W., Mold, C., Paterson, P. Y., Alroy, J., Gewurz, H. (1981). Localization of C-reactive protein in inflammatory lesions or experimental allergic encephalomyelitis. Clin. Exp. Immunol. **43**. 565-573

Du Clos, T. W., Zlock, L. T., Rubin, R. L. (1988). Analysis of the binding of C-reactive protein to histones and chromatin. The Journal of Immunology. **141**. 4266-4270.

Du Clos, T. W. (1989). C-reactive protein reacts with the U1 small nuclear ribonucleoprotein. The Journal of Immunology. **143**. 2553-2559.

Du Clos, T. W., Mold, C., Stump, R. F. (1990). Identification of a polypeptide sequence that mediates nuclear localization of the acute phase protein C-reactive protein. The Journal of Immunology. **145**. 3869-3875

Du Clos, T. W., Marnell, L., Zlock, L. R., Burlingame, R. W. (1991) (a). Analysis of the binding of C-reactive protein to chromatin subunits. The Journal of Immunology. **146**. 1220-1225.

Du Clos, T. W., Zlock, L. T., Marnell, L. (1991) (b). Definition of a C-reactive protein binding determinant on histones. The Journal of Biological Chemistry. **266**. 4. 2167-2171

Du Clos, T. W., Zlock, L. T., Hicks, P. S., Mold, C. (1994). Decreased autoantibody levels and enhanced survival of (NZB x NZW)1 mice treated with C-reactive protein. Clinical Immunology and Immunopathology. **70**. 1. 22-27

Duong, T., Acton, P. J., Johnson, R. A. (1998). The *in vitro* neuronal toxicity of pentraxins associated with Alzheimer's disease brain lesions. Brain Research. **813**. 303-312.

Dyck, R. F., Evans, D. J., Lockwood, C. M., Rees, A. J., Turner, D., Pepys, M. B. (1980). Amyloid P-component in human glomerular basement membrane. Abnormal patterns of immunofluorescent staining in glomerular disease. Lancet. **2**. 8195 pt 1. 606-609

Edagawa, T., Murata, M., Hattori, M., Onuma, M., Kodama, H. (1993). Cell surface C-reactive protein of rainbow trout lymphocytes. Developmental and Comparative Immunology. **17**. 2. 119-127.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research. **32**. 5. 1792-1797.

Edwards, K. M., Gewurz, H., Lint, T. F., Mold, C. (1982). A role for C-reactive protein in the complement-mediated stimulation of human neutrophils by type 27 *Streptococcus pneumoniae*. J. Immunol. **128**. 6. 2493-2496

Egenhofer, C., Alsdorff, K., Fehsel, K., Kolb-Bachofen, V. (1993). Membrane-associated C-reactive protein on rat liver macrophages is synthesised within the macrophages, expressed as neo-C-reactive protein and bound through a C-reactive protein-specific membrane receptor. Hepatology. **18**. 5. 1216-1223

Emsley, J., White, H. E., O'Hara, B. P., Oliva, G., Srinivasan, N., Tickle, I. J., Blundell, T. L., Pepys, M. B., Wood, S. P. (1994). Structure of pentameric human serum amyloid p component. Nature. **367**. 338-345.

Enghild, J. J., Thøgersen, I. B., Salvesen, G., Fey, G. H., Figler, N. L., Gonias, S. L., Pizzo, S. V. (1990). α-macroglobulin from *Limulus polyphemus* exhibits proteinase inhibitory activity and participates in a haemolytic system. Biochemistry. **29**. 10070-10080.

Familian, A., Zwart, B., Huisman, H. G., Rensink, I., Roem, D., Hordijk, P. L., Aarden, L. A., Hack, C. E. (2001). Chromatin-Independent binding of serum amyloid P component to apoptotic cells. The Journal of Immunology. **167**. 647-654.

Fernandez-Moran, H., Marchalonis, J. J., Edelman, G. M. (1968). Electron microscopy of a hemagglutinin from *Limulus polyphemus*. J. Mol. Biol. **32**. 467-469.

Fiedel, B. A., Gewurz, H. (1976a). Effects of C-reactive protein on platelet function. I. Inhibition of platelet aggregation and release reactions. J. Immunol. **116**. 5. 1289-1294

Fiedel, B. A., Gewurz, H. (1976b). Effects of C-reactive protein on platelet function. II. Inhibition by CRP of platelet reactivities stimulated by poly-L-lysine, ADP, epinephrine, and collagen. J. Immunol. **117**. 4. 1073-1078

Fiedel, B. A., Rent, R., Myhrman, R., Gewurz, H. (1976). Complement activation by interaction of polyanions and polycations. II. Precipitation and role of IgG, C1q and C1-INH during heparin-protamine-induced consumption of complement. **30**. 161-169

Fiedel, B. A., Simpson, R. M., Gewurz, H. (1982). Activation of platelets by modified C-reactive protein. Immunology. **45**. 439-447

Fiedel, B. A., Potempa, L. A., Frenzke, M. E., Simpson, R. M., Gewurz, H. (1982). Platelet inhibitory effects of CRP preparations are due to a co-isolating low molecular weight factor. Immunology. **45**. 15-22.

Fiedel, B. A., Ku, C. S. (1986). Further studies on the modulation of blood coagulation by human serum amyloid P component and its acute phase homologue C-reactive protein. Thromb. Haemost. **55**. 3. 406-409

Fu, T., Borensztajn, J. (2002). Macrophage uptake of low-density lipoprotein bound to aggregated C-reactive protein; possible mechanism of foam-cell formation in atherosclerotic lesions. Biochem. J. **366**. 195-201.

Fujii, N., Minetti, C. A. S. A., Nakhasi, H. L., Chen, S-W., Barbehenn, E., Nunes, P. H., Nguyen, N. Y. (1992). Isolation, cDNA cloning, and characterisation of an 18kDa hemaglutinin and amoebocyte aggregation factor from *Limulus polyphemus*. The Journal of Biological Chemistry. **267**. 31. 22452-22459.

Fujiki, K., Shin, D. H., Nanakao, M., Yano, T. (2000). Molecular cloning and expression analysis of carp (*Cyprinus carpio*) interleukin-1 beta, high affinity immunoglobulin E Fc receptor gamma subunit and serum amyloid A. Fish Shellfish Immunol. **10**. 3. 229-242

Funke, C., King, D. P., Brotheridge, R. M., Adelung, D., Stott, J. L. (1997). Harbor seal (*Phoca vitulina*) C-reactive protein (CRP): purification, characterization of specific monoclonal antibodies and development of an immuno-assay to measure serum CRP concentrations. Veterinary Immunology and Immunopathology. **59**. 151-162.

Furman, R. M., Pistole, T. G. (1976). Bactericidal activity of haemolymph from the horseshoe crab, *Limulus polyphemus*. Journal of Invertebrate Pathology. **28**. 2. 239-244.

Gaal, O., Medgyesi, G. A., Vereczkey, L. (1980). Electrophoresis in the separation of biological macromolecules. Chichester Wiley.

Gaboriaud, C., Juanhuix, J., Gruez, A., Lacroix, M., Darnault, C., Pignol, D., Verger, D., Fontecilla-Camps, J. C., Arlaud, G. J. (2003). The crystal structure of the globular head of complement protein C1q provides a basis for its versatile recognition properties. The Journal of Biological Chemistry. **278**. 47. 46974-46982

Ganapathi, M. K., Rzewnicki, D., Samols, D., Jiang, S. L., Kushner, I. (1991). Effect of combinations of cytokines and hormones on synthesis of serum amyloid A and C-reactive protein in Hep 3B cells. J. Immunol. **147**. 4. 1261-1265

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Research. **31**. 13. 3784-3788.

Gelsleichter, J., Musick, J. A., Nichols, S. (1999). Food habits of the smooth dogfish, *Mustelus canis*, dusky shark, *Carcharhinus obscurus*, Atlantic sharpnose shark, *Rhizoprionodon terraenovae*, and the sand tiger, *Carcharias taurus*, from the northwest Atlantic Ocean. Environmental Biology of Fishes. **54**. 205-217

Gershov, D., Kim, S., Brot, N., Elkon, K. B. (2000). C-reactive protein binds to apoptotic cells, protects the cells from assembly of the terminal complement components, and sustains an anti-inflammatory innate immune response: implications for systemic autoimmunity. Journal of Experimental Medicine. **192**. 9. 1353-1363.

Gewurz, H., Zhang, X-H, Lint, T. F. (1995). Structure and function of the pentraxins. Current Opinion in Immunology: 7. 54-64

Ghosh, S., Bhattacharya, S. (1992). Elevation of C-reactive protein in serum of *Channa punctatus* as an indicator of water pollution. Indian J. Exp. Biol. **30**. 8. 736-737

Gitlin, J. D., Gitlin, J. I., Gitlin, D. (1977). Localization of C-reactive protein in synovium of patients with rheumatoid arthritis. Arthritis and Rheumatism. **20**. 8. 1491-1498

Glaser, F., Pupko, T., Paz, I., Bell, I., Bechor-Shental, D., Martz, E., Ben-Tal, N. (2003). ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics. **19**. 1. 163-164

Gokudan, S., Muta, T., Tsuda, R., Koori, K., Kawahara, T., Seki, N., Mizunoe, Y., Wai, S. N., Iwanaga, S., Kawabata, S-I. (1999). Horseshoe crab acetyl group-recognising lectins involved in innate immunity are structurally related to fibrinogen. Proc. Natl. Acad. Sci. USA. **96**. 10086-10091.

Goldman, N. D., Liu, T-Y. (1987). Biosynthesis of human C-reactive protein in cultured hepatoma cells is induced by a monocyte factor(s) other than interleukin-1. The Journal of Biological Chemistry. **262**. 5. 2363-2368.

Gotschlich, E., Stetson, C. A. (1959). Immunologic cross-reactions among mammalian acute phase proteins. Journal of Experimental Medicine. **111**. 441-451.

Gotschlich, E. C., Edelman, G. M. (1965). C-reactive protein: a molecule composed of subunits. Proc. Natl. Acad. Sci. **54**. 558-566.

Gotschlich, E. C., Edelman, G. M. (1967). Binding properties and specificity of C-reactive protein. Proc. Natl. Acad. Sci. **57**. 706-712.

Gotschlich, E. C., Liu, T-Y., Oliviera, E. (1982). Binding of C-reactive protein to C-carbohydrate and PC-substituted protein. Annals of the New York Academy of Sciences. **389**. 163-171

Gould, J. M., Weiser, J. N. (2001) Expression of C-reactive protein in the human respiratory tract. Infection and Immunity. **69**. 3. 1747-1754

Griselli, M., Herbert, J., Hutchinson, W. L., Taylor, K. M., Sohail, M., Krausz, T., Pepys, M. B. (1999). C-reactive protein and complement are important mediators of tissue damage in acute myocardial infarction. J. Exp. Med. **190**. 12. 1733-1739

Hamazaki, H. (1986). Purification and characterization of a human lectin specific for penultimate galactose residues. The Journal of Biological Chemistry. **261**. 12. 5455-5459.

Hamazaki, H. (1987). Ca^{2+} -mediated association of human serum amyloid P component with heparin sulphate and dermatan sulphate. The Journal of Biological Chemistry. **262**. 4. 1456-1460.

Hamazaki, H. (1989). Calcium-dependent polymerization of human serum amyloid P component is inhibited by heparin and dextran sulphate. Biochim. Biophys. Acta. **998**. 3. 231-235

Hamazaki, H. (1990). Structure and significance on N-linked sugar unit of human serum amyloid P component. Biochimica et Biophysica Acta. **1037**. 435-438.

Hamazaki, H. (1995). Ca^{2+} -dependent binding of human serum amyloid P component to Alzheimer's β-amyloid peptide. The Journal of Biological Chemistry. **270**. 18. 10392-10394.

Hames, B. (1998). Gel electrophoresis of proteins: a practical approach. Third Edition. Oxford University Press.

Harrington, J. M., Chou, H-T, Gutsmann, T., Gelhaus, C., Stahlberg, H., Leippe, M., Armstrong, P. B. (2008). Membrane pore formation by pentraxin proteins from *Limulus*, the American horseshoe crab. Biochem. J. **413**. 2. 305-313.

