Andreia Margarida Silva Barro

Degree in Cell and Molecular Biology

Characterization of APRc as an immune evasion factor of *Rickettsia*

Dissertation to obtain the Master degree in Biotechnology

Supervisor: Dr. Isaura Isabel Gonçalves Simões, Auxiliary Investigator,
Centre for Neuroscience and Cell Biology, University of Coimbra

Examination Committee:

President: Dr. Susana Filipe Barreiros, Full Professor, NOVA School of Science and Technology

Examiner: Dr. Luís Jaime Gomes Ferreira da Silva Mota, Assistant Professor, NOVA School of Science and Technology

Member: Dr. Isaura Isabel Gonçalves Simões, Auxiliary Investigator, Centre for Neuroscience and Cell Biology, University of Coimbra

December 2020

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Andreia Margarida Silva Barro

Degree in Cell and Molecular Biology

Characterization of APRc as an immune evasion factor of *Rickettsia*

Dissertation to obtain the Master degree in Biotechnology

Supervisor: Dr. Isaura Isabel Gonçalves Simões, Auxiliary Investigator,
Centre for Neuroscience and Cell Biology, University of Coimbra

Examination Committee:

President: Dr. Susana Filipe Barreiros, Full Professor, NOVA School of Science and Technology

Examiner: Dr. Luís Jaime Gomes Ferreira da Silva Mota, Assistant Professor, NOVA School of Science and Technology

Member: Dr. Isaura Isabel Gonçalves Simões, Auxiliary Investigator, Centre for Neuroscience and Cell Biology, University of Coimbra

December 2020

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Characterization of APRc as an immune evasion factor of *Rickettsia*

Acknowledgements

First and foremost, I would like to express my profound appreciation towards my supervisor, Dr. Isaura Simões, for her dedication, support, patience, advice and guidance, fundamental for the realization of this dissertation. And, also for her contagious motivation and enthusiasm for science.

I also would like to extend my gratitude to the other members of the Microbial and Molecular Biotechnology Group. To assistant investigator Pedro Curto, for all the help, advice and guidance provided. And to master's student Bárbara Teixeira, for her support and motivation.

A special thanks to Ana Catarina, that besides being a great friend, has guide me throughout my academic life and always tried to encourage me and motivate me to achieve my goals.

I also give a special thanks to my beloved friends Inês, Marta, Rafael and Diogo, for all the fun moments that we experience, and for the support, advice and encouragement that made me overcome many obstacles. And to the best friends that university could give me, my dearest biologists, Rita, Joana, Telma, Beatriz, Mafalda, Sofia D. and Sofia M.. Their friendship, support and encouragement were essential.

And lastly, I would like to express my immense appreciation towards my family, that throughout my life always supported me and encouraged me to pursuit my dreams.

Abstract

Several pathogens, such as *Rickettsia*, have developed different tactics to evade immune responses in order to subsist and proliferate within the host. Rickettsial species are resistant to the serum bactericidal effects and can evade complement-mediated killing, suggesting that they developed mechanisms to inhibit recognition by host serum components. Interestingly, many pathogenic bacteria have also been shown to use non-immune immunoglobulin (Ig)-binding proteins to avoid recognition by innate immune serum components. With this work, we provide evidence for a novel Ig-binding protein from *Rickettsia*, anticipating additional immune-evasion tactics for this obligate pathogen. We demonstrate that the rickettsial retropepsin APRc is capable of binding to immunoglobulins from different species and classes. Also, our results suggest that the interaction of APRc with human IgG promotes the oligomeric stabilization of APRc. Mapping the APRc region responsible for binding revealed the segment between amino acids 150-166 as one of the interacting regions. Moreover, we demonstrate that the interaction occurs in the Fab domain but does not entail IgG cleavage. We also demonstrate IgG-APRc binding in serum samples. Finally, we demonstrate IgG-binding at the surface of *Rickettsia*. APRc-IgG binding and its localization at *Rickettsia's* outer membrane anticipate APRc as one of the proteins contributing to that activity. Altogether, we propose that APRc may act as a novel evasin by playing a role in protecting *Rickettsia* from complement-mediated killing through this non-immune IgG-binding activity.

Since *Rickettsia* are arthropod-borne pathogens that can originate severe diseases in humans, there is a growing concern about the increasing incidence of rickettsioses and their impact on global health. Therefore, with the lack of reliable protective vaccines for rickettsial diseases, the development of alternative therapeutics is critical. With this work, we provide valuable information for a deeper understanding of the molecular mechanisms underlying rickettsial pathogenesis, contributing to the future development of alternative therapeutics.

Keywords: *Rickettsia;* rickettsial retropepsin; APRc; non-immune Ig-binding proteins; immune evasion; rickettsial immune evasion toolbox.

Resumo

Para subsistir e proliferar no hospedeiro, vários patógenos (*e.g. Rickettsia*) desenvolveram diferentes táticas para escapar às respostas imunológicas do hospedeiro. Rickettsiae mostram resistência aos efeitos bactericidas do soro e aptidão para escapar à morte mediada pelo complemento, sugerindo a existência de mecanismos para inibir o reconhecimento pelos componentes do soro. Muitas bactérias patogénicas também usam proteínas capazes de ligação não-imune a imunoglobulinas para evitar esse reconhecimento. Neste trabalho, é identificada uma proteína de *Rickettsia* com capacidade de ligação não-imune a imunoglobulinas, antecipando-se novas táticas de evasão imunológica para este patógeno obrigatório. Demonstra-se que a retropepsina de *Rickettsia*, APRc, é capaz de ligar a imunoglobulinas de diferentes espécies e classes. Os nossos resultados também sugerem que a interação entre APRc e IgG humano promove a estabilização oligomérica da APRc. Identificámos pelo menos uma região na APRc importante para esta interação e demonstramos que a mesma ocorre na região Fab do anticorpo. No entanto, não se observa clivagem do anticorpo. Também é demonstrada a ligação de IgG-APRc em amostras de soro humano. Finalmente, demonstramos a ligação de IgG na superfície de *Rickettsia*, antecipando-se a APRc como uma das proteínas que contribui para essa atividade. Suportados nestas evidências, propomos uma nova função para a retropepsina APRc como uma nova evasina, contribuindo para a proteção de *Rickettsia* contra a morte mediada pelo complemento através desta ligação não-imune a IgG.

As espécies do género *Rickettsia* são transmitidas por artrópodes, podendo originar doenças infeciosas graves em humanos. As rickettsioses apresentam um aumento de incidência em todo o mundo, existindo uma preocupação crescente sobre o impacto destas doenças na saúde global. Com a falta de vacinas fiáveis, o desenvolvimento de tratamentos alternativos é extremamente importante. Este trabalho contribui para o conhecimento mais detalhado dos mecanismos de patogenicidade de *Rickettsia*, contribuindo para o futuro desenvolvimento de terapias alternativas.

Palavras-chave: *Rickettsia*; retropepsina; APRc; proteínas de ligação não imune a imunoglobulinas; mecanismos de evasão da resposta imune;

Table of Contents

List of Figures

List of Tables

List of Abbreviations

A_{220nm}	Absorbance at 220 nm
A_{280nm}	Absorbance at 280 nm
A_{450nm}	Absorbance at 450 nm
AG	Ancestral group
AIDS	Acquired immunodeficiency syndrome
APRc	Aspartic protease from *Rickettsia conorii*
BSA	Bovine serum albumin
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EIAV	Equine infectious anemia virus
ELISA	Enzyme-linked immunosorbent Assay
Fab	Fragment antigen-binding
Fc	Fragment crystallizable
FIV	Feline immunodeficiency virus
HIV-1	Human immunodeficiency virus type 1
HIV-2	Human immunodeficiency virus type 2
HPLC	High-performance liquid chromatography
HRP	Horseradish peroxidase
HSA	Human serum albumin
Ig	Immunoglobulin
IHC	Immunohistochemistry
IMAC-Ni^{2+}	Immobilized metal affinity chromatography with Ni^{2+}
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	Kilodalton
LB	Luria Bertani broth
LMW	Low molecular weight
MIC	Minimum inhibitory concentration
MSF	Mediterranean spotted fever
MW	Molecular Weight
NHS	Normal human serum
NHSΔHSA	Human normal umbilical cord serum depleted from human serum albumin

OD_{600nm}	Optical density at 600nm
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline containing Tween 20
PCR	Polymerase chain reaction
PFA	Peroxyformic acid
PFU	Plaque forming units
PLA2	Phospholipase A2
PNK	Polynucleotide Kinase
PpL	Protein L from *Finegoldia magna*
RMSF	Rocky Mountain spotted fever
RSV	Rous sarcoma virus
S:E	Substrate:Enzyme
Sca	Surface cell antigens
ScFv	Single Chain Variable Fragment
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SFG	Spotted Fever Group
SIV	Simian immunodeficiency virus
SpA	Staphylococcal protein A
SpG	Streptococcal protein G
TBS	Tris buffered saline
TBS-T	Tris buffered saline containing Tween 20
TFA	Trifluoroacetic acid
TG	Typhus group
TRG	Transitional group

1 Introduction

1.1 *Rickettsia*

1.1.1. Bacteriology

Rickettsiae were first described by Ricketts in 1909[1]. These Gram-negative obligate intracellular bacteria are small (0.8-2.0 by 0.3-0.5 μm) coccobacilli that require eukaryotic cells in which to replicate. *Rickettsia* multiply in the cytosol or within the nucleus of the host cells and divide by binary fission[2]. Although they stain poorly with conventional Gram techniques, when stained by the Gimenez method, they retain basic fuschin[3].

1.1.2. Taxonomy and Phylogeny

Currently, within the class Alphaproteobacteria, which comprises a diverse group of Gram-negative bacteria, lies the order *Rickettsiales*[4,5]. This group of obligate or facultative intracellular organisms accommodates the family *Anaplasmataceae*, containing the genera *Anaplasma*, *Ehrlichia*, *Wolbachia*, and *Neorickettsia*; and the family *Rickettsiaceae*, containing the genus *Orientia* and the genus *Rickettsia*[4].

Until now, more than 47 bacterial species have been described within the genus *Rickettsia* (www.ncbi.nlm.nih.gov/taxonomy). They are classified into four groups: the spotted fever group (SFG), the typhus group (TG), the ancestral group (AG), and the transitional group (TRG) (Figure 1.1)[4]. The SFG, the largest cluster, is represented by tick-borne species, where several of them are of unknown pathogenicity, but some are pathogenic agents that cause spotted fevers (*e.g.*, *Rickettsia rickettsii* and *Rickettsia conorii*). Within the TG, we have highly pathogenic species that cause murine typhus and epidemic typhus and are transmitted by rat fleas (*Rickettsia typhi*) and human body lice (*Rickettsia prowazekii*), respectively. AG *Rickettsia* includes non-pathogenic and tick-borne species (*Rickettsia bellii* and *Rickettsia canadensis*). The TRG includes *Rickettsia akari* (mite-borne), *Rickettsia felis* (flea-borne), *Rickettsia asembonensis*, *Rickettsia australis* ,and *Rickettsia hoogstraalii*[4,5].

Figure 1.1- Phylogenetic classification of the bacterial species belonging to the genus *Rickettsia*. This classification is substantiated with a phylogenomic approach. The genus *Rickettsia* is classified into four main clusters: the spotted fever group (SFG), the typhus group (TG), the ancestral group (AG), and the transitional group (TRG). Adapted from Diop, A. *et al*. (2020)[4].

Despite the recent molecular biology approaches and the whole genome sequence analysis used for the species delineation, the strict intracellular lifestyle of the species within the genus *Rickettsia* and their small genomes make a consensual classification difficult. Thereby, many rearrangements in the phylogenetic classification of *Rickettsia* have been proposed, making it a matter of debate until today[4,6].

1.1.3. Rickettsioses

Epidemiology

The pathologies caused by obligate intracellular bacteria from the family *Rickettsiaceae*, associated with arthropod vectors (*e.g.*, ticks, lice, mites, and fleas), are designated rickettsioses and they constitute a big percentage of the emerging infectious diseases, representing a significant burden on global health and economics[7,8]. These pathogens are endemic worldwide, except for Antarctica, and their geographic distribution is dependent on climatic conditions and vector and natural host limitations[7]. Rickettsioses are transmissible from animals to humans and can originate mild to severe illnesses (Table 1.1)[9].

SFG *Rickettsia* are tick-transmitted and worldwide distributed rickettsioses. Two of the most severe rickettsioses are Rocky Mountain spotted fever (RMSF) and Mediterranean spotted fever (MSF). RMSF is caused by *R. rickettsii* (considered the most pathogenic rickettsial species among SFG *Rickettsia*) and has been reported in the Americas. *Dermacentor variabilis* (in the southeast and south-central states of the United States), *Dermacentor andersoni* (in the western mountainous states of the United States), *Rhipicephalus sanguineus* (in the tribal lands of Arizona and in Mexico) and a variety of *Amblyomma* species (in Central and South America) constitute tick vectors of *R. rickettsii*. MSF is caused by SFG *R. conorii* that is transmitted by *R. sanguineus* in Europe, Africa, and Asia[9].

TG rickettsioses include epidemic typhus and murine typhus. Epidemic typhus is caused by a lice-born TG *Rickettsia* (*R. prowazekii*); it is transmitted in unhygienic conditions by body louse (*Pediculus humanus humanus*) and constitutes one of the most severe rickettsioses (when conditions favor transmission of the body lice, often epidemics occur). Murine typhus is caused by the flea-borne TG *R. typhi* transmitted by the rat flea (*Xenopsylla cheopis*), especially in tropical and sub-tropical seaboard regions, or by the cat flea (*Ctenocephalides felis*) in southern California and south Texas[9].

TRG rickettsioses comprise Queensland tick typhus, Rickettsialpox, and Flea-borne spotted fever. Queensland tick typhus is caused by *R. australis* and transmitted by *Ixodes holocyclus* ticks along the eastern coast of Australia. Rickettsialpox is caused by *R. akari*, described in the United States, Mexico, Ukraine, Croatia, and Turkey, and is transmitted by the mouse mites *Liponyssoides sanguineus*. The Flea-borne spotted fever is caused by *R. felis*, having a worldwide distribution, and it is mostly transmitted by cat fleas (*C. felis*)[9].

Table 1.1- Rickettsioses and their epidemiologic features (associated organism and respective species group, geographic distribution, associated arthropod vector, and severity of the disease). Adapted from Blanton, L. S. (2019)[9].

Disease	Organism	Group	Distribution	Vector	Severity
Rocky Mountain spotted fever	*R rickettsii*	Spotted fever	Americas	Tick	+++++
Mediterranean spotted fever	*Rickettsia conorii*	Spotted fever	Europe, Africa, Asia	Tick	+++
Siberian tick typhus	*Rickettsia sibirica*	Spotted fever	Eurasia, Africa	Tick	++
Japanese spotted fever	*Rickettsia japonica*	Spotted fever	Japan, eastern Asia	Tick	++
Flinders Island spotted fever	*Rickettsia honei*	Spotted fever	Australia, Asia	Tick	++
Far Eastern spotted fever	*Rickettsia heilongjiangensis*	Spotted fever	Eastern Asia	Tick	++
African tick bite fever	*Rickettsia africae*	Spotted fever	Sub-Saharan Africa, Caribbean islands	Tick	++
Maculatum disease	*Rickettsia parkeri*	Spotted fever	Americas	Tick	++
Tick-borne lymphadenopathy	*Rickettsia slovaca*	Spotted fever	Europe, Asia	Tick	+
Tick-borne lymphadenopathy	*Rickettsia raoultii*	Spotted fever	Europe, Asia	Tick	+
Unnamed	*Rickettsia massiliae*	Spotted fever	South America, Europe	Tick	+
Pacific Coast tick fever	*Candidatus* R. philippii	Spotted fever	United States	Tick	+
Unnamed	*Rickettsia aeschlimannii*	Spotted fever	Europe, Africa	Tick	+
Unnamed	*Rickettsia monacensis*	Spotted fever	Europe	Tick	+
Unnamed	*Rickettsia helvetica*	Spotted fever	Europe	Tick	+
Asymptomatic or mild illness with seroconversion	*Rickettsia amblyommatis*	Spotted fever	Americas	Tick	±
Typhus	*R prowazekii*	Typhus	South America, Africa, Eurasia	Body louse, ectoparasites of flying squirrels	++++
Murine typhus	*R typhi*	Typhus	Worldwide	Flea	+++
Rickettsialpox	*R akari*	Transitional	North America, Eurasia	Mouse mite	++
Queensland tick typhus	*R australis*	Transitional	Eastern Australia	Tick	++
Flea-borne spotted fever	*R felis*	Transitional	Worldwide	Flea	+

Disease symptoms and therapeutics

Rickettsioses present a multitude of clinical signs and symptoms, being generally characterized by high fevers. Headache, myalgia, nausea, vomiting, and abdominal pain constitute other typical symptoms. Macular or maculopapular rashes have been described in patients with SFG, TG, and TRG rickettsioses, being present in most patients with MSF, RSMF, epidemic typhus, and Queensland tick typhus. Local dermal and epidermal necrosis, forming an eschar at the site of tick or mite inoculation, have been described in patients with SFG and TRG rickettsioses, such as African tick bite fever, MSF, RMSF (rarely), rickettsialpox, and Queensland tick typhus. Moreover, lymphadenopathy is observed in patients with less severe SFG rickettsioses, being sometimes noted in RMSF patients and patients with TRG rickettsioses[9,10].

Untreated rickettsioses, caused by highly pathogenic rickettsiae, can be lethal and develop severe injuries, sometimes evolving to multiorgan failure. Systemic vascular infection in RMSF can progress to encephalitis, leading to stupor, coma and seizures, interstitial pneumonia, non-cardiogenic pulmonary edema, and adult respiratory distress syndrome. In the most severe cases of this type of rickettsioses, hypovolaemia and hypotensive shock can progress to acute renal failure. In another highly severe rickettsiosis, epidemic typhus, the disease can also progress to life-threatening conditions if not treated

early with a suitable antibiotic. The aggravation of the latent infection caused by *R. prowazekii* (responsible for epidemic typhus) in convalescent individuals can lead to Brill–Zinsser disease, characterized by fever, rash, and less-severe illness[9,10].

The diagnosis of rickettsial diseases is accomplished by understanding of the complete clinical history (exposure to a potential source of rickettsial disease is evaluated) followed by laboratory tests, such as serology tests, molecular detection, isolation and culture of pathogens, and immunohistochemistry (IHC) analysis. The treatment of rickettsial diseases relies on the administration of antibiotics. The antibiotic class of choice are tetracyclines, with a minimum inhibitory concentration (MIC) 0.06 to 0.25 μg/mL. Doxycycline is the preferred antibiotic and minocycline also shows to be effective. Furthermore, chloramphenicol and fluoroquinolones were shown to be good alternatives for the treatment of rickettsioses[7,9].

Concerns associated with Rickettsioses

There is growing concern about rickettsioses and their severe impact on global health. Some of these rickettsial diseases are very severe, being considered life-threatening diseases. The rickettsial organisms responsible for these pathologies continue to emerge and re-emerge worldwide, and an increasing incidence of these diseases has been noticed since 1970[9,11]. Climate change and social changes (*e.g.*, international trade, travel, human behavior, demographics, and availability of public health care infrastructures) have been described as possible agents responsible for the emergence of these diseases[12].

Another factor of great concern is the currently limited spectrum of drugs for the treatment of rickettsial diseases. The treatment of rickettsioses is based on a few antibiotics and, until this day, there are no protective vaccines. Also, indications of antibiotic resistance in some rickettsial species were observed, and it was detected that some rickettsiae could persist after the treatment with antibiotics. Therefore, the identification of new potential therapeutic targets and vaccines that can prevent rickettsioses are extremely needed. For that, more information about the molecular mechanisms underlying rickettsial pathogenesis and the protective adaptive immune responses against rickettsiae is required[13].

1.1.4. Pathogenesis

Rickettsia pathogenesis comprises several steps. The sequence of events begins with the acquisition, maintenance and transmission of rickettsiae by the arthropod vectors, followed by the entry in the vertebrate host, adhesion to the host cell, evasion of host defences, slow growth with acquisition of nutrients from the host cell and spread throughout the other host cells (induction of cell injury)[10,14].

Life cycle

In the rickettsial life cycle, rickettsial organisms are acquired, maintained, and transmitted to humans, non-human mammals, and birds by hematophagous arthropods (*e.g.*, fleas, lice, ticks, or mites). Without the arthropod vector, *Rickettsia* would have little impact on its interactions with vertebrate hosts[15].

The arthropod vectors acquire and maintain rickettsial species through vertical transmission (by transovarial transmission, as from the adult female to the egg, or by trans-stadial passage, as from egg to larva to nymph to adult) and/or through horizontal transmission during the feeding of the infected arthropod vector[10,15].

When the arthropod vectors are in their feeding behavior, they bite and may infect humans. When feeding on the blood generated after the bite, they salivate into the wound and release immunomodulatory and anti-inflammatory chemicals, increasing the blood flow and hindering the host's mechanisms from eliminating the parasite. *Rickettsia* are secreted with the saliva into the bite site and take advantage of this area of low immune activity to establish infection[15]. The infection of human hosts can also occur by the feces of infected louse or flea released in the skin during the feeding. By scratching this area, the rickettsial organisms can enter the skin throughout the bite site[10,15]. Therefore, humans are considered accidental hosts for *Rickettsia*[15].

Invasion and dissemination within host cells

As an obligate intracellular bacteria, *Rickettsia* need to adhere to and invade target host cells to survive, multiply, and successfully establish the infection. Vascular endothelial cells are considered the main target cells of rickettsial organisms; however, it has been reported that rickettsial species are able

to interact with cells other than the endothelium[16,17]. As a way of example, pathogenic SFG *Rickettsia* are able to survive and proliferate within THP-1 macrophage-like cells and epithelial cells[17].

Viability of the invading bacteria and metabolic activity of the host cell are required for the intracellular uptake of rickettsiae. After transmission, interactions between rickettsial ligands on the outer membrane and their specific host cell surface receptors, and consequent activation of downstream signaling mechanisms, lead to the internalization of the invading bacteria by a process designated as "induced phagocytosis"[16].

Several proteins involved in rickettsial-host cell interactions have been identified. Among them, there are the surface cell antigens (Sca), encoding proteins similar to autotransporter proteins from Gram-negative bacteria. Sca0 (OmpA) (a surface-exposed protein conserved in SFG and missing in TG rickettsiae), Sca1 (an adhesin responsible for the attachment of *R. conorii* to mammalian endothelial cells and with no role in invasion), Sca2 (a protein that participates in mediating interactions with mammalian target cells and is conserved in SFG and missing in TG rickettsiae) and Sca5 (OmpB) (a surface-exposed protein conserved in SFG rickettsiae), have important roles in rickettsial adhesion and/or invasion. rOmpB binds to its host receptor, Ku70 (a subunit of nuclear DNA-dependent protein kinase) activating a host-signaling cascade. This pathway involves the ubiquitination of Ku70, mediated by the ubiquitin ligase c-cbl, the GTPase Cdc42, the phosphoinositide 3-kinase and the activation of tyrosine kinases such as c-Src, leading to the activation of the Arp2/3 complex responsible for the host cell actin polymerization, and subsequently rickettsial internalization. The internalization process is dependent on clathrin and caveolin-2-dependent endocytosis[16].

The internalization of rickettsiae is followed by the escape from the phagosome into the host cytosol. The degradation of the phagosomal membrane is likely induced by hemolysin C, phospholipase D, and phospholipase A2 (PLA2). Once *Rickettsia* reach the cytosol, they utilize the host cell actin cytoskeleton, facilitating their movements and spreading from cell-to-cell. The rickettsial protein RickA (a nucleation-promoting factor that activates the Arp2/3 complex) and the autotransporter protein Sca2 have important roles in the actin-based mobility[16]. This is extremely important to establishing and disseminating the rickettsial infection within the endothelium, where they grow and stimulate oxidative stress, causing cell injury[10,16].

Rickettsial evasion of the bacterial effects of complement

For the survival and proliferation of rickettsial organisms, they require adhesion and invasion to target host cells[16]. Nevertheless, *Rickettsia* spend an extension of time outside the host cell, for example, during the feeding of the arthropod vector or during dissemination. In these periods, *Rickettsia* do not have the relative protection of the host cell[18,19], being exposed to the bactericidal effects of the host's complement system. The rickettsial survival is dependent on their ability to evade killing until being able to invade a proper cellular host[19,20]. It was demonstrated that *Rickettsia* are resistant to the serum bactericidal effects, and can evade complement-mediated killing[21]. Thus far, it has been demonstrated that (at least) three rickettsial surface proteins mediate this mechanism. The rickettsial autotransporter protein (rOmpB), which specifically interacts with factor H (a soluble host complement inhibitor)[20] and the rickettsial outer membrane proteins, Adr1 and Adr2, that interact with the terminal complement complex inhibitor vitronectin (Vn)[18,19,22]. Since all these protein factors seem to mediate partial survival of *Rickettsia* in human serum, it is reasonable to speculate that rickettsial species may have evolved multiple mechanisms to evade complement-mediated killing which are still to be uncovered.

1.2 Bacterial Immunoglobulin-Binding Proteins

To subsist and proliferate within the host, pathogens typically need to go through the protective barriers of the host (*e.g.*, epithelia). Some accomplish it by sliding between cells, and others succeed by promoting their uptake into the host epithelial cells and establishing infection[23]. However, the internalization of the pathogens is not always that advantageous. The host organism holds strategies to obliterate the invading microorganisms, such as the overlaying of the pathogen by professional phagocytes (macrophages, neutrophils, and dendritic cells)[23]. During the process of phagocytosis, the internalization of pathogens into the host cells can occur by interaction of the host cell receptors with the cognate ligands on the surface of pathogens or by injection of effectors, promoting entry. Then, the pathogen remains in the phagosome, a membrane-bound vacuole that undergoes maturation, promoting the degradation and killing of the invading microorganism (Figure 1.2)[23].

Thereby, some pathogens have established different tactics in order to evade phagocytosis and establish infection. With the intention to either promote or evade the uptake into the host cells, those tactics include mechanisms of impairing the maturation of the phagosome; mechanisms of escape from the phagosome; mechanisms of survival within a mature phagosome, and mechanisms to evade the internalization. This last set of mechanisms comprises the encapsulation and the complement/antibody inactivation of the pathogen, which will prevent the recognition of the pathogen by the phagocytes.

Moreover, it also includes mechanisms that require attachment to the host cell without the following internalization. These are the manipulation of the cytoskeleton, resulting in adherence/effacement, and the interference with signaling (Figure 1.2)[23].

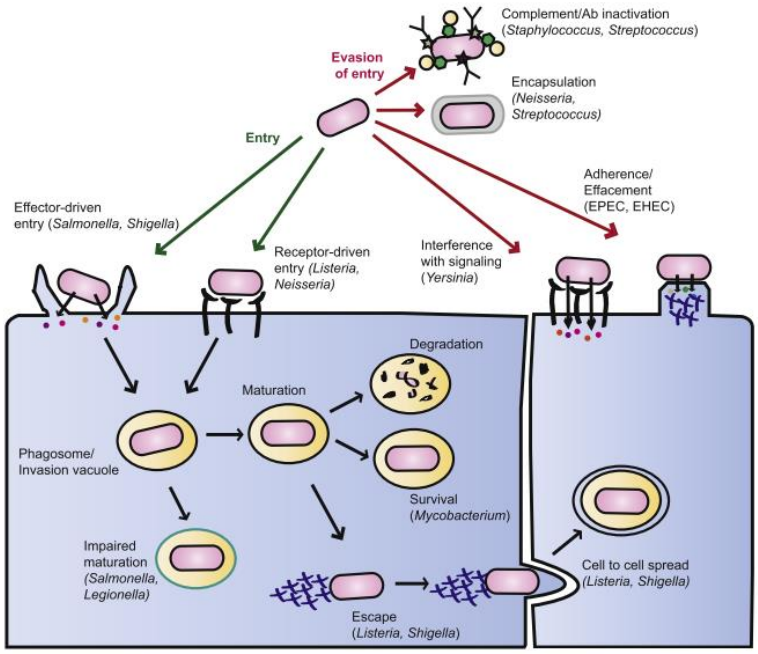


Figure 1.2- Subversion of phagocytosis for pathogen survival. Different strategies used by pathogens to evade the phagocytosis process, which is employed by host cells as a defence mechanism against infection. Adapted from Sarantis, H. *et al*. (2012)[23].

The recognition of pathogens by host cells can be promoted by the attachment of serum components (opsonins) to the bacterial surface. The host cells have opsonic receptors (*e.g.*, complement components' receptors and receptors for the fragment crystallizable region (Fc region) of immunoglobulins) that by recognizing the opsonization of the invading microorganisms, tag them for neutralization[23].