Harris, E. L. V., Angal, S. (1989). Protein purification methods a practical approach. Oxford IRL Press.

Harris, E. L. V., Angal, S. (1990). Protein purification applications a practical approach. Oxford IRL Press.

Hashii, T., Osaki, T., Koshiba, T., Kawabata, S-I. (2005). Characterization of a horseshoe crab keratin-like protein involved in transglutaminase-dependent protein cross-linking. Development of protein research. 2P-257

Haverkate, F., Thompson, S. G., Pyke, S. D. M., Gallimore, J. R., Pepys, M. B. (1997). Production of C-reactive protein and risk of coronary events in stable and unstable angina. The Lancet. **349**. 462-466

Hawkins, P. N., Rossor, M. N., Gallimore, J. R., Miller, B., Moore, E. G., Pepys, M. B. (1994). Concentration of serum amyloid P component in the CSF as a possible marker of cerebral amyloid deposits in Alzheimer's disease. Biochem. Biophys. Res. Commun. **201**. 2. 722-726

Heegaard, N. H., Heegaard, P. M., Roepstorff, P., Robey, F. A. (1996). Ligand-binding sites in human serum amyloid P component. European Journal of Biochemistry. **239**. 3. 850-856

Heftmann, E. (1983). Chromatography: Fundamentals and Applications of Chromatographic and Electrophoretic Methods. Part B Applications. Elsevier. Scientific Pub Co.

Heuertz, R. M., Schneider, G. P., Potempa, L. A., Webster, R. O. (2005). Native and modified C-reactive protein bind different receptors on human neutrophils. The International Journal of Biochemistry and Cell Biology. **37**. 320-335.

Hicks. P. S., Saunero-Nava, L., Du Clos, T. W., Mold, C. (1992). Serum amyloid P component binds to histones and activates the classical complement pathway. J. Immunol. **149**. 11. 3689-3694

Highton, J., Hessian, P. A., Palmer, D. G. (1985). The presence and possible significance of C-reactive protein in rheumatoid inflammation. J. Rheumatol. **12**. 5. 871-875

Hind, C. R. K., Collins, P. M., Renn, D., Cook, R. B., Caspi, D., Baltz, M. L., Pepys, M. B. (1984). Binding specificity of serum amyloid P component for the pyruvate acetal of galactose. J. Exp. Med. **159**. 1058-1069.

Hind, C. R. K., Collins, P. M., Baltz, M. L., Pepys, M. B. (1985). Human serum amyloid P component, a circulating lectin with specificity for the cyclic 4,6-pyruvate acetal of galactose. Biochem. J. **225**. 107-111.

Hirschfield, G. M., Hawkins, P. N. (2003). Amyloidosis: new strategies for treatment. The International Journal of Biochemistry and Cell Biology. **35**. 1608-1613

Hirschfield, G. M., Pepys, M. B. (2003). C-reactive protein and cardiovascular disease: new insights from an old molecule. Q J Med. **96**. 793-807

Hohenester, E., Hutchinson, W. L., Pepys, M. B., Wood, S. P. (1997). Crystal structure of a decameric complex of human serum amyloid P component with bound dAMP. Journal of Molecular Biology. **269**. 570-578.

Hokama, Y., Coleman, M. K., Riley, R. F. (1962). *In vitro* effects of C-reactive protein on phagocytosis. J. Bacteriol. **83**. 1017-1024.

Holm Nielsen, E., Nybo, M., Junker, K., Toftedal, P., Rasmussen, I. M., Svehag, S. E. (2000). Localization of human serum amyloid P component and heparin sulphate proteoglycan in *in vitro*-formed Aβ fibrils. Scand. J. Immunol. **52**. 110-112.

Holzer, T. J., Edwards, K. M., Gewurz, H., Mold, C. (1984). Binding of C-reactive protein to the pneumococcal capsule or cell wall results in differential localization of C3 and stimulation of phagocytosis. J. Immunol. **133**. 3. 1424-1430

Hoover, G. J., El-Mowafi, A., Simko, E., Kocal, T. E., Ferguson, H. W., Hayes, M. A. (1998). Plasma proteins of rainbow trout (*Oncorhynchus mykiss*) isolated by binding to lipopolysaccharide from *Aeromonas salmonicida*. Comparative Biochemistry and Physiology Part B. **120**. 559-569.

Hopkins, M., Flanagan, P. A., Bailey, S., Glover, I. D., Myles, D. A. A., Greenhough, T. J. (1994). Crystallization and preliminary X-ray analysis of C-reactive protein from rat. Journal of Molecular Biology. **235**. 767-771.

Horowitz, J., Volanakis, J. E., Briles, D. E. (1987). Blood clearance of *Streptococcus pneumoniae* by C-reactive protein. J. Immunol. **138**. 8. 2598-2603

Hu, S-I., Miller, S. M., Samols, D. (1986). Cloning and characterization of the gene for rabbit C-reactive protein. Biochemistry. **25**. 7834-7839.

Hughes, A. L. (1998). Protein phylogenies provide evidence of a radical discontinuity between arthropod and vertebrate immune systems. Immunogenetics. **47**. 4. 283-296

Hundt, M., Zielinska-Skowronek, M., Schmidt, R. E. (2001). Lack of specific receptors for C-reactive protein on white blood cells. Eur. J. Immunol. **31**. 3475-3483.

Hunt, R. C., Riegler, R., Davis, A. A. (1989). Changes in glycosylation alter the affinity of the human transferring receptor for its ligand. Journal of Biological Chemistry. **264**. 16. 9643-9648

Hurlimann, J., Thorbecke, G. J., Hochwald, G. M. (1965). The liver as the site of C-reactive protein formation. The Journal of Experimental Medicine. **123**. 365-378.

Hutchcraft, C. L., Gewurz, H., Hansen, B., Dyck, R. F., Pepys, M. B. (1981). Agglutination of complement-coated erthyrocytes by serum amyloid P component. Journal of Immunology. **126**. 3. 1217-1219

Hutchcraft, C., Hansen, B., Harling, V. O., Gewurz, H. (1982). Binding reactivities of serum amyloid P-component. Annals of the New York Academy of Sciences. **389**. 449-450

Hutchinson, W. L., Hohenester, E., Pepys, M. B. (2000). Human serum amyloid P component is a single uncomplexed pentamer in whole serum. Molecular medicine. **6**. 6. 482-493.

Iijiman, M., Matsuda, Y., Osaki, T., Yoshida, R., Koshiba, T., Kawabata, S-I. (2005). Cuticular chitin-binding proteins from horseshoe crab exoskeleton are involved in transglutaminase-dependent cross-linking. Development of protein research. 2P-255

Iwaki, D., Kawabata, S., Miura, Y., Kato, A., Armstrong, P. B., Quigley, J. P., Nielsen, K. L., Dolmer, K., Sottrup-Jensen, L., Iwanaga, S. (1996). Molecular cloning of *Limulus* α2-macroglobulin. Eur. J. Biochem. **242**. 822-831.

Iwaki, D., Osaki, T., Mizunoe, Y., Wai, S. N., Iwanaga, S., Kawabata, S. (1999). Functional and structural diversities of C-reactive proteins present in horseshoe crab haemolymph plasma. Eur. J. Biochem. **264**. 314-326.

Iwanaga, S., Miyata, T., Tokunaga, F., Muta, T. (1992). Molecular mechanism of hemolymph clotting system in *Limulus*. Thromb. Res. **68**. 1. 1-32

Iwanaga, S., Kawabata, S., Muta, T. (1998). New types of clotting factors and defense molecules found in horseshoe crab hemolymph: their structures and functions. J. Biochem. **123**. 1. 1-15

Jiang, H., Lint, T. F., Gewurz, H. (1991a). Defined chemically cross-linked oligomers of human C-reactive protein: characterization and reactivity with the complement system. Immunology. **74**. 725-731.

Jiang, H. X., Siegel, J. N., Gewurz, H. (1991b). Binding and complement activation by C-reactive protein via the collagen-like region of C1q and inhibition of these reactions by monoclonal antibodies to C-reactive protein and C1q. J. Immunol. **146**. 7. 2324-2330

Jinbo,T., Hayashi, S., Iguchi, K., Shimizu, M., Matsumoto, T., Naiki, M., Yamamoto, S. (1998). Development of monkey C-reactive protein (CRP) assay methods. Veterinary Immunology and Immunopathology. **61**. 195-202

Johansson, M. W. (1999). Cell adhesion molecules in invertebrate immunity. Developmental and Comparative Immunology. **23**. 303-315.

Kairies, N., Beisel, H-G., Fuentes-Prior, P., Tsuda, R., Muta, T., Iwanaga, S., Bode, W., Huber, R., Kawabata, S-I. (2001). The 2.0Å crystal structure of tachylectin 5A provides evidence for the common origin of the innate immunity and blood coagulation systems. PNAS. **98**. 24. 13519-13524.

Kalaria, R. N., Kroon, S. N. (1992). Complement inhibitor C4-binding protein in amyloid deposits containing serum amyloid P in Alzheimer's disease. Biochem. Biophys. Res. Commun. **186**. 1. 461-466

Kanost, M. R. (1999). Serine proteinase inhibitors in arthropod immunity. Developmental and Comparative Immunology. **23**. 291-301.

Kaplan, M. H., Volanakis, J. E. (1974). Interaction of C-reactive protein complexes with the complement system. I. The Journal of Immunology. **112**. 6. 2135-2147

Kawabata, S., Saito, T., Saeki, K., Okino, N., Mizutani, A., Toh, Y., Iwanaga, S. (1997). cDNA cloning, tissue distribution, and subcellular localization of horseshoe crab big defensin. Biol. Chem. **378**. 3-4. 289-292

Kawabata, S-I., Tsuda, R. (2002). Molecular basis of non-self recognition by the horseshoe crab tachylectins. Biochimica et Biophysica Acta. **1572**. 414-421

Kempka, G., Roos, P. H., Kolb-Bachofen, V. (1990). A membrane-associated form of C-reactive protein is the galactose-specific particle receptor on rat liver macrophages. J. Immunol. **144**. 3. 1004-1009

Kew, R. R., Hyers, T. M., Webster, R. O. (1990). Human C-reactive protein inhibits neutrophil chemotaxis *in vitro*: possible implications for the adult respiratory distress syndrome. J. Lab. Clin. Med. **115**. 3. 339-345

Kilpatrick, J. M., Volanakis, J. E. (1985). Opsonic properties of C-reactive protein. Stimulation by phorbol myristate acetate enables human neutrophils to phagocytize C-reactive protein-coated cells. J. Immunol. **134**. 5. 3364-3370

Kindmark, C-O. (1971). Stimulating effect of C-reactive protein on phagocytosis of various species of pathogenic bacteria. Clin. Exp. Immunol. **8**. 941-948.

Kinoshita, C. M., Ying, S-C., Hugli, T. E., Siegel, J. N., Potempa, L. A., Jiang, H., Houghten, R. A., Gewurz, H. (1989). Elucidation of a protease-sensitive site involved in the binding of calcium to C-reactive protein. Biochemistry. **28**. 9840-9848.

Kinoshita, C. M., Gewurz, A. T., Siegel, J. N., Ying, S-C, Hugli, T. E., Coe, J. E., Gupta, R. K., Huckman, R., Gewurz, H. (1992). A protease-sensitive site in the proposed Ca^{2+} binding region of human serum amyloid P component and other pentraxins. Protein Science. **1**. 700-709.

Kishore, U., Ghai, R., Greenhough, T. J., Shrive, A. K., Bonifati, D. M., Gadjeva, M. G., Waters, P., Kojouharova, M. S., Chakraborty, T., Agrawal, A. (2004). Structural and functional anatomy of the globular domain of complement protein C1q. Immunology Letters. **95**. 113-128

Kline, T. (1993). Handbook of affinity chromatography. Chromatographic Science series. Volume 63. Marcel Dekker, Inc.

Kodama, H., Yamada, F., Murai, T., Nakanishi, Y., Mikami, T., Izawa, H. (1989). Activation of trout macrophages and production of CRP after immunization with *Vibrio angillarum*. Developmental and Comparative Immunology. **13**. 2. 123-132.

Kokubun, M., Imafuku, Y., Okada, M., Ohguchi, Y., Ashikawa, T., Yamada, T., Yoshida, H. (2005). Serum amyloid A (SAA) concentration varies among rheumatoid arthritis patients estimated by SAA/CRP ratio. Clinica Chimica Acta. **360**. 97-102.

Kolb-Bachofen, V., Puchta-Teudt, N., Egenhofer, C. (1995). Expression of membrane-associated C-reactive protein by human monocytes: indications for a selectin-like activity participating in adhesion. Glycoconj. J. **12**. 2. 122-127

Kottgen, E., Hell, B., Kage, A., Tauber, R. (1992). Lectin specificity and binding characteristics of human C-reactive protein. The Journal of Immunology. **149**. 2. 445-453

Kubak, B. M., Gewurz, H., Potempa, L. A. (1988). Identification of multiple forms of the P component of amyloid isolated from human serum. Int. Arch. Allergy. Appl. Immunol. **87**. 2. 194-203

Kubak, B. M., Potempa, L. A., Anderson, B., Mahklouf, S., Venegas, M., Gewurz, H., Gewurz, A. T. (1988). Evidence that serum amyloid P component binds to mannose-terminated sequences of polysaccharides and glycoproteins. Mol. Immunol. **25**. 9. 851-858

Kushner, I., Kaplan, M. H. (1961). Studies of acute phase proteins. I. Journal of Experimental Medicine. **114**. 961-979.

Kushner, I., Rakita, L., Kaplan, M. H. (1963). Studies of acute-phase protein. II. Localization of Cx-reactive protein in heart in induced myocardial infarction in rabbits. Journal of Clinical Investigation. **42**. 2. 286-292

Kushner, I., Somerville, J. A. (1970). Estimation of the molecular size of C-reactive protein and Cx-reactive protein in serum. Biochimica et Biophysica Acta. **207**. 105-114

Kushner, I., Feldmann, G. (1978). Control of the acute phase response. Journal of Experimental Medicine. **148**. 466-477.

Kushner, I. (1982). The phenomenon of the acute phase response. Annals of the New York Academy of Sciences. **389.** 39-45

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., Ben-Tal, N. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Research. **33**. Web Server Issue. W299-302

Landsmann, P., Rosen, O., Pontet, M., Pras, M., Levartowsky, D., Shephard, E. G., Fridkin, M. (1994). Binding of human serum amyloid P component (hSAP) to human neutrophils. Eur. J. Biochem. **223**. 805-811.

Lasson, A., Göransson, J. (1999). No microheterogenous changes of plasma C-reactive protein found in man during various diseases. Scand. J. Clin. Lab. Invest. **59**. 4. 293-304

Le, P. T., Muller, M. T., Mortensen, R. F. (1982a). Acute phase reactants of mice. I. Isolation of serum amyloid P component (SAP) and its induction by a monokine. J. Immunol. **129**. 2. 665-672

Le, P. T., Muller, M. T., Mortensen, R. F. (1982b). Acute phase reactants of mice: isolation and induction of SAP and CRP. Annals of the New York Academy of Sciences. **389.** 454-455

Le Saux, A., Ng, P. M. L., Koh, J. J. Y., Low, D. H. P., Leong, G. E-L., Ho, B., Ding, J. K. (2008). The macromolecular assembly of pathogen-recognition receptors is impelled by serine proteases, via their complement control protein modules. J. Mol. Biol. **377**. 902-913.

Lee, R. T., Takagahara, I., Lee, Y. C. (2002). Mapping the binding areas of human C-reactive protein for phosphorylcholine and polycationic compounds. The Journal of Biological Chemistry. **277**. 1. 225-232.

Lee, R. T., Lee, Y. C. (2003). Carbohydrate-binding properties of human neo-CRP and its relationship to phosphorylcholine-binding site. Glycobiology. **13**. 1. 11-21.

Lei, K-J., Liu, T., Zon, G., Soravia, E., Liu, T-Y., Goldman, N. D. (1985). Genomic DNA sequence for human C-reactive protein. The Journal of Biological Chemistry. **260**. 24. 13377-13383.

Levin, J., Bang, F. B. (1964). The role of endotoxin in the extracellular coagulation of *Limulus* blood. Bulletin of the Johns Hopkins Hospital. **115**. 265-274

Levin, J. (1985). The history of the development of the *Limulus* amoebocyte lysate test. Prog. Clin. Biol. Res. **189**. 3-30

Li, S-P., Liu, T-Y., Goldman, N. D. (1990). Cis-acting elements responsible for Interleukin-6 inducible C-reactive protein gene expression. The Journal of Biological Chemistry. **265**. 7. 4136-4142.

Li, S-P., Goldman, N. D. (1996). Regulation of human C-reactive protein gene expression by two synergistic IL-6 responsive elements. Biochemistry. **35**. 9060-9068

Li, Y., Ng, M. L., Ho, B., Ding, J. L. (2007). A female-specific pentraxin, CrOctin, bridges pattern recognition receptors to bacterial phosphoethanolamine. Eur. J. Immunol. **37**. 3477-3488.

Lin, L., Liu, T-Y. (1993). Isolation and characterization of C-reactive protein (CRP) cDNA and genomic DNA from *Xenopus laevis*. The Journal of Biological Chemistry. **268**. 9. 6809-6815.

Liu, T-Y., Gotschilch, E. C. (1963). The chemical composition of Pneumococcal C-polysaccharide. The Journal of Biological Chemistry. **238**. 6. 1928-1934.

Liu, T., Lin, Y., Cislo, T., Minetti, C. A. S. A., Baba, J. M. K., Liu, T-Y. (1991). Limunectin. The Journal of Biological Chemistry. **266**. 22. 14813-14821

Liuzzo, G., Biasucci, L. M., Gallimore, J. R., Grillo, R. L., Rebuzzi, A. G., Pepys, M. B., Maseri, A. (1994). The prognostic value of C-reactive protein and serum amyloid A protein in severe unstable angina. The New England Journal of Medicine. **331**. 417-424.

Loveless, R. W., Floyd-O'Sullivan, G., Raynes, J. G., Yuen, C. T., Feizi, T. (1992). Human serum amyloid P is a multispecific adhesive protein whose ligands include 6-phosphorylated mannose and the 3-sulphated saccharides galactose, N-acetylgalactosamine, and glucuronic acid. The EMBO Journal. **11**. 3. 813-819

Lund, V., Olafsen, J. A. (1998a). A comparative study of pentraxin-like proteins in different fish species. Developmental and Comparative Immunology. **22**. 2. 185-194

Lund, V., Olafsen, J. A. (1998b). A typical phosphorylcholine-reactive protein from Atlantic salmon, *Salmo salar* L. Comparative Biochemistry and Physiology Part B. **119**. 471-477.

Lund, V., Olafsen, J. A. (1999). Changes in serum concentration of a serum amyloid P-like pentraxin in Atlantic salmon, *Salmo salar* L., during infection and inflammation. Developmental and Comparative Immunology. **23**. 61-70.

MacCarthy, E. M., Burns, I., Irnazarow, I., Polwart, A., Greenhough, T. J., Shrive, A. K., Hoole, D. (2008). Serum CRP-like protein profile in common carp *Cyprinus carpio* challenged with *Aeromonas hydrophila* and *Escherichia coli* lipopolysaccharide. Developmental and Comparative Immunology. **32**. 1281-1289

MacLeod, C. M., Avery, O. T. (1940a). The occurrence during acute infections of a protein not normally present in the blood. II. Journal of Experimental Medicine. **73**. 183-190.

MacLeod, C. M., Avery, O. T. (1940b). The occurrence during acute infections of a protein not normally present in the blood. III. Journal of Experimental Medicine. **73**. 191-200.

Mandal, C., Sinha, S., Mandal, C. (1999) Lectin like properties and differential sugar binding characteristics of C-reactive proteins purified from sera of normal and pollutant induced *Labeo rohita*. Glycoconjugate Journal. **16**. 741-750

Manolov, D. E., Röcker, C., Hombach, V., Nienhaus, G. U., Torzewski, J. (2004). Ultrasensitive confocal fluorescence microscopy of C-reactive protein interacting with FcγRIIa. Arterioscler. Thromb. Vac. Biol. **24**. 2372-2377.

Mantovani, A., Garlanda, C., Doni, A., Bottazzi, B. (2008). Pentraxin in innate immunity: from C-reactive protein to the long pentraxin PTX3. J. Clin. Immunol. **28**. 1-13.

Mantzouranis, E. C., Dowton, S. B., Whitehead, A. S., Edge, M. D., Bruns, G. A. P., Colten, H. R. (1985). Human serum amyloid P component. The Journal of Biological Chemistry. **260**. 12. 7752-7756.

Marchalonis, J., Edelman, G. M. (1965). Phylogenetic origins of antibody structure. I. Multichain structure of immunoglobulins in the smooth dogfish (*Mustelus canis*). Journal of Experimental Medicine. **122**. 3. 601-618

Marchalonis, J., Edelman, G. M. (1966). Polypeptide chains of immunoglobulins from the smooth dogfish (*Mustelus canis*). Science. **154**. 756. 1567-1568

Marchalonis, J., Edelman, G. M. (1968). Isolation and characterisation of a hemagglutinin from *Limulus polyphemus*. J. Mol. Biol. **32**. 453-465.

Marnell, L., Mold, C., Volzer, M. A., Burlingame, R. W., Du Clos, T. W. (1995). C-reactive protein binds to FcγRI in transfected COS cells. J. Immunol. **155**. 4. 2185-2193

Marnell, L., Mold, C., Du Clos, T. W. (2005). C-reactive protein: ligands, receptors and role in inflammation. Clinical Immunology. **117**. 104-111.

Martz, E. (2002). Protein Explorer: easy yet powerful macromolecular visualization. Trends in Biochemical Sciences. **27**. 2. 107-109

Matsuda, Y., Osaki, T., Koshiba, T., Kawabata, S-I. (2005). Characterization of a recombinant caraxin-1, a horseshoe crab cuticualr protein for transglutaminase-dependent cross-linking. Development of protein research. Seikagaku. **77**. 8. 783

Matsuda, Y., Koshiba, T., Osaki, T., Suyuma, H., Arisaka, F., Toh, Y., Kawabata, S-I. (2007a). An arthropod cuticular chitin-binding protein endows injured sites with transglutaminase-dependent mesh. The Journal of Biological Chemistry. **282**. 52. 37316-37324

Matsuda, Y., Osaki, T., Hashii, T., Koshiba, T., Kawabata, S-I. (2007b). A cysteine-rich protein from an arthropod stabilizes clotting mesh and immobilizes bacteria at injury sites. The Journal of Biological Chemistry. **282**. 46. 33545-33552

Maudsley, S., Pepys, M. B. (1987). Immunochemical cross-reactions between pentraxins of different species. Immunology. **62**. 17-22.

McGrath, F. D., Brouwer, M. C., Arlaud, G. J., Daha, M. R., Hack, C. E., Roos, A. (2006). Evidence that complement protein C1q interacts with C-reactive protein through its globular head region. J. Immunol. **176**. 5. 2950-2957

Melchior, R., Quigley, J. P., Armstrong, P. B. (1995). α2-macroglobulin-mediated clearance of proteases from the plasma of the American horseshoe crab, *Limulus polyphemus*. The Journal of Biological Chemistry. **270**. 22. 13496-13502.

Merril, C. R., Goldman, D., Sedman, S. A., Ebert, M. H. (1981). Ultrasensitive stain for proteins in polyacrylamide gels shows regional variation in cerebrospinal fluid proteins. Science. **211**. 4489. 1437-1438

Mineo, C., Gormley, A. K., Yuhanna, I. S., Osborne-Lawrence, S., Gibson, L. L., Hahner, L., Shohet, R. V., Black, S., Salom, J. E., Samols, D., Karp, D. R., Thomas, G. D., Shaul, P. W. (2005). FcγRIIB mediates C-reactive protein inhibition of endothelial NO synthase. Circulation Research. **97**. 1124-1131

Minetti, C. A. S. A., Lin, Y., Cislo, T., Liu, T-Y. (1991). Purification and characterization of an endotoxin-binding protein with protease inhibitory activity from *Limulus* amebocytes. The Journal of Biological Chemistry. **266**. 31. 20773-20780.