Among the several strategies described above to evade phagocytosis, pathogens also have specific strategies to escape the opsonization, and therefore evade the recognition by the host cells[23]. Certain bacteria have developed the ability to produce immunoglobulin (Ig)-binding proteins, that without the implication of the antigen-binding sites are able to bind to immunoglobulins (non-immune binding). These proteins have the purpose of scavenging the opsonizing antibodies, impairing the recognition by the opsonic receptors of the host cells (complement/antibody inactivation of the pathogen)[23,24]. Therefore, since they play an important role in protecting bacteria against the action of the protective mechanisms of the host's body, they constitute important factors of pathogenicity of bacteria[24].

Ig-binding proteins of bacteria can be located at the surface of cells or appear in the culture medium and their molecular weights are in the range of 20 to 350 kDa[24]. This group of proteins integrates several proteins that highly diverge in molecular structure, location within the bacteria, and binding features. They can bind to different classes and subclasses of immunoglobulins and have different binding sites[24]. Furthermore, a single bacterial cell can simultaneously have many Ig-binding proteins with different structures and functions[24].

The biological function of these proteins can also diverge according to the type of bacterial cell envelope. In Gram-positive bacteria, the proteins that interact with the environment are covalently attached to the peptidoglycan and present the LPXTG motif responsible for inducing cell-wall anchoring of proteins. Whereas in Gram-negative bacteria, those proteins are anchored in the outer lipid membrane. Thus, the binding properties of the Ig-binding proteins differ between the Gram-positive and Gram-negative bacteria. The study of Ig-Binding proteins from Gram-positive bacteria is more extensive, compared to Gram-negative bacteria, which have a lower number of known Ig-Binding proteins and limited information about them[24,25]. Ig-binding proteins from different bacterial species and their main properties and biological activities, further described in this chapter, are exemplified in Table 1.2.

Table 1.2- Immunoglobulin-binding proteins from different bacterial species and their associated properties and biological activities. Adapted from Collin, M. *et al.* (2014)[25].

Protein	Species	Mode of Action	Biological Effects/ Additional Features
Protein A	*S. aureus*	IgG Fc binding	Inhibition of opsonophagocytic clearance, superantigen, B-cell depletion
Protein G	Group G *Streptococcus*	IgG Fc binding	Also binds IgG Fab
Protein L	*F. magna*	κ light-chain binding	Superantigen, mast cell activation
Protein H	*S. pyogenes*	IgG Fc binding	Inhibition of opsonophagocytic killing, complement inhibition
M proteins	*S. pyogenes*	IgG Fc binding	Inhibition of opsonophagocytic killing
Protein ARP	*S. pyogenes*	IgA binding	Also binds the complement regulator C4BP and thereby inhibits complement at the bacterial surface
Protein Sir	*S. pyogenes*	IgG, IgA binding	Also binds the complement regulator C4BP and thereby inhibits complement at the bacterial surface
β protein	*S. agalactiae*	IgA binding	Protection against IgA-mediated bacterial killing
MID	*M. catharralis*	IgD binding	Activates B cells
Protein D	*H. influenzae*	IgD binding	Histamine release, degranulation of mast cells
Sbi	*S. aureus*	IgG binding	Inhibition of opsonization, complement inhibition

1.2.1 IgG-Binding Proteins of Gram-Positive Bacteria

Six different patterns of non-immune binding to the Fc region of different subclasses of IgG from different mammalian species have been described in Gram-positive bacteria. Therefore, IgG Fc-binding proteins of Gram-positive bacteria are classified into six types. Each of those profiles appears to be correlated to antigenically different proteins and related to different bacterial species. Type I bacterial IgG Fc-binding proteins are expressed by most *Staphylococcus aureus* strains and include the well-characterized staphylococcal protein A (SpA). Type II is associated with group A streptococcal strains, displaying a high heterogeneity in their binding patterns, and includes the M-proteins. Type III is found in most human isolates of groups C and G streptococcal, displaying the biggest diversity of species and subclasses reactivity, and includes the streptococcal protein G (SpG). Type IV is correlated with certain

bovine *β*-hemolytic group G streptococci. This type of Ig-binding proteins has a low affinity and limited species IgG reactivity. They have restricted reactivity towards human subclasses IgG1, IgG3, and IgG4, and some show significant Ig-binding activity only with human, rabbit, and pig immunoglobulins. Given their low affinity, the information about them is limited. Type V is expressed by certain strains of *Streptococcus zooepidemicus* within non-human isolates of group C streptococci. This type of Ig-binding proteins has a relative molecular weight of 45 kDa reacts to IgG1, IgG2, and IgG4, but not IgG3, as described for SpA; however, differs from SpA in their reactivity with IgG from various species. Besides those subclasses of human IgG, they strongly react only to pig, guinea pig, and rabbit IgGs. As for Type IV, the information about Type V is limited. Type VI is associated with one strain of *S. zooepidemicus* that displays a different immunoglobulin species-binding pattern from the one found in type V. Type VI Ig-binding proteins react to IgG from goat, pig, human, rabbit, sheep, rat, and cow. [26,27,28].

Besides the Ig-binding proteins comprised in this classification, many others have been identified to other immunoglobulin isotypes in several other different bacterial species[26]. Such as IgA Fc-Binding proteins associated with groups A and B streptococci[26], for example protein ARP an IgA-binding protein and protein Sir an IgG and IgA-binding protein associated with *Streptococcus pyogenes* that show ability to bind C4BP (a complement regulator), inhibiting complement at the bacterial surface (Table 1.2)[25,29,30,31] and β protein an IgA-binding protein associated with *Streptococcus agalactiae* that participates in the protection of the bacteria against IgA-mediated bacterial killing (Table 1.2)[25,32,33]. IgD-binding proteins, for example MID an IgD-binding protein associated with isolates of *Branhamella catarrhalis* that shows B cell stimulatory properties (Table 1.2)[34,35] and protein D a bacterial Ig-binding protein associated with *Haemophilus influenzae* that shows affinity to human IgD and a role in histamine release and degranulation of mast cells (Table 1.2)[25,36,37]. Also, the protein L (a light chain IgG-binding protein from a strain of *Finegoldia magna*)[38] and the protein P (a IgM Fc-binding protein from *Brucella abortus*)[27]. This chapter will emphasize more about IgG-binding proteins.

IgG-binding proteins of staphylococci

S. aureus presents Type I IgG Fc-binding activity. The first and far most studied bacterial non-immune Ig-binding protein is the type I bacterial IgG Fc-binding protein expressed in most *S. aureus* strains and usually designated as staphylococcal protein A (SpA)[27]. SpA was first identified by Forsgren *et al.* in 1966, describing the pseudoimmune reactivity of a surface protein from *S. aureus*[39]. Since then, the Ig-binding properties of this protein have been very well characterized[27]. SpA can be found on the cell surface of *S. aureus*, more particularly in the external layer of the cell wall, and in the culture

medium, being dependent on liquid nutrient medium. Furthermore, the location and expression of SpA within the cell seems to be dependent on the culture medium and on the bacterial strain[24,40]. SpA from different *S. aureus* strains show a range in their molecular weights from 45 to 57 kDa and show differences in their structure and functional activity[24,41]. The structure of this Ig-binding protein comprises a signaling sequence; five highly homologous IgG-binding domains (A-E) of 58-62 amino acid residues, that are highly hydrophilic and resistant to proteases, lack cysteine residues and are constituted by tightly packed antiparallel α-helices, stabilized by hydrophobic interactions between the α-helices; and a C-terminal region of 150 amino acids responsible for the protein fixation into the cell wall. The variability of SpA in weight and properties is due to the lack of some regions of these different elements[24].

SpA interacts with different classes of immunoglobulins (IgG, IgM, IgA, and IgE) and different subclasses of human or animal IgGs. The Ig-binding protein from *S. aureus* strongly interacts with human IgG1, IgG2, and IgG4 and with mouse IgG2a, IgG2b, and IgG3. Its interaction with IgG occurs in the Fc-fragment region between C_H2 and C_H3 domains of the heavy chain of the immunoglobulin, forming a complex with residues of the second and first α-helices of the Ig-binding domain of SpA, stabilized by hydrophobic interaction and some polar contacts. However, this interaction can also occur in the Fab-fragment of IgG, forming a complex between the second and third α-helices of SpA and the four β-strands of the variable region of the heavy chain[24,42].

This surface Ig-binding protein shows immune-modulatory attributes, being involved in the immune evasive strategies of *S. aureus* (Table 1.2). By interacting to the complement binding (Fcγ) portion of antibodies, on the cell surface of *S. aureus*, SpA impairs the recognition of surface antigens, promoting the inhibition of opsonophagocytic clearance. Also, SpA acts as a superantigen, modulating B cell responses to infection (this is promoted by the formation of SpA-IgM complexes on the surface of B cells) and promotes B cell depletion, when associated with complement factors[25,43].

Besides protein A, *S. aureus* encodes other proteins with IgG-binding activity, Sbi, IsaB, and SSL10[24]. Sbi is a predicted cell-surface protein with Ig-binding specificities similar to the ones of SpA. This protein interacts with the Fc-fragment of IgG and interacts with other serum proteins, such as the complement component C3 and the complement regulator Factor H contributing to immune evasion by stimulating inhibition of opsonization and complement inhibition (Table 1.2)[25,44,45,46]. The protein IsaB is an immunodominant protein that interacts with IgM and IgG[24,47]. The protein SSL10 is a staphylococcal superantigen-like protein that binds IgG and inhibits the interaction between IgG and C1q-component of the complement, impairing the complement activation through the classical pathway[24,48].

IgG-binding proteins of group A streptococci

Group A streptococci present Type II IgG Fc-binding proteins. The majority is encoded by M-protein genes[24,28]. M-proteins constitute a superfamily of proteins comprising the families Emm (class I and II), Mrp (FcrA), and Enn. They differ in structure and functional properties and play an important role in group A streptococci virulence[24]. M-proteins can be found on the microbial cell surface, forming structures similar to pili and fimbria of Gram-negative bacteria. The primary structure of these proteins consists of four regions (A-D) of repeated polypeptides and the conserved sorting signal region responsible for the protein fixation into the cell wall. The N-terminal A-repeats region constitutes the hypervariable region of the protein that is responsible for the serotypic differences between M-proteins. The central B-repeats region is variable in individual serotypes, but the C-repeats sequence is conserved among all serotypes. Variability among M-proteins is dependent on the regions A- and B-repeats and on the number of repeating units among them[24,49].

Not only M-proteins interact with different subclasses of human and animal IgGs, but they also interact with proteins of blood plasma, such as IgA, albumin, fibrinogen, kininogens, and the C4BP protein. The Ig-binding proteins from group A streptococci that interact with horse, swine, rabbit, and human IgG (only the IgG1, IgG2, and IgG4 subclasses) are grouped in the subtype of Ig-binding proteins IIa. The ones that interact only with the human IgG3 subclass are grouped in the subtype IIb. The subtype IIc includes Ig-binding proteins that interact with horse, swine, rabbit, and human IgG (only the IgG1 and IgG4 subclasses). Subtype IIo includes Ig-binding proteins that interact with horse, swine, rabbit and all subclasses of human IgG, and the subtype II'o represents the ones that only interact with rabbit and all subclasses of human IgG[24]. The IgG-binding sites of M-proteins differ among the several families. For example, in H proteins (Enn family), subtype IIa, there are two IgG-binding sites in the N-terminal A and B regions, whereas in M1 proteins (family Emm I), subtype IIb, there is only IgG-binding site located in the C-terminal sorting signal region[24,50]. As described for SpA, the interaction of Fc-fragment region between the C_H2 and C_H3 domains in IgG is also observed in M-protein, more specifically in FgBP protein[24].

M-proteins constitute virulence factors of group A streptococci. This cell surface Ig-binding protein promotes the inhibition of opsonophagocytic killing, that appears to be related to the binding of M-proteins to factor H leading to the regulation of complement deposition. Furthermore, protein H from the Enn family of M-proteins, besides showing to inhibit opsonophagocytotic killing, also shows complement inhibition, where the interaction between IgG and protein H inhibit the binding of the complement component Cq1 components to the surface of immunoglobulins (Table 1.2)[25,49,51].

Strains of group A streptococci also express other Ig-binding proteins besides M-proteins. Similar to what happens in *S. aureus*, the expression of the Ig-binding protein Sib was also observed in group A streptococci[24].

IgG-binding proteins of C and G group streptococci

C and G group streptococci present Type III IgG Fc-binding proteins. The first described Ig-binding protein of this type was the streptococcal protein G (SpG)[24,26]. SpG can be found on the bacterial surface and comprises a C-terminal sorting signal region, responsible for the protein fixation into the cell wall, a C-terminal IgG-binding domain with 55 amino acid residues formed by two antiparallel and two parallel β-strands with the α-helix situated along its diagonal, and an N-terminal Albumin-binding domain. This Ig-binding protein interacts with albumin, kininogen, the proteases α_2-macroglobulin inhibitor, and with all four subclasses of human and animal IgG. Interaction with IgG occurs through its C-terminal part of the α-helix, its N-terminal part of the third β-strand, and the loop region connecting these two structural elements, with the Fc-fragment region between the C_H2 and C_H3 domains. However, this interaction can also occur through the constant domain (C_H1) of the Fab-fragment heavy-chain, where SpG forms a complex with the Fab-fragment involving two β-strands that form an antiparallel β-sheet with the last β-strand of the C_H1 domain (Table 1.2)[24,52].

Strains of C and G group streptococci also express other Ig-binding proteins besides SpG, such as FOG, Dem A, Dem B, and FgBP that belongs to the family of M-proteins from the group A streptococci. They shared similar properties with type IIa Ig-binding proteins from group A streptococci[24].

Protein L from *Finegoldia magna*

The Gram-positive anaerobic microorganism *F. magna* also presents Ig-binding proteins, designated as PpL[24,38]. PpL is located at the cell surface and has the unique feature of non-immune Ig-binding activity towards the variable regions of the Fab-fragment located in the κ-type light chains of IgG (VL). This allows the binding of PpL to any subclass of antibodies. PpL from different strains of *F. magna* show a range in their molecular weights from 76 to 106 kDa, and is constituted by the C-terminal region responsible for the protein fixation into the cell wall and by four or five highly homologous Ig-binding domains. The interaction between PpL and the Fab-fragment forms a complex with a β-sheet formed by four β-strands and a central α-helix, stabilized by hydrophobic interactions and some

hydrogen bonds. PpL has two independent binding sites (with different affinities) that interact with two similar variable regions of the Fab-fragment located in the κ-type light chains[24,38,42].