Mitra, S., Bhattacharya, S. (1992). Purification of C-reactive protein from *Channa punctatus* (Bloch). Indian J. Biochem. Biophys. **29**. 6. 508-511

Miyagawa, N., Okamoto, Y., Miyagawa, S. (1988). Effect of C-reactive protein on peritoneal macrophages. I. Human C-reactive protein inhibits migration of guinea pig peritoneal macrophages. Microbiol. Immunol. **32**. 7. 709-719

Miyazawa, K., Inoue, K. (1990). Complement activation induced by human C-reactive protein in mildly acidic conditions. J. Immunol. **145**. 2. 650-654

Mohan, S., Thambi Doran, D., Srimal, S., Bachhawat, B. K. (1982). Binding studies of a sialic acid-specific lectin from the horseshoe crab *Carcinoscorpius rotundicauda* with various sialoglycoproteins. Biochem. J. **203**. 253-261.

Mold, C., Nakayama, S., Holzer, T. J., Gewurz, H., Du Clos, T. W. (1981a). C-reactive protein is protective against *Streptococcus pneumoniae* infection in mice. Journal of Experimental Medicine. **154**. 1701-1708.

Mold, C., Rodgers, C. P., Richards, R. L., Alving, C. R., Gewurz, H. (1981b). Interaction of C-reactive protein with liposomes. III. Membrane requirements for binding. J. Immunology. **126**. 3. 856-860

Mold, C., Edwards, K. M., Gewurz, H. (1982). Effect of C-reactive protein on the complement-mediated stimulation of human neutrophils by *Streptococcus pneumoniae* serotypes 3 and 6. Infection and Immunity. **37**. 3. 987-992

Mold, C., Gresham, H. D., Du Clos, T. W., (2001). Serum amyloid P component and C-reactive protein mediate phagocytosis through murine FcγRs. The Journal of Immunology. **166**. 1200-1205.

Mold, C., Baca, R., Du Clos, T. W. (2002). Serum amyloid P component and C-reactive protein opsonize apoptotic cells for phagocytosis through Fcγ receptors. J. Autoimmunity. **19**. 3. 147-154

Mold, C., Du Clos, T. W. (2006). C-reactive protein increases cytokine response to *Streptococcus pneumoniae* through interactions with Fcγ receptors. The Journal of Immunology. **176**. 7598-7604

Morley, J. J., Kushner, I. (1982). Serum C-reactive protein levels in disease. Annals of the New York Academy of Sciences. **389**. 406-416

Morrow, D. A., Rifai, N., Antman, E. M., Weiner, D. L., McCabe, C. H., Cannon, C. P., Braunwald, E. (1998). C-reactive protein is a potent predictor of mortality independently of and in combination with troponin T in acute phase coronary syndromes: a TIMI 11A study. Thrombolysis in Myocardial Infarction. J. AM. Coll. Cardiol. **31**. 7. 1460-1465

Mortensen, R. F., Osmand, A. P., Gewurz, H. (1975). Effects of C-reactive protein on the lymphoid system. I. Journal of Experimental Medicine. **141**. 821-839.

Mortensen, R. F., Osmand, A. P., Lint, T. F., Gewurz, H. (1976). Interaction of C-reactive protein with lymphocytes and monocytes: complement-dependent adherence and phagocytosis. J. Immunol. **117**. 3. 774-781

Mortensen, R. F., Duszkiewicz, J.A. (1977). Mediation of CRP-dependent phagocytosis through mouse macrophage Fc-receptors. J. Immunol. **119**. 5. 1611-1616

Mortensen, R. F., Zhong, W. (2000). Regulation of phagocytic leukocyte activities by C-reactive protein. Journal of Leukocyte Biology. **67**. 495-500.

Motie, M., Brockmeier, S., Potempa, L. A. (1996). Binding of model soluble immune complexes to modified C-reactive protein. J. Immunol. **156**. 11. 4435-4441.

Müller, h., Fehr, J. (1986). Binding of C-reactive protein to human polymorphonuclear leukocytes: evidence for association of binding sites with Fc receptors. J. Immunol. **136**. 6. 2202-2207

Murai, T., Kodama, H., Naiki, M., Mikami, T., Izawa, H. (1990). Isolation and characterization of rainbow trout C-reactive protein. Developmental and Comparative Immunology. **14**. 1. 49-58.

Murata, M., Onuma, M., Kodama, H. (1994). Isolation and characterization of rainbow trout (*Oncorhynchus mykiss*) serum amyloid P component (SAP). J. Vet. Med. Sci. **56**. 4. 661-665

Murata, M., Kodama, H., Onuma, M. (1995). Characterization of rainbow trout C-polysaccharide binding proteins. J. Vet. Med. Sci. **57**. 3. 419-425

Muta, T., Hahimoto, R., Miyata, T., Nishimura, H., Toh, Y., Iwanaga, S. (1990). Proclotting enzyme from horseshoe crab hemocytes. The Journal of Biological Chemistry. **265**. 36. 22426-22433.

Muta, T., Miyata, T., Misumi, Y., Tokunaga, F., Nakamura, T., Toh, Y., Ikehara, Y., Iwanaga, S. (1991). *Limulus* Factor C. The Journal of Biological Chemistry. **266**. 10. 6554-6561.

Myles, D. A. A., Bailey, S., Rule, S. A., Jones, G. R., Greenhough, T. J. (1990). Preliminary crystallographic study of C-reactive protein from *Limulus polyphemus*. Journal of Molecular Biology. **213**. 223-225.

Nagpurkar, A., Mookerjea, S. (1981). A novel phosphorylcholine-binding protein from rat serum and its effect on heparin-lipoprotein complex formation in the presence of calcium. The Journal of Biological Chemistry. **256**. 14. 7440-7448.

Nakamura, T., Furunaka, H., Miyata, T., Tokunaga, F., Muta, T., Iwanaga, S. (1988). Tachyplesin, a class of antimicrobial peptide from the hemocytes of the Horseshoe crab (*Tachypleus tridentatus*). The Journal of Biological Chemistry. **263**. 32. 16709-16713.

Nakanishi, Y., Kodama, H., Murai, T., Mikami, T., Izawa, H. (1991). Activation of rainbow trout complement by C-reactive protein. Am. J. Vet. Res. **52**. 3. 397-401

Narkates, A. J., Volanakis, J. E. (1982) C-reactive protein binding specificities: artificial and natural phospholipids bilayers. Annals of the New York Academy of Sciences. **389**. 172-182.

Nakayama, S., Mold, C., Gewurz, H., Du Clos, T. W. (1982). Opsonic properties of C-reactive protein *in vivo*. J. Immunol. **128**. 6. 2435-2438

Nakayama, S., Gewurz, H., Holzer, T., Du Clos, T. W., Mold, C. (1983). The role of the spleen in the protective effect of C-reactive protein in *Streptococcus pneumoniae* infection. Clin. Exp. Immunol. **54**. 319-326

Ng, P. M. L., Jin, Z., Tan, S. S. H., Ho, B., Ding, J. K. (2004). C-reactive protein: a predominant LPS-binding acute phase protein responsive to *Pseudomonas* infection. Journal of Endotoxin Research. **10**. 3. 164-174.

Ng, P. M. L., Le Saux, A., Lee, C. M., Tan, N. S., Lu, J., Thiel, S., Ho, B., Ding, J. L. (2007). C-reactive protein collaborates with plasma lectins to boost immune response against bacteria. The EMBO Journal. **26**. 3431-3440.

Nguyen, N. Y., Suzuki, A., Cheng, S-M., Zon, G., Liu, T-Y. (1986a). Isolation and characterisation of *Limulus* C-reactive protein genes. The Journal of Biological Chemistry. **261**. 22. 10450-10455.

Nguyen, N. Y., Suzuki, A., Boykins, R. A., Liu, T-Y. (1986b). The amino acid sequence of *Limulus* C-reactive protein. The Journal of Biological Chemistry. **261**. 22. 10456-10465.

Nishiguchi, S., Maeda, S., Araki, S., Shimada, K. (1988). Structure of the mouse serum amyloid P component gene. Biochemical and Biophysical research communications. **155**. 3. 1366-1373

Noursadeghi, M., Bickerstaff, M. C. M., Gallimore, J. R., Herbert, J., Cohen, J., Pepys, M. B. (2000). Role of serum amyloid P component in bacterial infection: protection of the host or protection of the pathogen. PNAS. **97**. 26. 14584-14589.

Nybo, M., Hackler, R., Kold, B., Holm Nielsen, E., Steinmetz, A., Svehag, S-E. (1998). Isoforms of murine and human serum amyloid P component. Scand. J. Immunol. **48**. 350-356.

Ohnishi, S., Maeda, S., Nishiguchi, S., Arao, T., Shimida, K. (1988). Structure of the mouse C-reactive protein gene. Biochemical and Biophysical Research Communications. **156**. 2. 814-822

Oliveira, E. B., Gotschlich, E. C., Liu, T-Y. (1979). Primary structure of human C-reactive protein. The Journal of Biological Chemistry. **254**. 2. 489-502.

Oliveira, E. B., Gotschlich, E. C., Liu, T-Y. (1980). Comparative studies on the binding properties of human and rabbit C-reactive protein. The Journal of Immunology. **124**. 3. 1396-1402.

Oppenheim, J. D., Nachbar, M. S., Salton, M. R. J., Aull, F. (1974). Purification of a hemagglutinin from *Limulus polyphemus* by affinity chromatography. Biochemical and Biophysical Research Communications. **58**. 4. 1127-1134.

Osmand, A. P., Mortensen, R. F., Siegel, J., Gewurz, H. (1975). Interactions of C-reactive protein with the complement system. III. The Journal of Experimental Medicine. **142**. 1065-1077

Osmand, A. P., Friedenson, B., Gewurz, H., Painter, R. H., Hofmann, T., Shelton, E. (1977). Characterization of C-reactive protein and the complement subcomponent C1t as homolgous proteins displaying cyclic pentameric symmetry (pentraxins). Proc. Natl. Acad. Sci. USA. **74**. 2. 739-743.

Osmand, A. P., Gewurz, H., Friedenson, B. (1977). Partial amino-acid sequences of human and rabbit C-reactive proteins: homology with immunoglobulins and histocompatibility antigens. Proc. Natl. Acad. Sci. USA. **74**. 3. 1214-1218.

Ozaki, A., Ariki, S., Kulman, J., Koshiba, T., Kawabata, S-I. (2005). Identification of complement factors in horseshoe crab *Tachypleus tridentatus*. Development of protein research. 2P-254

Padilla, N. D., Bleeker, W. K., Lubbers, Y., Rigter, G. M. M., Van Mierlo, G. J., Daha, M. R., Hack, C. E. (2003). Rat C-reactive protein activates the autologous complement system. Immunology. **109**. 564-571

Page, R. D. M. (1996). Treeview: An application to display phylogenetic trees on personal computers. CABIOS Applications Note. **12**. 4. 357-358.

Painter, R. H., De Escallón, I, Massey, A., Pinteric, L, Stern, S. B. (1982). The structure and binding characteristics of serum amyloid protein. Annals of the New York Academy of Sciences. **389**. 199-213

Parish, W. E. (1976). Studies on vasculitis. VII. C-reactive protein as a substance perpetuating chronic vasculitis. Occurrence in lesions and concentrations in sera. Clin. Allergy. **6**. 6. 543-550

Parton, A., Forest, D., Kobayashi, H., Dowell, L., Bayne, C., Barnes, D. (2007). Cell and molecular biology of SAE, a cell line from the spiny dogfish shark, *Squalus acanthias*. Comparative Biochemistry and Physiology C. Toxicology and Pharmacology. **145**. 1. 111-119.

Pasceri, V., Willerson, J. T., Yeh. E. T. H. (2000). Direct proinflammatory effect of C-reactive protein on human endothelial cells. Circulation. **102**. 2165-2168.

Patterson, L. T., Higginbotham, R. D. (1965). Mouse C-reactive protein and endotoxin-induced resistance. Journal of Bacteriology. **90**. 6. 1520-1524

Paul, I., Mandal, C., Mandal, C. (1998). Effect of environmental pollutants on the C-reactive protein of a freshwater major carp, *Catla catla*. Developmental and Comparative Immunology. **22**. 5/6. 519-532.

Pepys, M. B., Dash, A. C., Munn, E. A., Feinstein, A., Skinner, M., Cohen, A. S., Gewurz, H., Osmand, A. P., Painter, R. H. (1977). Isolation of amyloid P component (protein AP) from normal serum as a calcium –dependent binding protein. The Lancet. **1**. 1029-1031.

Pepys, M. B., Dash, A. C., Fletcher, T. C., Richardson, N., Munn, E. A., Feinstein, A. (1978). Analogues in other mammals and in fish of human plasma proteins, C-reactive protein and amyloid P component. Nature. **273**. 168-170.

Pepys, M. B. (1979). Isolation of serum amyloid P component (protein SAP) in the mouse. Immunology. **37**. 637-641

Pepys, M. B., Baltz, M., Gomer, K., Davies, A. J. S., Doenhoff, M. (1979a). Serum amyloid P-component is an acute-phase reactant in the mouse. Nature. **278**. 259-261

Pepys, M. B., Dyck, R. F., De Beer, F. C., Skinner, M., Cohen, A. S. (1979b). Binding of serum amyloid P-component (SAP) by amyloid fibrils. Clin. Exp. Immunol. **38**. 284-293

Pepys, M. B., Butler, P. J. G. (1987). Serum amyloid P component is the major calcium-dependent specific DNA binding protein of the serum. Biochemical and Biophysical Research Communications. **148**. 1. 308-313

Pepys, M. B., Booth, S. E., Tennent, G. A., Butler, P. J. G., Williams, D. G. (1994). Binding of pentraxins to different nuclear structures: C-reactive protein binds to small nuclear ribonucleoprotein particles, serum amyloid P component binds to chromatin and nucleoli. Clin Exp Immunol. **97**. 152-157

Pepys, M. B., Rademacher, T. W., Amatayakul-Chantler, S., Williams, P., Noble, G. E., Hutchinson, W. L., Hawkins, P. N., Nelson, S. R., Gallimore, J. R., Herbert, J., Hutton, T., Dwek, R. A. (1994). Human serum amyloid P component is an invariant constituent of amyloid P deposits and has a uniquely homogenous glycostructure. Proc. Natl. Acad. Sci. USA. **91**. 12. 5602-5606.

Pepys, M. B., Herbert, J., Hutchinson, W. L., Tennent, G. A., Lachmann, H. J., Gallimore, J. R., Lovat, L. B., Bartfai, T., Alanine, A., Hertel, C., Hoffmann, T., Jakob-Roetne, R., Norcross, R. D., Kemp, J. A., Yamamura, K., Suzuki, M., Taylor, G. W., Murray, S., Thompson, D., Purvis, A., Kolstoe, S., Wood, S. P., Hawkins, P. N. (2002). Targeted pharmacological depletion of serum amyloid P component for treatment of human amyloidosis. Nature. **417**. 254-259

Pepys, M. P., Hirchfield, G. M. (2003). C-reactive protein: a critical update. J. Clin. Invest. **111**. 1805-1812

Pepys, M. B. (2005). CRP or not CRP? That is the question. Arterioscler. Thromb. Vasc. Biol. **25**. 1091-1094

Pepys, M. B., Hawkins, P. N., Kahan, M. C., Tennent, G. A., Gallimore, J. R., Graham, D., Sabin, C. A., Zychlinsky, A., de Diego, J. (2005). Proinflammatory effects of bacterial recombinant human C-reactive protein are caused by contamination with bacterial products, not by C-reactive protein itself. Circulation Research. **97**. 97-103

Pepys, M. B., Hirschfield, G. M., Tennet, G. A., Gallimore, J. R., Kahan, M. C., Bellotti, V., Hawkins, P. N., Myers, R. M., Smith, M. D., Ploara, A., Cobb, A. J. A., Ley, S. V., Aquilina, J. A., Robinson, C. V., Sharif, I., Gray, G. A., Sabin, C. A., Jenvey, NM. C., Kolstoe, S. E., Thompson, D., Wood, S. P. (2006). Targeting C-reactive protein for the treatment of cardiovascular disease. Nature. **44**. 1217-1221

Pepys, M. B. (2008). C-reactive protein is neither a marker nor a mediator of atherosclerosis. Nat. Clin. Pract. Neprol. **4**. 5. 234-235

Perlmutter, L. S., Barrón, E., Myers, M., Saperia, D., Chui, H. C. (1995). Localization of amyloid P component in human brain: vascular staining patterns and association with Alzheimer's disease lesions. J. Comp. Neurol. **352**. 1. 92-105

Pied, S., Nussler, A., Pontet, M., Miltgen, F., Matile, H., Lambert, P-H., Mazier, D. (1989). C-reactive protein protects against pre-erythrocytic stages of malaria. Infection and Immunity. **57**. 1. 278-282

Pinteric, L., Painter, R. H. (1979). Electron microscopy of serum amyloid protein in the presence of calcium; alternative forms of assembly of pentagonal molecules in two-dimensional lattices. Can. J. Biochem. **57**. 6. 727-736

Pistole, T. G., Britko, J. L. (1978). Bactericidal activity of amebocytes from the horseshoe crab, *Limulus polyphemus*. Journal of Invertebrate Pathology. **31**. 3. 376-382.

Potempa, L. A., Siegel, J. N., Gewurz, H. (1981). Binding reactivity of C-reactive protein for polycations. II. The Journal of Immunology. **127**. 4. 1509-1514.

Potempa, L. A., Gewurz, H. (1983). Influence of heparin on interactions between C-reactive protein and polycations. Mol. Immunol. **20**. 5. 501-509

Potempa, L. A., Maldonado, B. A., Laurent, P., Zemel, E. S., Gewurz, H. (1983). Antigenic, electrophoretic and binding alterations of human C-reactive protein modified selectively in the absence of calcium. Mol. Immunol. **20**. 11. 1165-1175

Potempa, L. A., Siegel, J. N., Fiedel, B. A., Potempa, R. T., Gewurz, H. (1987). Expression, detection and assay of a neoantigen (Neo-CRP) associated with a free, human C-reactive protein subunit. Mol. Immunol. **24**. 5. 531-541

Potempa, L. A., Zeller, J. M., Fiedel, B. A., Kinoshita, C. M., Gewurz, H. (1988). Stimulation of human neutrophils, monocytes, and platelets by modified C-reactive protein (CRP) expressing a neoantigenic specificity. Inflammation. **12**. 4. 391-405

Prelli, F., Pras, M., Frangione, B. (1985). The primary structure of human tissue amyloid P component from a patient with primary idiopathic amyloidosis. The Journal of Biological Chemistry. **260**. 24. 12895-12898.

Pritchard, D. G., Volanakis, J. E., Slutsky, G. M., Greenblatt, C. L. (1985). C-reactive protein binds leishmanial excreted factors. Proc. Soc. Exp. Biol. Med. **178**. 3. 500-503

Quigley, J. P., Armstrong, P. B. (1983). An endopeptidase inhibitor, similar to mammalian α2-macroglobulin, detected in the haemolymph of an invertebrate, *Limulus polyphemus*. The Journal of Biological Chemistry. **258**. 13. 7903-7906.

Quigley, J. P., Ikai, A., Arakawa, H., Osada, T., Armstrong, P. B. (1991). Reaction of proteinase with α2-macroglobulin from the American horseshoe crab, *Limulus*. The Journal of Biological Chemistry. **266**. 29. 19426-19431.

Quigley, J., Misquith, S., Surolia, A., Srimal, S., Armstrong, P. (1994). Preliminary investigation of the molecular basis for the functional differences between the two pentraxins Limulin and C-reactive protein from the plasma of the American horseshoe crab, *Limulus polyphemus*. Biol. Bull. **187**. 229-230.

Raju, T. S. (2008). Terminal sugars of Fc glycans influence antibody effector functions of IgGs. Current Opinion in Immunology. **20**. 4. 471-478

Ramadan, M. A. M., Shrive, A. K., Holden, D., Myles, D. A. A., Volanakis, J. E., DeLucas, L. J., Greenhough, T. J. (2002). The three-dimensional structure of calcium depleted human C-reactive protein from perfectly twinned crystals. Acta Crystallographica Section **D58**. 992-1001.

Rassouli, M., Sambasivam, H., Azadi, P., Dell, A., Morris, H. R., Nagpurkar, A., Mookerjea, S., Murray, R. K. (1992). Derivation of the amino acid sequence of rat C-reactive protein from cDNA cloning with additional studies on the nature of its dimeric component. The Journal of Biological Chemistry. **267**. 5. 2947-2954.

Reid, K. B. M., Lowe, D. M., Porter, R. R. (1972). Isolation and characterization of C1q, a subcomponent of the first component of complement, from human and rabbit sera. Biochem. J. **130**. 749-763.

Reid, K. B. M., Sim, R. B., Faiers, A. P. (1977). Inhibition of the reconstitution of the haemolytic activity of the first component of human complement by a pepsin-derived fragment of subcomponent C1q. Biochem. J. **161**. 239-245.

Richards, R. L., Gewurz, H., Osmand, A. P., Alving, C. R. (1977). Interactions of C-reactive protein and complement with liposomes. Proc. Natl. Acad. Sci. USA. **74**. 12. 5672-5676.

Richards, R. L., Gewurz, H., Siegel, J., Alving, C. R. (1979). Interactions of C-reactive protein and complement with liposomes. II. Influence of membrane composition. J. Immunol. **122**. 4. 1185-1189

Richardson, M. D., Gray, C. A., Shankland, G. S. (1991). Opsonic effect of C-reactive protein on phagocytosis and intracellular killing of virulent and attenuated strains of *Candida albicans* by human neutrophils. FEMS Microbiol. Immunol. **3**. 6. 341-344

Richardson, M. D., Shankland, G. S., Gray, C. A. (1991). Opsonozing activity of C-reactive protein in phagocytosis of *Aspergillus fumigatus condidia* by human neutrophils. Mycoses. **34**. 3-4 141-143

Ridker, P. M., Hennekens, C. H., Buring, J. E., Rifai, N. (2000). C-reactive protein and other markers of inflammation in the prediction of cardiovascular disease in women. The New England Journal of Medicine. **342**. 836-843

Ridker, P. M., Cannon, C. P., Morrow, D., Rifai, N., Rose, L. M., McCabe, C. H., Pfeffer, M. A., Braunwald, E. (2005). C-reactive protein levels and outcomes after statin therapy. The New England Journal of Medicine. **352**. 1. 20-28

Riley, R. F., Coleman, M. K. (1970). Isolation of C-reactive proteins of man, monkey, rabbit, and dog by affinity chromatography phosphorylated cellulose. Clinica Chimica Acta. **30**. 2. 483-496.

Robey, F. A., Liu, T-Y. (1981). Limulin: A C-reactive protein from *Limulus polyphemus*. The Journal of Biological Chemistry. **256**. 2. 969-975.

Robey, F. A., Tanaka, T., Liu, T-Y. (1983). Isolation and characterization of two major serum proteins from the Dogfish, *Mustelus canis*, C-reactive protein and amyloid P component. The Journal of Biological Chemistry. **258**. 6. 3889-3894.

Robey, F. A., Liu, T-Y. (1983). Synthesis and use of new spin labelled derivatives of phosphorylcholine in a comparative study of human, dogfish, and *Limulus* C-reactive proteins. The Journal of Biological Chemistry. **258**. 6. 3895-3900.

Robey, F. A., Jones, K. D., Tanaka, T., Liu, T-Y. (1984). Binding of C-reactive protein to chromatin and nucleosome core particles. The Journal of Biological Chemistry. **259**. 11. 7311-7316.

Robey, F. A., Ohura, K., Futaki, S., Fujii, N., Yajima, H., Goldman, N., Jones, K. D., Wah, S. (1987). Proteolysis of human C-reactive protein produces peptides with potent immunomodulating activity. The Journal of Biological Chemistry. **262**. 15. 7053-7057.