This bacterial surface Ig-binding protein is associated with the virulence of *F. magna*. PpL constitutes a superantigen inducing basophils and mast cell (FcεRI^{+} cells) activation. Its interaction with the κ-type light chains of the IgE isotype promotes the secretion of important proinflammatory mediators, such as IL-3 and IL-4 from FcεRI+ cells that are implicated in bacterial infections (Table 1.2)[25,53].

1.2.2 IgG-Binding Proteins of Gram-Negative Bacteria

Gram-negative bacteria also express non-immune IgG-binding proteins, despite the limited number of known Ig-binding proteins and limited information about them. They can be found in the cell envelope, in the capsular material, or the culture medium[24]. The Gram-negative bacteria that present IgG-binding proteins include *Escherichia coli, Histophilus somni,* within the pathogenic *Yersinia* genus, *Y. pestis and Y. pseudotuberculosis*, *Stenotrophomonas maltophilia,* and *Helicobacter pylori*[24]. *E. coli* presents a family of Ig-binding proteins, Eib, comprising 6 elements, EibA, -C, -D, -E, -F, and -G. All members bind to the Fc-fragment of IgG, but some also bind to human IgA[54]. *H. somni* presents high-molecular-weight proteins that bind bovine IgG2, IgA, and IgM with high affinities, or a major 41 kDa outer membrane protein with weaker binding activities for both IgG1 and IgG2, and also for IgA and IgM. Besides bovine immunoglobulins, they also react with horse, rabbit, pig, cat, dog, and sheep IgG[24,55]. *Y. pestis* presents an IgG Fc-binding protein, PsaA, with a high molecular weight, forming pili on the surface of the plague microbe. PsaA only interacts with human IgG1, IgG2, and IgG3[24,56]. *Y. pseudotuberculosis* presents the thermostable hydrophilic 14.3 kDa protein and the OmdH/Skp chaperon protein. Skp has a compact β-structured central domain and a domain consisting of two elongated α-helix that form a hairpin with ends in the central domain. Skp interacts with the Fc-fragment of human and rabbit IgG, forming a complex with IgG1 through its α-helical region and the Fc-fragment of IgG1[24,57,58]. *S. maltophilia* presents a 30 kDa protein that interacts with the four subclasses of human IgG and with rabbit IgG, and also with mouse IgG and IgA[24,59].

As described for Gram-positive bacteria, for the PpL protein of *P. magnus* strains, gram-negative bacteria also express non-immune binding with immunoglobulins through the Fab region. *H. pylori* presents a 60 kDa protein, Hsp60, which interacts with the Fab-fragment region of human IgG1, IgG3, and IgM. This protein does not show reactivity to human IgA, rabbit IgG, and mouse IgG and its binding to Igs is inhibited by light κ-chains of human IgG[24,60].

1.3 Aspartic proteases of the retropepsin type

Aspartic proteases are proteolytic enzymes characterized for the use of an activated water molecule as the nucleophile responsible for hydrolyzing the peptide bond. The water molecule is activated by two aspartic acid residues in the active site of the enzyme, which bind to the water molecule and operate collectively to activate it[61]. According to the *MEROPS* database (a repository that offers a hierarchical, structure-based, classification of the proteolytic enzymes into classes, clans, and families), aspartic proteases are subdivided into 6 clans and 16 families (Table 1.3)[62].

Table 1.3- Hierarchical classification of aspartic proteases into clans and families, according to the *MEROPS* database. Adapted from Rawlings, N. D. *et al.* (2018)[62].

CLAN	FAMILY	TYPE PROTEASE
AA	A1	pepsin A (*Homo sapiens*)
	A2	HIV-1 retropepsin (human immunodeficiency virus 1)
	A3	cauliflower mosaic virus-type peptidase (cauliflower mosaic virus)
	A9	spumapepsin (human spumaretrovirus)
	A11	Copia transposon peptidase (*Drosophila melanogaster*)
	A28	DNA-damage inducible protein 1 (*Saccharomyces cerevisiae*)
	A32	PerP peptidase (*Caulobacter crescentus*)
AC	A8	signal peptidase II (*E. coli*)
AD	A22	presenilin 1 (*Homo sapiens*)
	A24	type 4 prepilin peptidase 1 (*Pseudomonas aeruginosa*)
AE	A25	gpr peptidase (*Bacillus megaterium*)
	A31	HybD peptidase (*E. coli*)
AF	A26	omptin (*E. coli*)
UNASSIGNED	A5	thermopsin (*Sulfolobus acidocaldarius*)
	A36	sporulation factor SpoIIGA (*Bacillus subtilis*)
	A37	sso1175 g.p. (*Sulfolobus solfataricus*)

One of the largest and best-characterized families of aspartic proteases is the family A2 (AA clan family), representing the aspartic proteases of the retropepsin type, also designated as retroviral proteases[63]. This family of proteases includes polyprotein-processing enzymes from positive-strand RNA viruses and retrotransposons[61]. The structures of several retroviral proteases have been reported, such as retroviral proteases from human immunodeficiency virus type 1 (HIV-1), type 2 (HIV-2), simian immunodeficiency virus (SIV), Rous sarcoma virus (RSV), feline immunodeficiency virus (FIV) and equine infectious anemia virus (EIAV)[63]. The vast interest in the study of these proteases is due to their essential role in the life cycle of viruses, being important therapeutic targets for several human diseases. The great necessity for developing antiviral drugs for HIV/AIDS fuelled the study of retroviral proteases, especially of the HIV-1 retroviral protease[64].

Retroviral proteases are very singular enzymes in their structural features, despite exhibiting some common characteristics of aspartic proteases, mostly of the A1 family (pepsin-like proteases). Like many other aspartic proteases, they are inhibited by pepstatin (at least partially) and inactivated by mutation of the active site aspartic acid residues[63]. In opposition to non-viral proteases characterized by single monomers with two topologically similar (not identical) domains, retroviral proteases are active as symmetric dimers with two identical monomers and one single active site constituted by residues of each one of the monomers. This is a unique feature among enzymes[63]. They are also much smaller than those monomers of more than 300 amino acids, characteristic of non-viral proteases. Each one of the retroviral monomers is in the range of 10-15 kDa[63,64].

This family of aspartic proteases is characterized by secondary structures that are based on a structural template shared with non-viral aspartic proteases. Retroviral proteases are homodimers, whose monomeric molecules resemble the single domains of non-viral aspartic proteases. The monomer is established by the duplication of four structural elements: a hairpin (A1), a wide loop (B1), containing the catalytic aspartate residue, an α-helix (C1) and a second hairpin (D1), the second identical monomer contains the same four structural elements, designated by A2, B2, C2, and D2 (Figure 1.3). Nonetheless, there is some variation of these elements in the known structures of retroviral proteases. The hairpin D2 is replaced by a β-strand. Only EIAV retroviral protease exhibits the α-helix C1; in RSV and FIV retroviral proteases this is displayed as a single helical turn; and in HIV (type 1 and 2) and SIV retroviral protease is substitute by a loop. The structural elements A1 and A2 also vary. The length of those loops and the length and conformation of the segments connecting them differs between the different retroviral proteases[63,65].

One of these structural elements, the β-loop D1, is very important for the catalytic activity of retroviral proteases. The flexible β-loop D1 (known as "flap" in non-viral retroviral proteases) changes orientation when binding to the ligand (substrate or inhibitor), resulting in several interactions between them. Since retroviral proteases are symmetric dimers, each one of the monomers has one of these flaps. The catalytic mechanism of retroviral proteases is characterized by an initial binding of the substrate to the active site, followed by the conformational change of the flaps (from open to close) by a downward moving, promoting the interactions between the enzyme and the substrate and strengthening the binding. The active site is located in loop B1 and is composed by two copies of the conserved Asp25, Thr26 (replaced by Ser38 in RSV), and Gly27. The two aspartic acid residues (Asp25 and Asp25' in HIV-1, HIV-2, SIV, and EIAV retroviral proteases, Asp30/30' in FIV retroviral protease, and Asp37/37' in RSV retroviral protease) are also important for catalytic activity. They are bridged by a water molecule, which is positioned by a hydrogen-bond length of the oxygen atoms of the Asp carboxylates. These interactions with the water molecule were also described in non-viral proteases. Therefore, the water molecule is likely to be the catalytic water necessary for the hydrolysis of the peptide bond[63,65].

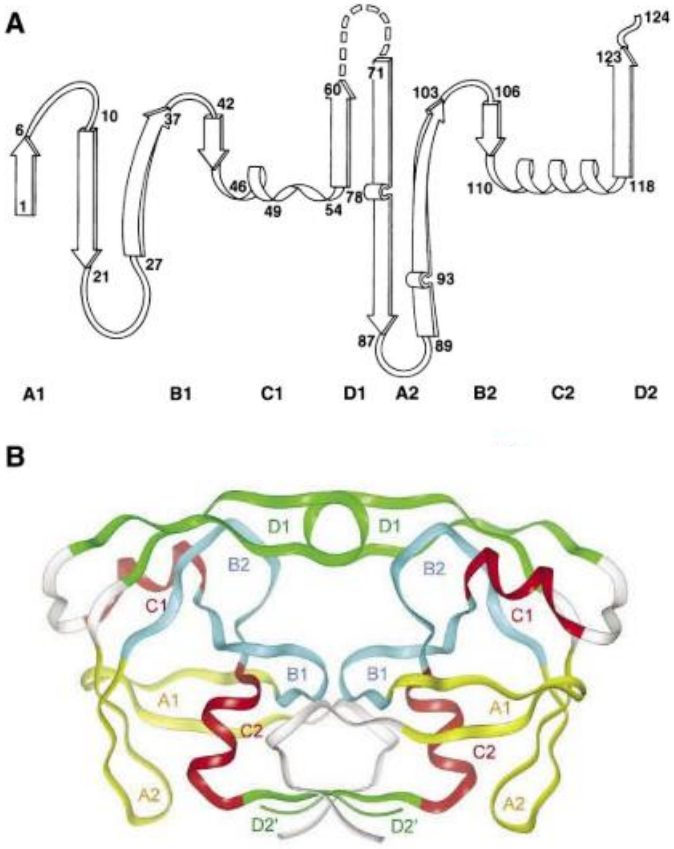


Figure 1.3- Secondary structure illustration of aspartic proteases of the retropepsin type. (A) Diagram of the secondary structure of retroviral proteases based on the RSV retroviral protease. (B) Three-dimensional representation of the symmetric retroviral dimer. The loops A1 and A2 are displayed in yellow, the loops B1 and B2, which contain the catalytic residues are displayed in blue. The helical segments C1 and C2 are displayed in red and the loop D1, that establishes the double flap structure, and the loop D2 are displayed in green. Adapted from Wlodawer, A. *et al*. (2000) and Dunn, B. *et al.* (2002)[63,65].

Retroviral proteases have maximum catalytic activity on peptide substrates under mildly acidic conditions and require dimerization to activity (the N and C termini form a four-stranded antiparallel β-sheet)[64]. The substrates for retroviral proteases are the polyproteins of their own virus, but other non-viral protein substrates have been reported (*e.g.*, vimentin, desmin, glial fibrillary acidic protein, ribonuclease A, lactate dehydrogenase, calcium-free calmodulin, actin, troponin C, Alzheimer amyloid precursor protein, and pro-interleukin 1β). The processing of substrates follows a precise order, beginning with the autocleavage of the retroviral protease in the N-terminal region. This autoprocessing activity promotes the release of the mature protease from the larger polyprotein precursor, allowing subsequent cleavage of other substrates[63,65].

The inhibition of proteolysis affects the virus replication; therefore, the inhibitors of these retroviral proteases constitute antiviral agents that can be used as therapeutic drugs[64]. As mentioned above, pepstatin inhibits both retroviral proteases and non-viral proteases[63]. For the HIV protease, the most studied retroviral protease, there are many inhibitors identified, for example Saquinavir, Ritonavir, Indinavir, Nelfinavir, Amprenavir, Lopinavir, Fosamprenavir Atazanavir, Tipranavir, and Darunavir[64,66].

Rickettsia's Aspartic Protease of the retropepsin type

Recently was described a novel aspartic protease with properties of retropepsin-like enzymes highly conserved in *Rickettsia*, designated as Aspartic protease from *Rickettsia conorii* (APRc). This constitutes the first indication of retropepsin-like enzymes in gram-negative intracellular bacteria[67].

The retropepsin-like aspartic protease is encoded by the gene RC1339 from *R. conorii*, which is highly conserved in 55 *Rickettsia* genomes. Despite the high degree of identity of APRc among *Rickettsia* spp., the similarity with aspartic proteases from other organisms is not significant, except for the conservation of the active site consensus motif Asp-Thr-Gly and the hydrophobic-hydrophobic-glycine motif, also conserved in other retroviral (and pepsin-like) proteases. Furthermore, the enzymatic characterization of the soluble catalytic domain of RC1339/APRc showed that this rickettsial enzyme is very similar in its properties with the family of retropepsin-like aspartic proteases. APRc shows autoprocessing activity affected by the mutation of the catalytic aspartate residue, accumulating in the dimeric form, even though most of the protein accumulates as a monomer, and optimal activity at pH 6, with no significant activity below pH 5. APRc also displays partial inhibition by specific HIV-1 retroviral protease inhibitors (*e.g.*, indinavir, nelfinavir, saquinavir, amprenavir, and atazanavir) with a slightly delay of the auto-activation process by pepstatin, and shows a unique specificity pattern with amino acid preferences similar to those of retroviral and pepsin-like proteases[67].

APRc is integrated into the outer membrane of *R. conorii* and is predicted to comprise three transmembrane domains at the N terminus and a soluble catalytic domain (Arg87-Tyr231), extracellular oriented, at the C terminus with the active site motif Asp140-Thr141-Gly142 [Figure 1.4 (A)]. Autoprocessing studies in *E. coli* revealed three auto-cleavage sites (Tyr92-Ala93, Met98-Ser99, and Ser104-Tyr105), with accumulation of the mature form Tyr105-Tyr23. APRc's enzymatic activity appears to be dependent on that mature form[67].