Roche, A-C., Monsigny, M. (1974). Purification and properties of Limulin: a lectin (agglutinin) from haemolymph of *Limulus polyphemus*. Biochimica et Biophysica Acta. **371**. 242-254.

Roche, A. C., Schauer, R., Monsigny, M. (1975). Protein-sugar interactions. Purification by affinity chromatography of Limulin, an N-acyl-neuraminidyl-binding protein. FEBS lett. **57**. 3. 245-249

Röcker, C., Manalov, D. E., Kuzmenkina, E. V., Tron, K., Slatosch, H., Torzewski, J., Nienhaus, G. U. (2007). Affinity of C-reactive protein toward FcγRI is strongly enhanced by the γ-chain. The American Journal of Pathology. **170**. 2. 755-763

Rodriguez, J. A., Bodman-Smith, K. B., Raynes, J. G. (2004). Neutrophil responses to CRP are not dependent on polymorphism of human FcγRIIA (R131H). Clin Exp Immunol. **138**. 2. 271-277

Romero, I. R., Morris, C., Rodriguez, M., Du Clos, T. W., Mold, C. (1998). Inflammatory potential of C-reactive protein complexes compared to immune complexes. Clin. Immunol. Immunopathol. **87**. 2. 155-162

Roth, R. I., Su, D., Child, A. H., Wainwright, N. R., Levin, J. (1998). *Limulus* antilipopolysaccharide factor prevents mortality late in the course of endotoxemia. The Journal of Infectious Diseases. **177**. 388-394

Roux, K. H., Kilpatrick, J. M., Volanakis, J. E., Kearney, J. F. (1983). Localization of the phosphocholine-binding sites on C-reactive protein by immunoelectron microscopy. J. Immunol. **131**. 5. 2411-2415

Rubio, N., Sharp, P. M., Rits, M., Zahedi, K., Whitehead, A. S. (1993). Structure, expression, and evolution of guinea pig serum amyloid P component and C-reactive protein. J. Biochem. **113**. 277-284.

Rudnick, C. M., Dowton, S. B. (1993a). Serum amyloid P (female protein) of the Syrian hamster. The Journal of Biological Chemistry. **268**. 29. 21760-21769.

Rudnick, C. M., Dowton, S. B. (1993b). Serum amyloid P-component of the Armenian hamster; gene structure and comparison with structure and expression of the SAP gene from Syrian hamster. Scand. J. Immunol. **38**. 5. 445-450

Ryu, J., Lee, C. W., Shin, J. A., Parks, C. S., Kim, J. J., Park, S. J., Han, K. H. (2007). FcγRIIa mediates C-reactive protein-induced inflammatory responses of human vascular smooth muscle cells by activating NADPH oxidase 4. Cardiovasc Res. **75**. 3. 555-556.

Saito, T., Kawabata, S., Shigenaga, T., Takayenoki, Y., Cho, J., Nakajima, H., Hirata, M., Iwanaga, S. (1995). A novel big defensin identified in horseshoe crab hemocytes: isolation, amino acid sequence, and antibacterial activity. J. Biochem. **117**. 5. 1131-1137

Saitou, N., Nei, M. (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**. 4. 406-425.

Salonen, E-M., Vartio, T., Hedman, K., Vaheri, A. (1984). Binding of fibronectin by the acute phase reactant C-reactive protein. The Journal of Biological Chemistry. **259**. 3. 1496-1501.

Sambrook, J., Russell, D. W. (2001). Molecular Cloning. A laboratory manual. Volume 2. Third Edition. CSHL Press.

Samudzi, C. T., Nguyen, N. Y., Rubin, J. R. (1993). Crystallization and preliminary X-ray diffraction studies of Dogfish C-reactive protein. PROTEINS: Structure, Function, and Genetics. **15**. 100-102.

Saxena, U., Nagpurkar, A., Mookerjea, S. (1985). Contrasting effect of phosphorylcholine-binding protein from rat and rabbit on heparin-lipoprotein interaction: a role of sialic acid. Can. J. Biochem. Cell. Biol. **63**. 9. 1014-1021.

Schägger, H., von Jagow, G. (1987). Tricene-sodium dodecyl sulphate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100kDa. Analytical Biochemistry. **166**. 2. 368-379.

Schwalbe, R. A., Dahlback, B., Coe, J. E., Nelsestuen, G. L. (1992). Pentraxin fanily of proteins interact specifically with phosphorylcholine and/or phosphorylethanolamine. Biochemistry. **31**. 4907-4915.

Sekiguchi, K. (1988). Biology of horseshoe crabs. Science House Co., Ltd.

Shephard, E. G., Anderson, R., Strachan, A. F., Kühn, S. H., De Beer, F. C. (1986). CRP and neutrophils: functional effects and complex uptake. Clin. Exp. Immunol. **63**. 718-727

Shephard, E. G., Anderson, R., Beer, S. M., Jansen van Rensburg, C. E., De Beer, F. C. (1988). Neutrophil lysosomal degradation of human CRP: CRP –derived peptides modulate neutrophil function. Clin. Exp. Immunol. **73**. 139-145.

Shimizu, S., Ito, M., Niwa, M. (1977). Lectins in the hemolymph of Japanese horseshoe crab, *Tachypleus tridentatus*. Biochimica et Biophysica Acta. **500**. 71-79

Shirahama, T., Cohen, A. S. (1967). High-resolution electron microscopic analysis of the amyloid fibril.The Journal of Cell Biology. **33**. 679-708

Shrive, A. K., Cheetham, G. M. T., Holden, D., Myles, D. A. A., Turnell, W. G., Volanakis, J. E., Pepys, M. B., Bloomer, A. C., Greenhough, T. J. (1996). Three dimensional structure of human C-reactive protein. Nature Structural Biology. **3**. 4. 346-354.

Shrive, A. K., Metcalfe, A. M., Cartwright, J. R., Greenhough, T. J. (1999). C-reactive protein and SAP-like pentraxin are both present in *Limulus polyphemus* haemolymph: crystal structure of *Limulus* SAP. J. Mol. Biol. **290**. 997-1008.

Shuster, C. N., Barlow, R. B., Brockman, H. J. (2003). The American Horseshoe Crab. Published by Harvard University Press.

Siegel, J., Rent, R., Gewurz, H. (1974). Interactions of C-reactive protein with the complement system. I. The Journal of Experimental Medicine. **140**. 631-647.

Sinha, S., Mandal, C., Allen, A. K., Mandal, C. (2001). Acute phase response of C-reactive protein of *Labeo rohita* to aquatic pollutants is accompanied by the appearance of distinct molecular forms. Arch. Biochem. Biophys. **396**. 2. 139-150

Sipe, J. D., Vogel, S. N., Sztein, M. B., Skinner, M., Cohen, A. S. (1982). The role of interleukin 1 in acute phase serum amyloid A (SAA) and serum amyloid P (SAP) biosynthesis. Annals of the New York Academy of Sciences. **389**. 137-150

Siripont, J., Tebo, J. M., Mortensen, R.F. (1988). Receptor-mediated binding of the acute-phase reactant mouse serum amyloid P-component (SAP) to macrophages. Cell. Immunol. **117**. 2. 239-252

Skinner, M., Sipe, J. D., Yood, R. A., Shirahama, T., Cohen, A. S. (1982). Characterization of P component (AP) isolated from amyloidotic tissue: half-life studies of human and murine AP. Annals of the New York Academy of Sciences. **389**. 190-198

Skoglund, C., Wetterö, J., Skogh, T., Sjöwall, C., Tengvall, P., Bengtsson, T. (2008). C-reactive protein and C1q regulate platelet adhesion and activation on adsorbed immunoglobulin G and albumin. Immunol Cell Biol. **86**. 5. 466-474

Soelter, J., Uhlenbruck, G. (1986). The role of phosphate groups in the interaction of human C-reactive protein with galactan polysaccharides. Immunology. **58**. 139-144.

Sørensen, I. J., Andersen, O., Nielsen, E. H., Svehag, S. E. (1995). Native human serum amyloid P component is a single pentamer. Scand. J. Immunol. **41**. 3. 263-267

Sottrup-Jensen, L., Borth, W., Hall, M., Quigley, J. P., Armstrong, P. B. (1990). Sequence similarity between alpha 2-macroglobulin from the horseshoe crab, *Limulus polyphemus*, and proteins of the alpha 2-macroglobulin family from mammals. Comparative Biochemistry and Physiology B. **96**. 3. 621-625

Srimal, S., Dorai, D. T., Somasundaran, M., Bachhawat, B. K., Miyata, T. (1985). A new haemagglutinin from the amoebocytes of the horseshoe crab *Carcinoscorpius rotundicauda*. Biochem. J. **230**. 321-327.

Srinivasan, N., White, H. E., Emsley, J., Wood, S. P., Pepys, M. B., Blundell, T. L. (1994). Comparative analyses of pentraxins: implications for protomer assembly and ligand binding. Structure. **2**. 1017-1027.

Steel, D. M., and Whitehead, A. S. (1994). The major acute phase reactants: C-reactive protein, serum amyloid P component and serum amyloid A protein. Immunology Today. **15**. 2. 81-88

Stein, M-P., Edberg, J. C., Kimberley, R. P., Mangan, E. K., Bharadwaj, D., Mold, C., Du Clos, T. W. (2000). C-reactive protein binding to FcγRIIa on human monocytes and neutrophils is allele specific. The Journal of Clinical Investigation. **105**. 3. 369-376

Stein, M-P., Mold, C., Du Clos, T. W. (2000). C-reactive protein binding to murine leukocytes requires Fcγ receptors. The Journal of Immunology. **164**. 1514-1520.

Stenstad, T., Magnus, J. H., Syse, K., Husby, G. (1993). On the association between amyloid fibrils and glycosaminoglycans; possible interactive role of Ca^{2+} and amyloid P component. Clin. Exp. Immunol. **94**. 189-195.

Størmer, L. (1952). Phylogeny and taxonomy of fossil horseshoe crabs. Journal of Paleontology. **26**. 4. 630-639

Stults, N. L., Lee, Y. C., Hoppe, C. A., Kawaguchi, K., Kohda, S., Takagahara, I., Koishi, T., Liu, T-Y. (1987). Preparation of phosphorylcholine derivatives of bovine serum albumin and their application to the affinity chromatography of C-reactive protein. Analytical Biochemistry. **161**. 567-573.

Suresh, M. V., Singh, S. K., Agrawal, A. (2004). Interaction of calcium-bound C-reactive protein with fibronectin is controlled by pH. The Journal of Biological Chemistry. **279**. 50. 52552-52557.

Swanson, S. J., McPeek, M. M., Mortensen, R. F. (1989). Characteristics of the binding of human C-reactive protein (CRP) to laminin. J. Cell. Biochem. **40**. 1. 121-132

Swanson, S. J., Mortensen, R. F. (1990). Binding and immunological properties of a synthetic peptide corresponding to the phosphorylcholine-binding region of C-reactive protein. Mol. Immunol. **27**. 7. 679-687

Swanson, S. J., Lin, B. F., Mullenix, M. C., Mortensen, R. F. (1991) (a). A synthetic peptide corresponding to the phosphorylcholine (PC)- binding region of human C-reactive protein possesses the TEPC-15 myeloma PC-idiotype. J. Immunol. **146**. 5. 1596-1601

Swanson, S. J., Mullenix, M. C., Mortensen, R. F. (1991) (b). Monoclonal antibodies to the calcium-binding region peptide of human C-reactive protein alter its conformation. J. Immunol. **147**. 7. 2248-2252

Swanson, S. J., Christner, R. B., Mortensen, R. F. (1992). Human serum amyloid P-component (SAP) selectively binds to immobilized or bound forms of C-reactive protein (CRP). Biochim. Biophys. Acta. **1160**. 3. 309-316

Swarnakar, S., Asokan, R., Quigley, J. P., Armstrong, P. B. (2000). Binding of α2-macroglobulin and limulin: regulation of the plasma haemolytic system of the American horseshoe crab, *Limulus*. Biochem. J. **347**. 679-685.