Structurally, APRc also resembles the retroviral family. The fold of the APRc monomer follows the structural template of retropepsins, with all of the structural elements (the hairpin A1, the wide loop, containing the catalytic aspartate residue, the α-helix and the second hairpin) present in the monomer of APRc [Figure 1.4 (B)]. However, the quaternary structure of its dimer does not resemble the canonical dimer of retropepsins. Since the soluble domain of APRc is an active enzyme, as described in Cruz, R. *et al* (2014), the observed dimer must be an artefact of the expression and/or crystallization conditions[68].

When compared with other retropepsins, APRc exhibits the α-helix C1 (residues 150–156), just like EIAV retroviral protease and the eukaryotic retropepsin Ddi1. The wide loop (residues 157-166) is longer and presents a distinct conformation from all the other structures of retropepsins to which it was compared to. This region usually is among the most variable regions in retropepsins. Also, the singularity of this wide loop has been reported to grant diverse immunological properties to several retroviral proteases. Overall, the major structural differences between the APRc monomers and retroviral proteases are in the lengths of the surface loops and in the lengths and conformations of the segments that connect them[68].

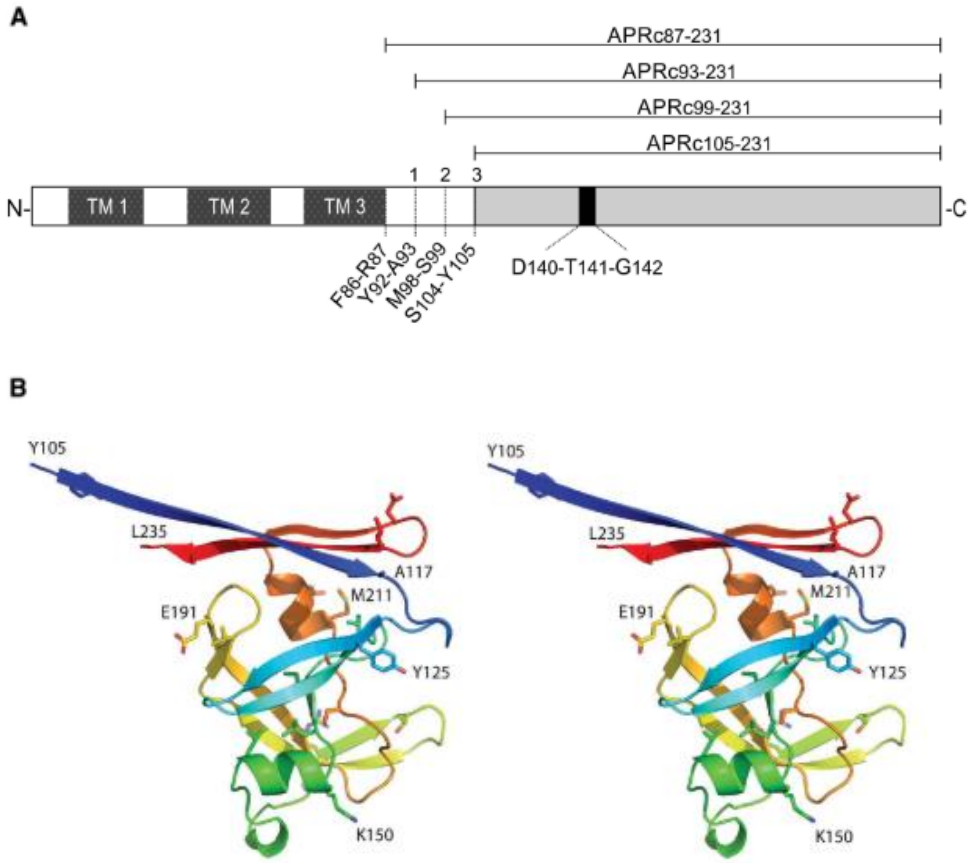


Figure 1.4 (previous page)- Representation of the aspartic protease from *Rickettsia conorii* (APRc). (A) Schematic representation of full-length APRc. APRc is predicted to include three transmembrane domains (TM1-3) at the N terminus, three auto-cleavage sites (1-Tyr92-Ala93, 2-Met98-Ser99 and 3-Ser104-Tyr105) and a soluble catalytic domain (Arg87-Tyr231) at the C terminus with the active site motif Asp140-Thr141-Gly142. (B) Secondary structure representation of the APRc monomer in rainbow colors, from blue at the N terminus to red at the C terminus. A β-strand starts after the N-terminal methionine and continues through Ala117. This is followed by two strands (Tyr125-Val130 and Val133-Val139), that form a β-hairpin, and by the central strand of a four-stranded β-sheet (Ile146-Leu148). From Lys150 through Leu156 there is an α-helix that is followed by a β-hairpin formed by the strands Arg167-Thr171 and Gly174-Val187. The last part of this strand forms another β-hairpin, including the strand Glu191–Gly200. Both of these strands belong to the four-stranded β-sheet, which is complete with the strand Ser207-Gly210. Met211-Glu215 forms an α-helix, that is followed by the last β-hairpin form by the strands Gly219–Asp223 and Leu226–Ala234. The last authentic residue in the sequence of APRc is the Tyr231, therefore the C-terminal strand goes through this residue, including part of the linker region to the C-terminal His-tag. Adapted from Cruz, R. *et al* (2014) and Li, M. *et al* (2015)[67,68].

This novel rickettsial protease is expressed in several rickettsial species [*R. conorii, R. rickettsii*[67], *R. parkeri, R. africae, R. massiliae, and R. montanensis* (unpublished data)] accumulating in the outer membrane with a soluble catalytic domain extracellular oriented, as shown in *E. coli*. Also, an up-regulation of APRc's gene expression in virulent *R. prowazekii* has been reported. This evidence sustains an important role of this protease in rickettsial life cycle[67,69]. Due to the unique features of APRc (the observable nonstringent sequence specificity, accumulation in the outer membrane with a catalytic domain extracellular oriented, and the autoprocessing activity that may indicate a release of the soluble catalytic domain of APRc from the surface of the cells by an ectodomain shedding-like process) it is hypothesized that this rickettsial protease comprises more than one functional role. It was reported that APRc catalyses the *in vitro* processing of rickettsial autotransporter proteins located at the outer membrane (Sca0/OmpA and Sca5/OmpB), that do not display autocatalytic activity as other autotransporter proteins from Gram-negative bacteria. Moreover, APRc may be involved in the degradation of host tissues for providing bacteria with nutrients, as reported for other proteases from pathogens, and might also sustain the spread of the infection and dissemination of bacteria into deeper tissue by shedding adhesion molecules at the cell surface or by disabling components of the host immune system[67].

1.4 Project aims

The very severe and life-threatening character of rickettsial diseases, together with their increasing incidence worldwide, results in a growing concern about rickettsioses and their severe impact on global health[9,11]. Currently, the treatment for these diseases is based solely on the use of antibiotics. Given the lack of reliable protective vaccines and the fact that the treatment is based on a few antibiotics (where antibiotic resistance was already observed), the identification of new protein factors that may work as potential therapeutic targets and/or the development of vaccines is extremely important. To achieve it, more information about the molecular mechanisms underlying the rickettsial pathogenesis and the protective innate and adaptive immune responses against rickettsiae is required[13].

In order to subsist and proliferate within the host, many pathogens have established different tactics to evade the host's immune responses and establish infection[23]. It was demonstrated that *Rickettsia* are resistant to the serum bactericidal effects and can evade complement-mediated killing[21] in a mechanism that is mediated by (at least) three rickettsial surface proteins (OmpB, Adr1, and Adr2) [18,19,22]. This clearly suggests that rickettsial species developed mechanisms to inhibit recognition by host serum components.

The novel rickettsial protease APRc is predicted to participate in the spreading of infection and dissemination of bacteria into deeper tissue by shedding adhesion molecules at the cell surface or by disabling components of the host immune system[67]. In the host laboratory, it was observed that this protein established a non-immune interaction with IgG. Since many bacteria present non-immune IgG-binding proteins with the purpose of scavenging the opsonizing antibodies, impairing the recognition by the opsonic receptors of the host organism and subsequently evading phagocytosis[23] or, alternatively, interfering with the deposition of complement components inhibiting the alternative pathway, we hypothesize that APRc is an IgG-binding protein that may act as a novel evasin contributing to rickettsial immune evasion toolbox.

The main goal of this project is to characterize in detail this non-immune Ig-APRc binding activity. APRc will be recombinantly produced in *E.coli* and then purified. The interaction between APRc and different immunoglobulins (immunoglobulins from different classes and species) will be evaluated. Since APRc is a protease, it will be evaluated if APRc combines Ig-binding with Ig-cleavage activity. Also, the regions in APRc and Ig responsible for binding will be mapped. The IgG-APRc binding activity in a more complex context (Normal Human Serum) will be evaluated. Moreover, non-immune Ig-binding at the surface of *Rickettsia* species will also be evaluated. This characterization will allow a better understanding of the rickettsial evasion mechanisms, providing essential information for the future development of new therapeutics for rickettsial diseases.

2 Materials and Methods

2.1 Reagents

The reagents used in all of the experiments were acquired from Sigma-Aldrich (Missouri, U.S.A), from Merck (Germany), or from NZYTech (Portugal). All primers were acquired from Integrated DNA Technologies (IDT), BVBA (Leuven, Belgium).

2.2 DNA Constructs

The expression construct pET_APRc$_{110-231}$-His (construct that includes the coding sequence of APRc amino acids 110-231 fused with the C-terminal tag sequence AAALEHHHHHH; pET-23d-based vector) was available in the laboratory and was generated from the construct that comprises the full-length APRc, as described in Li, M. *et al* (2015)[68]. The construct pET_APRc$_{110-231}$-HisShort (construct that includes the coding sequence of APRc amino acids 110-231 fused with the C-terminal tag sequence HHHHHH) and the active site mutants [pET_APRc$_{110-231}$-His(D140N) and pET_APRc$_{110-231}$-HisShort(D140N)], where the active site aspartic acid residue was replaced by asparagine, were also available in the lababoratory.

Different truncated forms of APRc were produced by PCR-based cloning. pET_APRc$_{110-231}$-HisShort was used as a template. The PCRs were performed according to the following conditions: 1x of Phusion buffer (containing 1.5 mM of $MgCl_2$), 200 µM of dNTP mix, 0.5 µM of each primer, 100 ng of DNA template, and 1 U of Phusion DNA Polymerase, in a total volume of 50 µL [the reagents were acquired from New England Biolabs (NEB) (Massachusetts, U.S.A)]. The PCR mixtures were incubated at 98˚C for 5 min, for the initial denaturation, followed by 32 cycles of incubations at 98˚C for 15 seconds, at 60˚C for 30 seconds, and at 72˚C for 6 minutes. And, for the final extension step, the PCR mixtures were incubated at 72ºC for 10 minutes.

For the generation of the construct pET_APRc$_{144-231}$-HisShort (construct coding amino acids 144-231), the sequence was amplified with the forward primer (5'-GCC TCT GAT ATT GCA CTG AC-3') and the reverse primer (5'-CAT ATG TAT ATC TCC TTC TTA AAG TTA AAC AAA-3'); for the construct pET_APRc(ΔLTKLK)$_{110-231}$-HisShort (construct coding amino acids 110-231 with the deletion of the sequence LTKLK, residues 160 to 164), the sequence was amplified with the forward primer (5'-TAT ACC CGT ACG TAC CTG ACG-3') and the reverse primer (5'-ATC GAA ACC CAG TTT CTG C-3'); for the construct pET_APRc(ΔGFDLTKLKYT)$_{110-231}$-HisShort (construct coding amino acids 110-231 with the deletion of the sequence GFDLTKLKYT, residues 157 to 166), the sequence was amplified with the forward primer (5'-CGT ACG TAC CTG ACG GCC-3') and the reverse primer (5'-CAG TTT CTG CGC ATC TTC TT-3'); for the construct

pET_APRc(ΔKEDAQKLGFDLTKLKYT)$_{110-231}$-HisShort (construct coding amino acids 110-231 with the deletion of the sequence KEDAQKLGFDLTKLKYT, residues 150 to 166), the sequence was amplified with the forward primer (5'- CGT ACG TAC CTG ACG GCC -3') and the reverse primer (5'-CGT CAG TGC AAT ATC AGA GG-3'); for the construct pET_APRc$_{110-173}$-HisShort (construct coding amino acids 110-173), the sequence was amplified with the forward primer (5'-CAC CAC CAC CAC CAC CAC -3') and the reverse primer (5'-GTT GGC CGT CAG GTA CGT-3'); for the construct pET_APRc$_{110-189}$-HisShort (construct coding amino acids 110-189), the sequence was amplified with the forward primer (5'-CAC CAC CAC CAC CAC CAC -3') and the reverse primer (5'-GCC GAT AAC CAC GCT GTT-3'); for the construct pET_APRc$_{110-218}$-HisShort (construct coding amino acids 110-218), the sequence was amplified with the forward primer (5'-CAC CAC CAC CAC CAC CAC -3') and the reverse primer (5'-TTT AAA ACG TTC CAG CAG AGA C-3'); and for the construct pET_APRc$_{110-225}$-HisShort (construct coding amino acids 110-225), the sequence was amplified with the forward primer (5'-CAC CAC CAC CAC CAC CAC -3') and the reverse primer (5'-ATC TTT ATC GAT GCG GAA ACC-3').

The amplification products were confirmed by agarose gel electrophoresis, as detailed in section 2.14, and then digested with DpnI (1 µL) (NEB) for 2 hours at 37ºC. Afterwards, the PCR products were liquid purified using the Macherey-Nagel™ NucleoSpin™ Gel and PCR Clean-up kit (Macherey-Nagel, Germany) and quantified. Phosphorylation and ligation were performed for 3 hours at room temperature, according to the following conditions: 1X T4 DNA Ligase Reaction Buffer (50 mM Tris-HCl, pH 7.5, containing 10 mM MgCl2, 1 mM ATP and 10 mM DTT), 1X T4 DNA Ligase, 1 µL of T4 Kinase PNK, 1 µL of T4 DNA ligase, and 100 ng of DNA, in a total volume of 20 µL (the reagents were acquired from NEB). The ligation product was transformed in *E. coli* TOP10F' competent cells. Positive clones were inoculated in 10 mL of LB medium containing 100 µg/mL ampicillin, overnight at 37ºC with agitation. The cells were harvested by centrifugation at 6000 x*g* for 10 minutes in a Heraeus Megafuge 40 centrifuge (Thermo Fisher), at 4ºC. Pure plasmid DNA was obtained with a Macherey-Nagel™ NucleoSpin Plasmid QuickPure™ kit (Macherey-Nagel) and confirmed by DNA sequencing.