Syin, C., Gotschlich, E. C., Liu, T. Y. (1986). Rabbit C-reactive protein. Biosynthesis and characterization of cDNA clones. The Journal of Biological Chemistry. **261**. 12. 5473-5479

Szalai, A. J., Norcum, M. T., Bly, J. E., Clem, L. W. (1992). Isolation of an acute-phase phosphorylcholine-reactive pentraxin from Channel catfish (*Ictalarus punctatus*). Comp. Biochem. Physiol. **102B**. 3. 535-543.

Szalai, A. J., Briles, D. A., Volanakis, J. E. (1995). Human C-reactive protein is protective against fatal *Streptococcus pneumoniae* infection in transgenic mice. The Journal of Immunology. **155**. 2557-2563.

Szalai, A. J., Briles, D. E., Volanakis, J. E. (1996). Role of complement in C-reactive-protein-mediated protection of mice from *Streptococcus pneumoniae*. Infection and Immunity. **64**. 11. 4850-4853

Szalai, A. J., Agrawal, A., Greenhough, T. J., Volanakis, J. E. (1997). C-reactive protein. Structural biology, gene expression, and host defense. Immunologic. Research. **16**. 2. 127-136

Szalai, A. J., Agrawal, A., Greenhough, T. J., Volanakis (1999). C-Reactive protein: structural biology and host defense function. Clin Chem Lab Med. **37**. 3. 265-270

Szalai, A. J., VanCott, J. L., McGhee, J. R., Volanakis, J. E., Benjamin, W. H. (2000a). Human C-reactive protein is protective against fatal *Salmonella enterica* serovar typhimurium infection in transgenic mice. Infection and Immunity. **68**. 10. 5652-5656.

Szalai, A. J., van Ginkel, F. W., Wang, Y., McGhee, J. R., Volanakis, J. E. (2000b). Complement-dependent acute phase expression of C-reactive protein and serum amyloid P component. The Journal of Immunology. **165**. 1030-1035.

Szalai, A. J. (2002). The antimicrobial activity of C-reactive protein. Microbes and Infection. **4**. 201-205.

Szalai, A. J. (2004). C-reacive protein (CRP) and autoimmune disease: facts and conjectures. Clinical and Developemental Immunology. **11**. 3/4. 221-226

Takiguchi, M., Fujinaga, T., Naiki, M., Mizuno, S., Otomo, K. (1990). Isolation, characterization, and quantitative analysis of C-reactive protein from horses. Am. J. Vet. Res. **51**. 8. 1215-1220

Tan, N. S., Ng, M. L. P., Yau, Y. H., Chong, P. K. W., Ho, B., Ding, J. L. (2000). Definition of endotoxin binding sites in horseshoe cran Factor C recombinant sushi proteins and neutralization of endotoxin by sushi peptides. FASEB. J. **14**. 1801-1813.

Tanaka,T., Horio, T., Matuo, Y. (2002). Secretory production of recombinant human C-reactive protein in *Escherica coli*, capable of binding with phosphorylcholine, and its characterisation. Biochemical and Biophysical Research Communications. **295**. 163-166

Tebo, J. M., Mortensen, R. F. (1990). Characterization and isolation of a C-reactive protein receptor from the human monocytis cell line U-937. J. Immunol. **144**. 1. 231-238

Tennent, G. A., Butler, P. J. G., Hutton, T., Woolfitt, A. R., Harvey, D. J, Rademacher, T. W., Pepys, M. B. (1993). Molecular characterization of *Limulus polyphemus* C-reactive protein. I. subunit composition. Eur. J. Biochem. **214**. 91-97.

Tennent, G. A., Lovat, L. B., Pepys, M. B. (1995). Serum amyloid P component prevents proteolysis of the amyloid fibrils of Alzheimer disease and systemic amyloidosis. Proc. Natl. Acad. Sci. USA. **92**. 4299-4303.

Tennent, G. A., Hutchinson, W. L., Kahan, M. C., Hirschfield, G. M., Gallimore, J. R., Lewin, J., Sabin, C. A., Dhillon, A. P., Pepys, M. B. (2008). Transgenic human CRP is not pro-atherogenic, pro-atherothrombotic, or pro-inflammatory in apoE-/- mice. Atherosclerosis. **196**. 248-255

TeWinkel, L. E. (1950). Notes on ovulation, ova, and early development in the smooth dogfish, *Mustelus canis*. Biological Bulletin. **99**. 3. 474-486

Thambi Dorai, D., Bachhawata, B. K., Bishayee, S., Kannan, K., Rajagopal Rao, D. (1981). Further characterization of the sialic acid-binding lectin from the horseshoe crab *Carcinoscorpius rotundicauda*. Archives of Biochemistry and Biophysics. **209**. 1. 325-333.

Tharia, H. A., Shrive, A. K., Mills, J. D., Arme, C., Williams, G. T., Greenhough, T. J. (2002). Complete cDNA structure sequence of SAP-like pentraxin from *Limulus polyphemus*: Implications for pentraxin evolution. J. Mol. Biol. **316**. 583-597.

Thomas-Rudolph, D., Du Clos, T. W., Snapper, C. M., Mold, C. (2007). C-reactive protein enhances immunity to Streptococcus pneumoniae by targeting uptake to FcγR on dendritic cells. The Journal of Immunology. **178**. 7283-7291

Thompson, A. R., Enfield, D. L. (1978). Human plasma P component: isolation and characterization. Biochemistry. **17**. 20. 4304-4310

Thompson, D., Pepys, M. B., Wood, S. P. (1999). The physiological structure of human C-reactive protein and its complex with phosphocholine. Structure. **7**. 2. 169-177.

Thompson, D., Pepys, M. B., Tickle, I., Wood, S. (2002). The structures of crystalline complexes of human serum amyloid P component with its carbohydrate ligand, the cyclic pyruvate acetal of galactose. J. Mol. Biol. **320**. 1081-1086.

Thompson, J. D., Higgins, D. G., Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research. **22**. 22. 4673-4680.

Tillet, W. S., Francis, T. (1930). Serological reactions in pneumonia with a non-protein somatic fraction of pneumococcus. Journal of Experimental Medicine. **52**. 561-571.

Togashi, S., Lim, S. K., Kawano, H., Ito, S., Ishihara, T., Okada, Y., Nakano, S., Kinoshita, T., Horie, K., Episkopou, V., Gottesman, M. E., Costantini, F., Shimida, K., Maeda, S. (1997). Serum amyloid P component enhances induction of murine amyloidosis. Lab. Invest. **77**. 5. 525-531

Tomasz, A. (1967). Choline in the cell wall of a bacterium: novel type of polymer-linked choline in Pneumococcus. Science. **157**. 789. 694-697

Tron, K., Manalov, D. E., Röcker, C., Kächele, M., Torzewski, J., Nienhaus, G. U. (2008). C-reactive protein specifically binds to Fcγ receptor type I on a macrophage cell line. Eur J Immunol. **38**. 5. 1414-1422

Tseng, J., Mortensen, R. F. (1986). Binding specificity of mouse serum amyloid P-component for fibronectin. Immunol. Invest. **15**. 8. 749-761

Tseng, J. M., Diiorio, L., Mortensen, R. F. (1988). Monoclonal antibodies to human C-reactive protein (CRP) that recognize epitopes in functional regions. Hybridoma. **7**. 2. 185-191

Tseng, J., Mortensen, R. F. (1988). Binding of human C-reactive protein (CRP) to plasma fibronectin occurs via the phosphorylcholine-binding site. Mol. Immunol. **25**. 8. 679-686.

Tseng, J., Mortensen, R. F. (1989). The effect of human C-reactive protein on the cell-attachment activity of fibronectin and laminin. Exp. Cell. Res. **180**. 2. 303-313.

Tsuboi, I., Matsukawa, M., Sato, N. (1993a). Isolation and characterisation of a sialic acid-specific lectin from haemolymph of the southeast asian horseshoe crab *Tachypleus gigas*. Biosci. Biotech. Biochem. **57**. 8. 1237-1242.

Tsuboi, I., Matsukawa, M., Sato, N., Kimura, S. (1993b). Isolation and characterization of a sialic acid-specific binding lectin from the haemolymph of Asian horseshoe crab, *Tachypleus tridentatus*. Biochimica et Biophysica Acta. **1156**. 255-262

Tsuboi, I., Yanagi, K., Wada, K., Kimura, S., Ohkuma, S. (1993c). Isolation and characterisation of a novel sialic acid-specific lectin from haemolymph of *Limulus polyphemus*. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology. **104**. 1. 19-26.

Tsuboi, I., Yanagi, K., Matsukawa, M., Kubota, H., Yamakawa, T. (1996). Isolation of a novel lectin from the haemolymph of horseshoe crabs *Limulus polyphemus* and its hemagglutinating properties. Comp. Biochem. Physiol. **113B**. 1. 137-142.

Tucci, A., Goldberger, G., Whitehead, A. S., Kay, R. M., Woods, D. E., Colten, H. R. (1983). Biosynthesis and postsynthetic processing of human C-reactive protein. The Journal of Immunology. **131**. 5. 2416-2419

Uhlenbruck, G., Sölter, J., Janssen, E., Haupt, H. (1982). Anti-galactan and anti-haemocyanin specificity of CRP. Annals of the New York Academy of Sciences. **389**. 476-479

Urbányi, Z., Medzihradszky, D. (1992). Rapid method to isolate serum amyloid P component from human plasma. Characterisation of the isolated protein. J. Chromatogr. **578**. 1. 130-133

Urbányi, Z., Lakics, V., Erdő, S. L. (1994). Serum amyloid P component-induced cell death in primary cultures of rat cerebral cortex. Eur. J. Pharmacol. **270**. 4. 375-378

Urbányi, Z., Lászlo, L., Tomasi, T. B., Tóth, E., Mekes, E., Sass, M., Pázmány, T. (2003). Serum amyloid P component induces neuronal apoptosis and β-amyloid immunoreactivity. Brain Research. **988**. 69-77.

Urbányi, Z., Forrai, E., Sárvári, M., Likó, I., Illé, J., Pázmány, T. (2005). Glycosaminoglycans inhibit neurodegenerative effects of serum amyloid P component *in vitro*. Neurochemistry International. **46**. 471-477.

Vallespi, M. G., Glaria, L. A., Reyes, O., Garay, H. E., Ferrero, J., Arana, M. J. (2000) A *Limulus* antilipopolysaccharide factor-derived peptide exhibits a new immunological

activity with potential applicability in infectious diseases. Clinical and Diagnostic Laboratory Immunology. **7**. 4. 669-675

Vallespi, M. G., Alvarez-Obregon, J. C., Rodriguez-Alonso, I., Montero, T., Garay, H., Reyes, O., Arana, M. J. (2003). A *Limulus* anti-LPS factor-derived peptide modulates cytokine gene expression and promotes resolution of bacterial acute infection in mice. International Immunopharmacology. **3**. 247-256.

Vetter, M. L., Gewurz, H., Hansen, B., James, K., Baum, L. L. (1983). Effects of C-reactive protein on human lymphocyte responsiveness. J. Immunol. **130**. 5. 2121-2126

Volanakis, J. E., Kaplan, M. H. (1971). Specificity of C-reactive protein for choline phosphate residues of pneumococcal C-polysaccharide. Proc. Soc. Exp. Biol. Med. **136**. 2. 612-614

Volanakis, J. E., Kaplan, M. H. (1974). Interaction of C-reactive protein complexes with the complement system. II. The Journal of Immunology. **113**. 1. 9-17.

Volanakis, J. E., Wirtz, K. W. A. (1979). Interaction of C-reactive protein with artificial phosphatidylcholine bilayers. Nature. **281**. 155-157

Volanakis, J. E., Narkates, A. J. (1981). Interaction of C-reactive protein with artificial phosphatidylcholine bilayers and complement. The Journal of Immunology. **126**. 5. 1820-1825.

Volanakis, J. E., Kearney, J. F. (1981). Cross-reactivity between C-reactive protein and idiotypic determinants on a phosphocholine-binding murine myeloma protein. Journal of Experimental Medicine. **153**. 1604-1614.

Volanakis, J. E. (1982). Complement-induced solubilization of C-reactive protein-pneumococcal C-polysaccharide precipitates: evidence for covalent binding of complement proteins to C-reactive protein and to pneumococcal C-polysaccharide. The Journal of Immunology. **128**. 6. 2745-2750.