The construct used for the production of biotinylated APRc (pCoofy_HisAPRc$_{110-231}$-avi; pCoofy-1-based vector[70]) was available in the laboratory, and included the coding sequence of APRc amino acids 110-231 fused at the N terminus with a His-tag followed by a HRV 3C cleavage site, and at the C terminus with a 15 amino acid avi-tag peptide (biotin accepting peptide).

2.3 Small-scale expression screening of truncated forms of APRc

For small-scale expression screening, the following truncated forms of APRc: pET_APRc$_{110-231}$-HisShort, pET_APRc$_{144-231}$-HisShort, pET_APRc(ΔLTKLK)$_{110-231}$-HisShort,

pET_APRc(ΔGFDLTKLKYT)$_{110-231}$-HisShort, pET_APRc(ΔKEDAQKLGFDLTKLKYT)$_{110-231}$-HisShort, pET_APRc$_{110-173}$-HisShort, pET_APRc$_{110-189}$-HisShort, pET_APRc$_{110-218}$-HisShort, and pET_APRc$_{110-225}$-HisShort, were transformed in *E.coli* BL21 Star (DE3) competent cells. Then, the transformed cells were grown in 1 mL of LB medium containing 100 µg/mL ampicillin overnight at 37ºC with agitation. The pre-inocula were then diluted in 2 mL of LB medium containing 100 µg/mL ampicillin to an initial OD_{600nm} of 0.05, and then grown at 37ºC with agitation. When an OD_{600nm} of 0.6 was reached, the expression was induced with IPTG, at a final concentration of 0.1 mM,, overnight at 20ºC with agitation. Before cell lysis, E. *coli* growth was assessed (OD_{600nm}) and for protein extraction, a cell density corresponding to an OD_{600nm} of 3 (per condition) was harvested by centrifugation at 16 000 x*g* for 5 minutes in a Sigma 1-14 centrifuge (Sigma-Aldrich), at room temperature. The pellets were resuspended in BugBuster Protein Extraction Reagent (Novagen, Wisconsin, U.S.A) for 20 minutes at room temperature with agitation. The total fractions were denatured with 6xloading buffer (Tris-HCl (0.35 M)/SDS (0.28%) buffer pH 6.8, with 30% Glycerol, 10% SDS, 0.012% Bromophenol Blue and with 0.6 M DTT), 9 min at 90ºC. Protein samples were separated by SDS-PAGE followed by Coomassie blue staining and analysed by Far-Western with rabbit anti-Mouse IgG [HRP] antibody, as detailed in section 2.14.

2.4 Scale-up expression and purification of recombinant APRc

2.4.1 Expression for production of recombinant protein

For scale-up expression of APRc$_{110-231}$-HisShort, APRc$_{110-231}$-His and the respective active site mutants [APRc$_{110-231}$-HisShort(D140N), APRc$_{110-231}$-His(D140N)], the constructs were independently transformed in *E.coli* BL21 Star (DE3) competent cells. Then, the transformed cells were diluted in 5 mL of LB medium containing 100 µg/mL Ampicillin. Fernbach flasks with 1 L of LB medium containing 100 µg/mL ampicillin were inoculated with 2 mL of the pre-inocula, and then grown at 37ºC with agitation. When an OD_{600nm} of 0.6 was reached, the expression was induced with IPTG, at a final concentration of 0.1 mM, for 3 hours at 30ºC with agitation. The cells were harvested by centrifugation at 6000 x*g* for 20 min at 4ºC in an Avanti J-26S XPI Centrifuge with a JLA 8.1000 rotor (Beckman Coulter Life Sciences, Indiana, U.S.A). The pellets were resuspended in 40 mL of 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl. At last, the harvested cells were frozen at -20ºC, until further use.

2.4.2 Purification by immobilized metal affinity chromatography with Ni^{2+} and ion-exchange chromatography

For purification by immobilized metal affinity chromatography with Ni^{2+} (IMAC-Ni^{2+}), the frozen cells were thawed at room temperature and then applied to an Emulsiflex^{R}-C3 (Avestin, Ontario, Canada), a high-pressure homogeniser, to induce cell lysis. Afterwards, the lysates were centrifuged at 24 000 x*g*, at 4ºC for 20 minutes, in an Avanti J-26S XPI Centrifuge with a JA 25.50 rotor, followed by ultracentrifugation at 315 594 x*g*, at 4ºC for 20 minutes, in an Optima l-100 XP ultracentrifuge with a 90Ti rotor (Beckman Coulter Life Sciences). The resultant supernatants were applied onto a Histrap HP 5 mL column (Cytiva, Massachusetts, U.S.A), previously equilibrated in 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl. Protein elution was carried out in a BioLogic DuoFlow^{TM} chromatography system (Bio-Rad, California, U.S.A) by implementation of a four-step gradient of imidazole (50 mM, 150 mM, 200 mM, and 500 mM) in 20 mM sodium phosphate pH 7.4 buffer containing 500 mM NaCl, at a flow rate of 5 mL/min. The proteins of interest were eluted in the 150 mM/200 mM of imidazole steps. Those fractions were independently pooled and dialyzed in Spectra/Por^{TM} 3 (3.5kDa) dialysis membranes (Spectrum Labs, California, U.S.A) against 20 mM HEPES pH 7.4 buffer, overnight at 4º C with agitation. The dialyzed proteins were ultracentrifuged at 315 594 x*g* at 4ºC for 20 minutes in an Optima l-100 XP ultracentrifuge with a 90Ti rotor, and the resultant supernatants were further purified by cation exchange chromatography, in a Mono-S 5/50 GL column (Cytiva), previously equilibrated with 20 mM HEPES pH 7.4 buffer. Protein elution was carried out in a BioLogic DuoFlow^{TM} chromatography system by implementation of a linear gradient of NaCl (0-1 M) in 20 mM HEPES pH 7.4 buffer, at a flow rate of 0.75 mL/min. Eluted protein samples were subsequently analysed by SDS-PAGE followed by Coomassie blue staining, as described in section 2.14.

2.4.3 Analytical-size Exclusion chromatography

The oligomerization state of the different recombinant forms of APRc was evaluated by analytical-size exclusion chromatography. The protein samples were injected in a Superdex 200 5/150 GL column (Cytiva) previous equilibrated in 20 mM HEPES buffer pH 7.4 containing 100 mM NaCl, using a Prominence Shimadzu HPLC system (Shimadzu, Tokyo, Japan). The molecular weight of proteins was estimated by calibration of the column with standard proteins from the Gel Filtration LMW calibration kit (Cytiva): aprotinin (6.5 kDa), ribonuclease A (13.7 kDa), carbonic anhydrase (29 kDa),

ovalbumin (43 kDa), and conalbumin (75 kDa). Protein elution was carried out at a flow rate of 0.25 mL/min and monitored at an absorbance of 220 nm.

2.4.4 Reverse-phase HPLC

The proteolytic activity of the different recombinant forms of APRc was evaluated by Reverse-Phase HPLC, using oxidized insulin β-chain as the substrate. 2.5 µg of each recombinant form of APRc were incubated overnight at 37ºC with 50 µg of oxidized insulin β-chain in 0.1 M sodium acetate buffer, pH 6. The samples were centrifuged for 10 minutes at 16 000 x*g* in a Sigma 1-14 centrifuge, at room temperature. The supernatant was applied in a KROMASIL 100 C18 250 4.6 mm column (Teknokroma, Spain) previous equilibrated in 0.1% trifluoroacetic acid (TFA). The peptides were separated by Reverse-Phase HPLC, using a Prominence Shimadzu HPLC system. The elution of peptides was carried out with a linear gradient of acetonitrile (0-80%) in 0.1% TFA, at a flow rate of 1 mL/min, and absorbance monitored at 220 nm.

2.5 Non-immune APRc Ig-binding and Ig-cleavage activity assays

To evaluate if APRc combines Ig-binding with Ig-cleavage activity, incubations of recombinant APRc with IgG from different species were performed under different conditions, as detailed below.

Recombinant $APRc_{110-231}$-HisShort was incubated with polyclonal rabbit IgG specific for a plant-protease epitope (produced by Genscript) in an equimolar ratio (529.35 ρmole) in PBS buffer pH 7.4 (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4), for 2 hours at 37ºC.

Recombinant $APRc_{110-231}$-HisShort and $APRc_{110-231}$-HisShort(D140N) (dimeric forms), were incubated with human IgG (#I2511, Sigma-Aldrich) at different molar ratios [equimolar (333.7 ρmole) /1:0.1/1:0. 05 Substrate:Enzyme (S:E)], in PBS buffer pH 7.4, for 4h at 37ºC.

Recombinant $APRc_{110-231}$-HisShort and $APRc_{110-231}$-HisShort(D140N) (dimeric forms), were incubated with human IgG (#I2511, Sigma-Aldrich) in an equimolar ratio (333.7 ρmole), in PBS buffer at different pH values (pH 7.4/pH 6), for 4h at 37ºC.

Recombinant $APRc_{110-231}$-HisShort, dimeric and monomeric forms, were incubated with human IgG (#I2511, Sigma-Aldrich) at a 50 µg:6 µg ratio (S:E), in PBS buffer pH 7.4, overnight at 37ºC.

The samples were denatured with 6x loading buffer (Tris-HCl (0.35 M)/SDS (0.28%) buffer pH 6.8, with 30% Glycerol, 10% SDS, 0.012% Bromophenol Blue) without DTT, 9 min at 90ºC, and analysed by Western blot with rabbit anti-APRc antibody followed by mouse anti-Rabbit IgG (M205) [HRP] antibody, or with quail anti-APRc antibody followed by rabbit anti-Chicken IgY (IgG) [HRP]

antibody, or with mouse anti-Human IgG Fc [HRP] antibody, or with rabbit anti-Mouse IgG [HRP] antibody, or with mouse anti-Rabbit IgG (M205) [HRP] antibody, as described in section 2.14.

2.6 Production of biotinylated APRc

2.6.1 Expression for production of recombinant protein

Biotinylated APRc was generated by enzymatic biotinylation, where the protein, fused with a biotin accepting peptide (avi-tag), was co-expressed with biotin protein ligase (birA). The avi-tag functions as a recognition site for birA, enabling biotinylation to occur in the presence of biotin. Therefore, pCoofy_HisAPRc$_{110-231}$-avi was transformed in *E. coli* BL21 Star (DE3) competent cells containing the pDW363ΔMBP plasmid, an expression vector that allows the expression of birA. Then, the transformed cells were inoculated in 100 mL of LB medium containing 100 µg/mL ampicillin and 50 ug/mL kanamycin, overnight at 37ºC with agitation. Fernbach flasks with 1 L of LB medium containing 100 ug/mL ampicillin, 50 ug/mL kanamycin, and 50 µM biotin were inoculated with 20 mL of the pre-inoculum, and then grown at 37ºC with agitation. When an OD_{600nm} of 0.6 was reached, the expression was induced with IPTG, at a final concentration of 0.1 mM, overnight, at 20ºC with agitation. The cells were harvested by centrifugation at 6 000 x*g* for 20 min at 4ºC in an Avanti J-26S XPI Centrifuge with a JLA 8.1000 rotor. The pellets were resuspended in 40 mL of 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl and frozen at -20ºC, until further use.

2.6.2 Purification by immobilized metal affinity chromatography with Ni^{2+} and ion-exchange chromatography

For purification by immobilized metal affinity chromatography with Ni^{2+} (IMAC-Ni^{2+}), the frozen harvested cells were thawed at room temperature and then applied to an Emulsiflex^{R}-C3, a high-pressure homogeniser, in order to induce cell lysis. Afterwards, the lysates were centrifuged at 24 000 x*g*, at 4ºC for 20 minutes, in an Avanti J-26S XPI Centrifuge with a JA 25.50 rotor, followed by ultracentrifugation at 315 594 x*g* at 4ºC for 20 minutes in an Optima l-100 XP ultracentrifuge with a 90Ti rotor. The resultant supernatants were applied onto a Histrap HP 5mL column, previously equilibrated in 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl. Protein elution was carried out in a BioLogic DuoFlow^{TM} chromatography system by implementation of a four-step gradient of imidazole (50 mM, 150 mM, 200 mM, and 500 mM) in 20

mM sodium phosphate pH 7.4 buffer containing 500 mM NaCl, at a flow rate of 5 mL/min. The protein of interest was eluted in the 150 mM of imidazole step. Those fractions were pooled and dialyzed in Spectra/Por™ 3 (3.5kDa) dialysis membrane against 20 mM Tris/HCl pH 8 buffer, overnight at 4ºC with agitation. The dialyzed proteins were ultracentrifuged at 315 594 x*g* at 4ºC for 20 minutes in an Optima l-100 XP ultracentrifuge with a 90Ti rotor, and the resultant supernatants were further purified by anion exchange chromatography in a Mono-Q 5/50 GL column (Cytiva), previously equilibrated in 20 mM Tris/HCl pH 8 buffer. Protein elution was carried out in a BioLogic DuoFlow™ chromatography system by implementation of a linear gradient of NaCl (0-1 M) in Tris/HCl pH 8 buffer, at a flow rate of 0.75 mL/min. Protein samples were separated by SDS-PAGE followed by Coomassie blue staining and analysed by Western-Blot with HRP-conjugated Streptavidin, as described in section 2.14.

2.6.3 His-tag cleavage

To remove the His-tag at the N terminus of recombinant APRc, the protein was digested at the HRV 3C consensus cleavage site downstream of the His-tag sequence. The collected fractions were incubated with 2 µL HRV 3C protease in 50 mM Tris-HCl pH 7.5 buffer containing 150 mM NaCl overnight at 4ºC. Afterwards, the digestion was incubated with Ni Sepharose High Performance beads (Cytiva) for 15 minutes at room temperature with agitation. The suspension was filtered through 0.2 µm filters to remove the Ni Sepharose High Performance beads. A dialysis in Spectra/Por™ 3 (3.5 kDa) dialysis membrane was performed against PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4, pH 7.4) for 3 hours at 4ºC with agitation. Protein samples were separated and analyzed by SDS-PAGE followed by Coomassie blue staining, as described in section 2.14.

2.7 ELISA

The interaction between biotinylated APRc and different immunoglobulins was quantified by sandwich enzyme-linked immunosorbent assay (ELISA) with streptavidin-biotin detection. Human IgG (#I2511, Sigma-Aldrich), mouse IgG (#I5381, Sigma-Aldrich), polyclonal rabbit anti-shewasin D antibody (produced by Genscript)[71], human IgA (#I4036, Sigma-Aldrich), and human IgM (#I8260, Sigma-Aldrich) were evaluated. Nunc MaxiSorp™ high protein-binding capacity 96 well ELISA plates (Thermo Fisher Scientific, Massachusetts, U.S.A) were coated with 1 µg of each immunoglobulin per well followed by incubation at 37ºC for 2 hours. The wells were washed 3 times with 200 µL of PBS-T buffer (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4, containing 0.05% Tween 20, pH 7.4). Afterwards, the remaining protein-binding sites in the coated wells were blocked by incubation

at 37ºC for 2 hours with 200 µL of PBS-T buffer containing 3% of BSA. The wells were washed 4 times with 200 µL of PBS-T buffer. Then, 50 µL of biotinylated APRc diluted in PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4, pH 7.4) at different concentrations (0 µg/mL, 20 µg/mL, 50 µg/mL, 200 µg/mL, 450 µg/mL, 600 µg/mL, and 800 µg/mL) were added per well, followed by incubation at 37ºC for 1 hour. The wells were washed 4 times with 200 µL of PBS-T buffer. HRP-conjugated Streptavidin (Cell Signaling Technology, Massachusetts, U.S.A) (1:2000) in PBS-T buffer containing 3% of milk was added to which well and incubated at 37ºC for 1 hour. The wells were washed 4 times with 200 µL of PBS-T buffer. Afterwards, 100 µL of 1-Strep™ Ultra TMB ELISA substrate (Thermo Fisher Scientific) was added to which well and incubated at room temperature for 15/20 minutes (in the dark). To stop the reaction, 100 µL of a solution of 2 M sulfuric acid was added per well. At last, the absorbance was measured at 450nm in a BioTek™ PowerWave™ Microplate Spectrophotometer (BioTek,Vermont, U.S.A).

2.8 Production and purification of different fragments of human IgG

For F(ab')$_2$ production, 0.79 mg/mL of human IgG (#I2511, Sigma-Aldrich) were digested with 0.04 mg/mL of pepsin in 50 mM sodium acetate pH 4.0 buffer containing 100 mM NaCl, for 8 hours and 30 minutes at 37ºC. To stop the digestion, 10 µM of pepstatin were added. The sample was diluted in 1 mL of PBS buffer (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4, pH 7.4) and then concentrated in a Vivaspin 500, 50 kDa PES column (Sartorius, France), previously equilibrated in PBS. This last procedure was repeated 3 times. To remove the Fc fragments resultant from the digestion of IgG with pepsin, the sample was incubated for 2 hours, with agitation, at room temperature, with 80 µL of protein A Mag Sepharose slurry (Cytiva), previously washed 3 times with 1 mL of PBS (for the washes, PBS buffer was added to the slurry, mixed, and the resins separated with a MagRack™ 6 (Cytiva) to remove the supernatant). After the incubation, the resins were separated with a MagRack™ 6 and the supernatant containing F(ab')$_2$ was collected. Protein samples were separated by SDS-PAGE followed by Coomassie blue staining and analysed by Western blot with mouse anti-human IgG Fc [HRP] antibody or with rabbit anti-Mouse IgG [HRP] antibody, as described in section 2.14.

For F(ab') production, 1.26 mg/mL of human F(ab')$_2$ (pepsin-digested samples containing Fc fragments) were digested with 0.2 mg/mL of papain in PBS pH 7.4 buffer containing 0.02 M EDTA and 0.02 M cysteine for 1 hour and 30 minutes at 37ºC. The sample was diluted in 20 mL of PBS pH 7.4 buffer, and then concentrated in an Amicon^R Ultra-4 Centrifugal Filter Unit (Ultracel 3000 Da) (Merck), previously equilibrated in PBS pH 7.4 buffer. To remove the Fc fragments, the sample was incubated for 2 hours, with agitation, at room temperature, with 80 µL of protein A Mag Sepharose slurry (Cytiva),

previously washed 3 times with PBS buffer (as described above). After the incubation, the resins were separated with a MagRack™ 6 and the supernatant was collected. To stop the digestion, 10 µM of E-64 protease inhibitor were added. Protein samples were separated by SDS-PAGE followed by Coomassie blue staining and analysed by Western blot with mouse anti-Human Ig Kappa Light Chain antibody followed by rabbit anti-Mouse IgG [HRP] antibody, or with mouse anti-human IgG Fc [HRP] antibody, as described in section 2.14.

2.9 Cross-linking reactions with glutaraldehyde

To evaluate the interaction region to APRc in IgG, cross-linking reactions with glutaraldehyde between APRc and different fragments of human IgG were performed. Recombinant $APRc_{110-231}$-HisShort and $APRc_{110-231}$-HisShort(D140N) were incubated with intact human IgG (#I2511, Sigma-Aldrich), purified human $F(ab')_2$, and human F(ab') fragments, in an equimolar ratio (166.85 ρmole), for 4 hours at 37ºC. Afterwards, the samples were treated with 2 µL of a 1.15% glutaraldehyde solution, in a total volume of 20 µL. As a control, similar incubations were treated with 2 µL of H_2O. All the samples were incubated for 4 min at 37ºC, and the reaction was stopped by adding 2 µL of a solution of 1 M Tris-HCl pH 8.0. The samples were denatured with 6x loading buffer (Tris-HCl (0.35 M)/SDS (0.28%) buffer pH 6.8, with 30% Glycerol, 10% SDS, 0.012% Bromophenol Blue and with 0.6 M DTT), 9 min at 90ºC. Protein samples were analysed by Western blot with rabbit anti-APRc antibody followed by mouse anti-Rabbit IgG (M205) [HRP] antibody, or with mouse anti-Human IgG Fc [HRP] antibody, or with mouse anti-Human Ig Kappa Light Chain antibody followed by rabbit anti-Mouse IgG [HRP] antibody, as described in section 2.14.

2.10 Pull-down assay with His Mag Sepharose Ni Beads

Non-immune APRc-Ig interaction in serum samples was evaluated by pull-down assay with His Mag Sepharose Ni Beads. $APRc_{110-231}$-HisShort (85 µg) or $APRc_{110-231}$-HisShort(D140N) (85 µg) diluted in 300 µL of PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4, pH 7.4), was incubated with agitation, for 1 hour and 30 minutes, at room temperature, with 25 µL of His Mag Sepharose Ni Beads slurry (Cytiva), previously washed 3 times with PBS buffer (for the washes, 1 mL PBS buffer was added to the slurry, mixed, and the beads separated with a MagRack™ 6 (Cytiva) to remove the supernatant). After incubation with APRc, the magnetic beads were separated with a MagRack™ 6 and the supernatant was discarded. The beads were washed 3x with 1 mL of PBS buffer, as described previously, and then incubated with 240 µL of normal human serum (NHS) (Zenbio, North

Carolina, U.S.A)/ or normal umbilical cord human serum) diluted 15x in PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4, pH 7.4), for 4 hours at 37º C with agitation. After incubation, the beads were separated with a MagRack™ 6 and the supernatant was collected. The beads were washed 3 times, first with 1 mL of PBS buffer, then with 1 mL of PBS buffer containing 200 mM NaCl, and at last with 1 mL of PBS buffer containing 0.1 % Tween. Lastly, 50 µL of elution buffer (loading buffer (Tris-HCl (0.35 M)/SDS (0.28%) buffer pH 6.8, with 30% Glycerol, 10% SDS, 0.012% Bromophenol Blue and with 0.6 M DTT) diluted 6x in PBS) was added to the separated beads, mixed, and denatured 9 min at 90ºC. The beads were separated with a MagRack™ 6 and the supernatant was collected. Protein samples were separated by SDS-PAGE followed by Coomassie blue staining and analysed by Western blot with mouse anti-Human IgG Fc [HRP] antibody, as described in section 2.14.

2.11 Immunoprecipitation assay with Protein A Mag Sepharose

Non-immune APRc-Ig interaction in normal human serum (NHS) samples was evaluated by immunoprecipitation assay with protein A Mag Sepharose. 240 µL of NHS (Zenbio) diluted 15x in PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4, pH 7.4) was incubated with 62.5 µg of $APRc_{110-231}$-HisShort or 62.5 µg $APRc_{110-231}$-HisShort(D140N), for 4 hours at 37ºC. Afterwards, these suspensions were incubated for 35 min, with agitation, at room temperature, with 40 µL of protein A Mag Sepharose slurry, previously washed 3 times with TBS buffer (50 mM Tris, 150 mM NaCl, pH7.5) (for the washes, 1 mL TBS buffer was added to the slurry, mixed, and the resins separated with a MagRack™ 6 (Cytiva) to remove the supernatant). After incubation, the resins were separated with a MagRack™ 6 and the supernatant was collected. The resins were washed 3x, first with 1 mL of TBS buffer, then with 1 mL of TBS buffer containing 250 mM NaCl, and at last with 1 mL of TBS buffer containing 1 % Octyl-β-Glucoside. Lastly, 50 µL of elution buffer (loading buffer (Tris-HCL (0.35 M)/SDS (0.28%) buffer pH 6.8, with 30% Glycerol, 10% SDS, 0.012% Bromophenol Blue and with 0.6 M DTT) diluted 6x in PBS) was added to the separated resins, mixed, and denatured 9 min at 90ºC. The resins were separated with a MagRack™ 6, and the supernatant was collected. Protein samples were separated by SDS-PAGE followed by Coomassie blue staining and analysed by Western blot with rabbit anti-APRc antibody followed by mouse anti-Rabbit IgG (M205) [HRP] antibody, as described in section 2.14.

2.12 APRc Ig-cleavage activity assay in serum samples

The APRc Ig-cleavage activity was evaluated in serum samples. Firstly, the depletion of human serum albumin (HSA) was performed. 750 µL of normal umbilical cord human serum diluted 15x in digestion buffer (50 mM sodium phosphate pH 6.5 buffer, containing 100 mM NaCl) was incubated with agitation for 35 minutes at room temperature with 80 µL of Cibacron blue agarose (Santa Cruz Biotechnology, Texas, U.S.A), previously washed 3 times with digestion buffer by centrifugation for 1 minute at 16 000xg in a Sigma 1-14 centrifuge at room temperature. After the incubation, the suspension was sedimented by centrifugation for 1 minute at 16 000x*g* in a Sigma 1-14 centrifuge at room temperature and the supernatant was collected. A digestion reaction of human normal umbilical cord serum depleted from HSA (NHSΔHSA) with APRc was then performed. 80 µg of $APRc_{110-231}$-HisShort or 80 µg of $APRc_{110-231}$-HisShort(D140N) were incubated with 213 µL of NHSΔHSA, for 5 hours at 37ºC in digestion buffer. As a control, the NHSΔHSA was incubated with digestion buffer and with APRc's dilution buffer (20 mM HEPES buffer pH 7.4 containing 100 mM NaCl), only. The samples were denatured with 6x loading buffer (Tris-HCl (0.35 M)/SDS (0.28%) buffer pH 6.8, with 30% Glycerol, 10% SDS, 0.012% Bromophenol Blue and with 0.6 M DTT), 9 min at 90ºC and separated by SDS-PAGE followed by Coomassie blue staining, as described in section 2.14.

2.13 Non-immune Ig-binding at the surface of *Rickettsia* species assay

Non-immune Ig-binding at the surface of *Rickettsia* species was evaluated. In this assay *R. africae* was used as a representative specie of SFG *Rickettsia*. *R. africae* (6.45×10^8 PFU) were fixed with 500 µL of 4% Peroxyformic acid (PFA), for 20 minutes at room temperature. The PFA-fixed *Rickettsia* were washed 2 times with 1 mL of PBS buffer (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4, pH 7.4) by centrifugation for 3 minutes at 16 000 x*g* in a Sigma 1-14 centrifuge at room temperature, and then independently incubated with 275 µg of human IgG (#I2511, Sigma-Aldrich), 50 µL of NHS diluted 2x in PBS buffer (contains approximately 275 µg of IgG), or 50 µL of PBS buffer, for 2 hours at 37ºC with agitation. Each sample was washed 2 times with 1 mL of PBS buffer by centrifugation for 3 minutes at 16 000x*g* in a Sigma 1-14 centrifuge at room temperature. The bacteria were resuspended in 40 µL of PBS buffer containing 1 M NaCl and incubated for 20 minutes at room temperature. The remaining bacteria was removed by centrifugation for 3 minutes at 16 000x*g* in a Sigma 1-14 centrifuge at room temperature and the cell-free elution fractions analysed by Western blot with mouse anti-human IgG Fc [HRP] antibody, as described in section 2.14.

2.14 Gel electrophoresis

2.14.1 SDS-PAGE analysis

Protein samples were analysed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Firstly, protein samples were denatured with 6x loading buffer (Tris-HCl (0.35 M)/SDS (0.28%) buffer pH 6.8, with 30% Glycerol, 10% SDS, 0.012% Bromophenol Blue, and with/without 0.6 M DTT), 9 min at 90 ºC, and then applied to 12.5% or 10% polyacrylamide gels. The gel electrophoresis was performed in a Mini-PROTEAN Tetra Cell system (Bio-Rad) with running buffer (100 mM Tris, 100 mM Bicine and 0.1% SDS), at 120 V and at room temperature. The gels were stained with Coomassie Brilliant Blue (50% methanol, 10% acetic acid and 0.2% Brilliant Blue R) followed by destaining with a solution of 25% methanol and 5% acetic acid. Images were acquired in a VWR imager (VWR international, Pensylvania, EUA). The apparent molecular weight of proteins was established by comparison with Precision Plus Protein™ Unstained Protein Standards (Bio-Rad).