Volanakis, J. E. (2001). Human C-reactive protein: expression, structure and function. Molecular Immunology. **38**. 189-197.

Volbeda, A., Hol, W. G. (1988). Structure of arthropodan hemocyanin. Prog. Clin. Biol. Res. **274**. 291-307

Voleti, B., Agrawal, A. (2005). Statins and nitric oxide reduce C-reactive protein production while inflammatory conditions persist. Molecular Immunology. **43**. 7. 891-896

Wang, C-M., Nguyen, N. Y., Yonaha, K., Robey, F., Liu, T-Y. (1982). Primary structure of rabbit C-reactive protein. The Journal of Biological Chemistry. **257**. 22. 13610-13615.

Wang, H.W., Wu, Y., Chen, Y., Sui, S. F. (2002). Polymorphism of structural forms of C-reactive protein. Int. J. Mol. Med. **9**. 6. 665-671

Weber, K., Osborn, M. (1969). The reliability of molecular weight determinations by dodecyl sulphate-polyacrylamide gel electrophoresis. The Journal of Biological Chemistry. **244**. 16. 4406-4412.

Whitehead, A. S., Bruns, G. A. P., Markham, A. F., Colten, H. R., Woods, D. E. (1983). Isolation of human C-reactive protein complementary DNA and localization of the gene to chromosome 1. Science. **221**. 4605. 69-71

Whitehead, A. S., Zahedi, K., Rits, M., Mortensen, R. F., Lelias, J. M. (1990). Mouse C-reactive protein. Biochem. J. **266**. 283-290.

Wijngaarden, S., van de Winkel, J. G., Bijlsma, J. W., Lafeber, F. P., van Roon, J. A. (2008). Treatment of rheumatoid arthritis patients with anti-TNF-alpha monoclonal antibody is accompanied by down-regulation of the activating Fcγ receptor I on monocytes. Clin Exp Rheumatol. **26**. 1. 89-95

Wilson, A. M., Ryan, M. C., Boyle, A. J. (2006). The novel role of C-reactive protein in cardiovascular disease: risk marker or pathogen. International Journal of Cardiology. **106**. 291-297.

Winkelhake, J. L., Vodicnik, M. J., Taylor, J. L. (1983). Induction in rainbow trout of an acute phase (C-reactive) protein by chemicals of environmental concern. Comp. Biochem. Physiol. C. **74**. 1. 55-58

Woo, P., Korenberg, J. R., Whitehead, A. S. (1985). Characterization of genomic and complementary DNA sequence of human C-reactive protein, and comparison with the complementary DNA sequence of serum amyloid P component. The Journal of Biological Chemistry. **260**. 24. 13384-13388.

Wood, H. F., McCarty, M., Slater, R. J. (1955). The occurrence during acute phase infections of a protein not normally present in the blood. V. Journal of Experimental Medicine. **100**. 1. 71-79.

Wood, S. P., Oliva, G., O'Hara, B. P., White, H. E., Blundell, T. L., Perkins, S. J., Sardharwalla, I., Pepys, M. B. (1988). A pentameric form of human serum amyloid P component. **202**. 169-173.

Woollard, K. J., Phillips, D. C., Griffiths, H. R. (2002). Direct modulatory effect of C-reactive protein on primary human monocyte adhesion to human endothelial cells. Clin Exp Immunol. **130**. 256-262

Wu, Y., Ji, S-R., Wang, H-W., Sui, S-F. (2002). Study of the spontaneous dissociation of rabbit C-reactive protein. Biochemistry (Moscow). **67**. 12. 1377-1382

Xia, D., Samols, D. (1997). Transgenic mice expressing rabbit C-reactive protein are resistant to endotoxemia. Proc. Natl. Acad. Sci. USA. **94**. 2575-2580.

Xing, D., Hage, F. G., Chen, Y. F., McCrory, M. A., Feng, W., Skibinski, G. A., Majid-Hassan, E., Oparil, S., Szalai, A. J. (2008). Exaggerated neointima formation in human C-reactive protein transgenic mice is IgG Fc receptor type 1(FcγRI)-dependent. Am J Pathol. **172**. 1. 22-30

Yang, H., Nan, B., Yan, S., Li, M., Yao, Q., Chen, C. (2005). C-reactive protein decreases expression of VEGF receptors and neuropilins and inhibits VEGF165-induced cell proliferation in human endothelial cells. Biochem. Biophys. Res. Commun. **333**. 3. 1003-1010.

Yang, J., Wezeman, M., Zhang, X., Lin, P., Wang, M., Qian, J., Wan, B., Kwak, L. W., Yu, L., Yi, Q. (2007). Human C-reactive protein binds activating Fcγ receptors and protects myeloma tumor cells from apoptosis. Cancer Cell. **12**. 3. 252-265

Yap, S. H., Moshage, H. J., Hazenberg, B. P. C., Roelofs, H. M. J., Bijzet, J., Limburg, P. C., Aarden, L. A., van Rijswijk, M. H. (1991). Tumor necrosis factor (TNF) inhibits interleukin (IL)-1 and/or IL-6 stimulated synthesis of C-reactive protein (CRP) and serum amyloid A (SAA) in primary cultures of human hepatocytes. Biochimica et Biophysica Acta. **1091**. 405-408

Yaron, G., Brill, A., Dashevsky, O., Yosef-Levi, I. M., Grad, E., Danenberg, H. D., Varon, D. (2006). C-reactive protein promotes platelet adhesion to endothelial cells: a potential pathway in atherotrhombosis. British Journal of Haematology. **134**. 426-431

Yasojima, K., Schab, C., McGeer, E. G., McGeer, P. L. (2000). Human neurons generate C-reactive protein and amyloid P: upregulation in Alzheimer's disease. Brain Research. **887**. 80-89.

Ying, S. C., Gewurz, H., Kinoshita, C. M., Potempa, L. A., Siegel, J. N (1989). Identification and partial characterization of multiple native and neoantigenic epitopes of human C-reactive protein by using monoclonal antibodies. J. Immunol. **143**. 1. 221-228

Ying, S. C., De Beer, F. C., Shephard, E., Siegel, J. N., Harris, D., Gewurz, B. E., Fridkin, M., Gewurz, H. (1992a). Localization of sequence-determined neoepitopes and neutrophil digestion fragments of C-reactive protein utilizing monoclonal antibodies and synthetic peptides. Mol. Immunol. **29**. 5. 677-687

Ying, S-C., Marchalonis, J. J., Gewurz, A. T., Siegel, J. N., Jiang, H., Gewurz, B. E., Gewurz, H. (1992b). Reactivity of anti-human C-reactive protein (CRP) and serum amyloid P component (SAP) monoclonal antibodies with limulin and pentraxins of other species. Immunology. **76**. 324-330.

Ying, S. C., Gewurz, A. T., Jiang, H., Gewurz, H. (1993). Human serum amyloid P component oligomers bind and activate the classical complement pathway via residues 14-26 and 76-92 of the A chain collagen-like region of C1q. Journal of Immunology. **150**. 1. 169-176

Yother, J., Forman, C., Gray, B. M., Briles, D. E. (1982). Protection of mice from infection with *Streptococcus pneumoniae* by anti-phosphocholine antibody. Infection and Immunity. **36**. 1. 184-188

Yother, J., Volanakis, J. E., Briles, D. E. (1982). Human C-reactive protein is protective against fatal *Streptococcus pneumoniae* infection in mice. The Journal of Immunology. **128**. 5. 2374-2376.

Zahedi, K., Mortensen, R. F. (1986). Macrophage tumoricidal activity induced by human C-reactive protein. Cancer research. **46**. 5077-5083

Zahedi, K., Tebo, J. M., Siripont, J., Klimo, G. F., Mortensen, R. F. (1989). Binding of human C-reactive protein to mouse macrophages is mediated by distinct receptors. J. Immunol. **142**. 7. 2384-2392

Zeller, J. M., Landay, A. L., Lint, T. F., Gewurz, H. (1986a). Aggregated C-reactive protein binds to human polymorphonuclear leukocytes and potentiates Fc receptor-mediated chemiluminescence. J. Lab. Clin. Med. **108**. 567-576

Zeller, J. M., Landay, A. L., Lint, T. F., Gewurz, H. (1986b). Enhancement of human peripheral blood monocyte respiratory burst activity by aggregated C-reactive protein. Journal of Leukocyte Biology. **40**. 769-783.

Zeller, J. M., Kubak, B. M., Gewurz, H. (1989). Binding sites for C-reactive protein on human monocytes are distinct from IgG Fc receptors. Immunology. **67**. 51-55.

Zeller, J. M., Sullivan, B. L. (1993). Monoclonal antibody to the type II Fc receptor for human IgG blocks potentiation of monocyte and neutrophil IgG-induced respiratory burst activation by aggregated C-reactive protein. Cell Immunol. **149**. 1. 144-154

Zhang, D., Jiang, S-L., Rzewnicki, D., Samols, D., Kushner, I. (1995). The effect of interleukin-1 on C-reactive protein expression in Hep3B cells is exerted at the transcriptional level. Biochem. J. **310**. 143-148.

Zhang, D., Sun, M., Samols, D., Kushner, I. (1996). STAT3 participates in transcriptional activation of the C-reactive protein gene by interleukin-6. The Journal of Biological Chemistry. **271**. 16. 9503-9509.

Zhu, Y., Thangamani, S., Ho, B., Ding, J. L. (2005). The ancient origin of the complement system. The EMBO Journal. **24**. 382-394.

Zouki, C., Beauchamp, M., Baron, C., Filep, J. G. (1997). Prevention of *in vitro* neutrophil adhesion to endothelial cells through shedding of L-selectin by C-reactive protein and peptides derived from C-reactive protein. J. Clin. Invest. **100**. 3. 522-529.

Zouki, C., Haas, B., Chan, J. S. D., Potempa, L. A., Filep, J. G. (2001). Loss of pentameric symmetry of C-reactive protein is associated with promotion of neutrophil-endothelial cell adhesion. The Journal of Immunology. **167**. 5355-5361.

Web resources

Biological Magnetic Resonance Data Bank. A repository for data from NMR spectroscopy on proteins, peptides and nucleic acids. Department of Biochemistry, University of Wisonsin.
www.bmrb.wisc.edu
Date accessed: August 2008

The ConSurf Server. Server for the identification of functional regions of proteins.
http://consurf.tau.ac.il
Date accessed: 2007 – 2008

European Bioinformatics Institute (EBI). European Molecular Biological Laboratory (EMBL).
www.ebi.ac.uk
Date accessed: 2004-2008

EMBL-EBI. MUSCLE: MUltiple Sequence Comparison by Log-Expectation
www.ebi.ac.uk/muscle
Date accessed: 2007- February 2008

ExPASy. ProtParam tool.
www.expasy.ch/tools/protparam.html
Date accessed: 2004-2008

ExPASy Translate Tool.
(www.expasy.ch/tools/dna.html).
Date accessed: 2007 – August 2008.

Invitrogen Corporation.
www.invitrogen.com
Date accessed: 2007

Marine Biological Laboratory. Biological Discoverey in Woods Hole.
(www.mbl.edu/marine_org/marine_org).
Date accessed: March 2008.

Maryland. Department of Natural Resources.
www.dnr.state.md.us.com
Date accessed: 2008

Mobyle. Boxshade: printouts from multiple-aligned protein or DNA sequences.
(http://mobile.pasteur.fr/cgi-bin/MobylPortal/portal.py?form=boxshade).
Date accessed: 2007 – August 2008

Mobyle. Clustalw: multiple alignment.
(http://mobile.pasteur.fr/cgl-bin/MobylPortal/portal.py?form=clutalw-mulialign).
Date accessed: 2007 – August 2008.

National Center for Biotechnology Information.
www.ncbi.nlm.nih.gov
Date accessed: 2004-2008

PHYLIP. By Joe Felsenstein, Depeartment of Genome Sciences and the Department of
Biology at the University of Washington.
http://evolution.genetics.washington.edu/phylip.html
Date accessed May 2008

RCSB Protein Data Bank. An information portal to biological macromolecular structures.
www.rcsb.org/pdb/home/home.do
Date accessed: 2004 – 2008