2.14.2 Western blot analysis

For the detection and identification of proteins, the samples were analysed by Western blot. After separation by SDS-PAGE, the protein samples were electrotransferred onto Amersham Hybond 0.2 µm PVDF membrane (Cytiva), using the buffer 25 mM Tris, 192 mM glycine and 20% methanol, for 1 hour and 40 minutes at 100 V, at 4 ºC, using a Trans-Blot Electrophoretic Transfer cell (Bio-Rad). The membrane was blocked by incubation with blocking buffer TBS-T 2% BSA or TBS-T 5% milk (20 mM Tris and 137 mM NaCl, containing 0.1% of Tween 20 and 2% of BSA or 5% milk, pH 7.6), depending on the antibody, for 1 hour with agitation. The membrane was then incubated with the primary antibody diluted in blocking buffer for 1 hour, with agitation, followed by 5 washes (5 minutes each) with TBS-T buffer (20 mM Tris and 137 mM NaCl containing 0.1% of Tween 20). The following primary antibodies were used accordingly: HRP-conjugated streptavidin (#3999S, Cell Signaling Technology) (1:4 000 in TBS-T 5% milk); polyclonal rabbit antibody anti-APRc (raised towards the sequence Cys-Tyr-Thr-Arg-Thr-Tyr-Leu-Thr-Ala-Asn-Gly-Glu-Asn-Lys-Ala), produced by GenScript (1:500 in TBS-T 2% BSA); polyclonal quail antibody anti-APRc (raised against the soluble domain of APRc (recombinant APRc110-231) (1:2 000 in TBS-T 2% BSA); monoclonal mouse anti-Human Ig Kappa Ligth Chain [MAB10050, R&D Systems, Bio-techne (Minnesota, U.S.A)] (1:259 in TBS-T 2% BSA).

Afterwards, the membrane was incubated with the secondary antibody diluted in blocking buffer for 1 hour with agitation, also followed by 5 washes (5 minutes each) with TBS-T buffer. The following secondary antibodies were used: polyclonal rabbit anti-Mouse IgG (whole molecule)-peroxidase (A9044, Sigma-Aldrich) (1:5 000/1:10 000 in TBS-T 2% BSA); polyclonal rabbit anti-Chicken IgY (IgG) (whole molecule)-peroxidase (A9046, Sigma-Aldrich) (1:2 000 in TBS-T 2% BSA); monoclonal mouse anti-Human IgG Fc[HRP] (A01854-200, GenScript) (1:20 000 in TBS-T 2% BSA); monoclonal mouse anti-rabbit IgG (M205) [HRP] (A01827-200, GenScript) (1:20 000 in TBS-T 2% BSA).

The chemiluminescent detection was performed by incubation of the membrane with the NZY advanced ECL substrate (NZYTech) for 4 minutes, and using a VWR imager for image acquisition. The apparent molecular weight of proteins was established by comparison with Precision Plus ProteinTM All Blue Prestained Protein Standards (Bio-Rad).

2.14.3 Agarose gel electrophoresis

The DNA fragments were separated according to their size by agarose gel electrophoresis. The DNA samples diluted in 10x loading buffer (50 mM Ficoll, 10 mM EDTA, 0.1% SDS, 0.36 mM Bromophenol Blue, and 0.46 mM xylene cyanol) were loaded onto an agarose gel (1%) containing ethidium bromide. The electrophoresis was performed in a Mini-Sub Cell GT Cell (Bio-Rad) in TAE buffer (40 mM Tris, 1 mM EDTA and 0.11% acetic acid), at 90 V for 25 minutes, at room temperature. The Gel-Doc system (Bio-rad) was used to visualize the gel and the size of the DNA fragments was determined by comparison with NZYDNA ladder III (NZYTech).

3 Results

3.1 Production of the recombinant retropepsin-like protease from *Rickettsia conorii* (APRc)

A shorter version of the mature form of APRc, $APRc_{110-231}$, fused with a C-terminal His-tag, and the respective active site mutant, were used to characterize the non-immune Ig-APRc binding activity. Previous experiments in the laboratory with the protein $APRc_{110-231}$-His (includes APRc amino acids 110-231 fused with the C-terminal tag sequence AAALEHHHHHH) revealed that the linker sequence to the C-terminal His-tag (AAALE) had an impact on the non-immune Ig-APRc binding activity. Therefore, for a more accurate characterization of the Ig-APRc binding activity, the new construct $APRc_{110-231}$-HisShort that includes APRc amino acids 110-231 directly fused with the C-terminal tag sequence HHHHHH and the respective active site mutant [$APRc_{110-231}$-HisShort(D140N)] were used throughout this work. The amino acid sequences of both constructs are represented in annexes (Figure 6.1).

$APRc_{110-231}$-HisShort and the respective active site mutant [$APRc_{110-231}$-HisShort(D140N)] were recombinantly expressed in *E. coli* (a scale-up expression was used in order to collect enough protein for all the experiments), purified by IMAC-Ni^{2+}, and further purified by cation-exchange chromatography. Since APRc was described as an active enzyme with properties similar to those of retropepsins, such as accumulation in the dimeric form that is essential for proteolytic activity[64,67], the oligomerization state and activity of these recombinant forms of APRc were also evaluated. In order to confirm that the removal of the linker would not impact $APRc_{110-231}$-HisShort properties, $APRc_{110-231}$-His and the respective active site mutant [$APRc_{110-231}$-His(D140N)] were also recombinantly produced, purified, and had their activity and oligomerization state evaluated for comparison.

3.1.1 Purification of recombinant APRc

The soluble domain pET_$APRc_{110-231}$-HisShort, pET_$APRc_{110-231}$-HisShort(D140N), pET_$APRc_{110-231}$-His and pET_$APRc_{110-231}$-His(D140N) were independently transformed in BL21 Star cells. Protein expression was induced at an OD_{600nm} of 0.6 for 3 hours at 30ºC, and the cells were harvested by centrifugation and frozen. In order to separate the recombinant proteins from the other components of the cell extracts, the soluble domains of the different cell extracts were purified by immobilized metal affinity chromatography with Ni^{2+}. The different forms of APRc have a C-terminal His-tag that has affinity for nickel ions, enabling the purification of the recombinant proteins using an IMAC column with coupled nickel (Histrap HP 5 mL column). The total soluble fractions were applied onto a Histrap HP 5 mL column, previously equilibrated in 20 mM sodium phosphate pH 7.4 buffer,

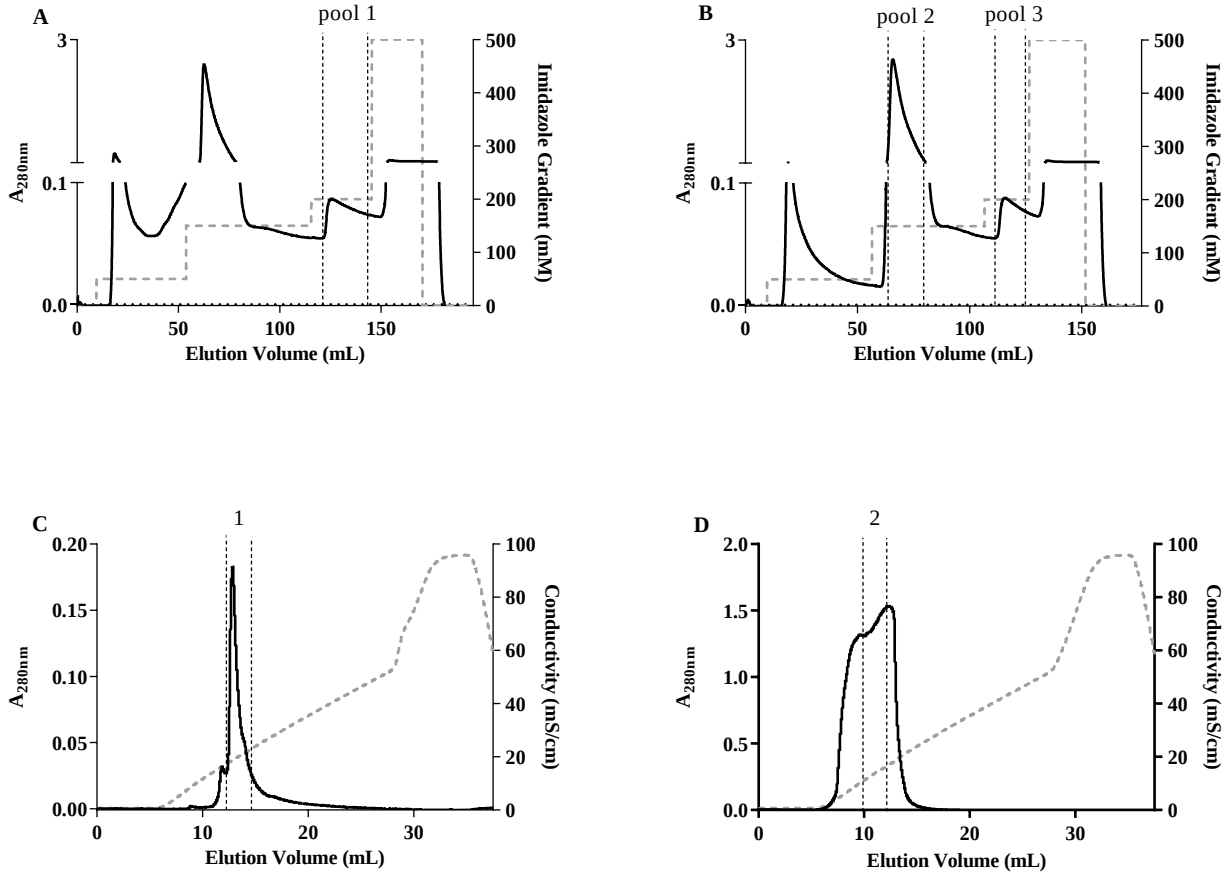
containing 10 mM imidazole and 500 mM NaCl. Protein elution was carried out by implementing a four-step gradient of imidazole (50 mM, 150 mM, 200 mM, and 500 mM) in 20 mM sodium phosphate pH 7.4 buffer containing 500 mM NaCl. Imidazole act as a metal ion ligand competing with the Ni^{2+}/Histidine complexes. The buffer with this compound was used to prevent protein non-specific binding, and, with increasing concentrations, for protein elution. Representative chromatograms of IMAC-Ni^{2+} purifications of $APRc_{110-231}$-HisShort, $APRc_{110-231}$-HisShort(D140N), $APRc_{110-231}$-His, and $APRc_{110-231}$-His(D140N) are shown in Figure 3.1 (A) and (B), Figure 3.2 (A) and (B), Figure 3.3 (A) and (B) and Figure 3.4 (A) and (B), respectively. Each chromatogram shows the monitoring of protein elution at an absorbance of 280 nm and the respective gradient of imidazole.

From prior purifications of $APRc_{110-231}$-His, it was observed that the fractions eluted with 200 mM imidazole corresponded to the dimeric forms of the protein (and with enzymatic activity, in the case of the wild-type). The proteins eluted at 150 mM imidazole were monomers and did not display activity. As for an accurate characterization of the non-immune Ig-APRc binding activity, dimeric and active forms of APRc were required, the eluted fractions at the 200 mM imidazole gradient step were selected for further purification. As a negative control, the monomeric recombinant form of $APRc_{110-231}$-HisShort was also purified from one of the IMAC Ni^{2+} purifications.

Given the low yields of the dimeric forms, two separately IMAC Ni^{2+} columns were performed per purification round. The proteins eluted at the 200 mM imidazole gradient steps were independently pooled and dialyzed overnight against 20 mM HEPES pH 7.4 buffer. After dialysis, the two pools of recombinant APRc were applied onto a cation-exchange column (Mono-S 5/50 GL column) previously equilibrated with 20 mM HEPES pH 7.4 buffer. These recombinant forms of APRc are positively charged at that pH, enabling the binding to the negatively charged column. This enables the separation of very similar types of molecules by separating ionizable molecules based on their total charge. Protein elution was carried out by implementing a linear gradient of NaCl (0-1 M) in 20 mM HEPES pH 7.4 buffer. With increasing concentrations of NaCl, the competition between the charged salt ions and the bound proteins for the charged resin functional groups also increases, enabling protein elution. In the case of the recombinant protein $APRc_{110-231}$-HisShort, the eluted protein at 150 mM imidazole was also subjected to the same purification steps. Representative chromatograms of the cation-exchange purifications of $APRc_{110-231}$-HisShort, $APRc_{110-231}$-HisShort(D140N), $APRc_{110-231}$-His, and $APRc_{110-231}$-His(D140N) are shown in Figure 3.1 (C), Figure 3.2 (C), Figure 3.3 (C), and Figure 3.4 (C), respectively. The representative chromatogram of the cation-exchange purification of the eluted fractions from the 150 mM imidazole gradient step of $APRc_{110-231}$-HisShort is shown in Figure 3.1 (D). Each chromatogram shows the monitoring of protein elution at an absorbance of 280 nm and the respective NaCl concentration given by buffer conductivity.

The selected eluted fractions resultant from the cation-exchange purification of each recombinant form of APRc are shown in each chromatogram. Samples from the different purification steps were analysed by SDS-PAGE followed by Coomassie blue staining, as shown in Figure 3.1 (E) ($APRc_{110-231}$-HisShort), Figure 3.2 (D) ($APRc_{110-231}$-HisShort(D140N)), Figure 3.3 (D) ($APRc_{110-231}$-His) and Figure 3.4 (D) ($APRc_{110-231}$-His(D140N)).

The recombinant forms of $APRc_{110-231}$-HisShort and $APRc_{110-231}$-His present similar profiles, suggesting that the removal of the linker sequence to the C-terminal His-tag (AAALE) did not impacted $APRc_{110-231}$-HisShort properties. Both show elution of the majority of the protein at 150 mM imidazole gradient step (this is confirmed in the SDS-PAGE analysis of the different selected fractions). And, as expected, under denaturing conditions, all purified recombinant forms of APRc present a molecular weight of approximately 14 kDa.



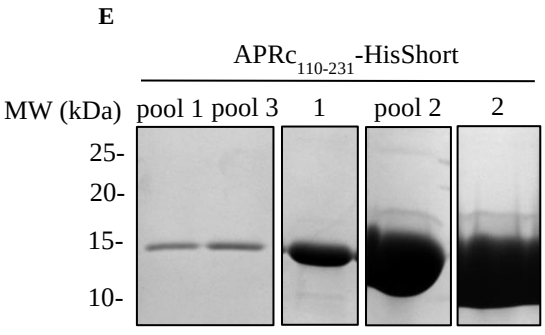


Figure 3.1 (previous page)- Purification of $APRc_{110-231}$-HisShort by IMAC Ni^{2+} and cation-exchange chromatography. (A) (B) The recombinant protein $APRc_{110-231}$-HisShort was purified by IMAC Ni^{2+} on a HisTrap HP 5 mL column, previously equilibrated in 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl. Protein elution was carried out by implementation of a four-step gradient of imidazole (50 mM, 150 mM, 200 mM, and 500 mM) in 20 mM sodium phosphate pH 7.4 buffer containing 500 mM NaCl (represented by the grey dashed line) at a flow rate of 5 mL/min and monitored at an absorbance of 280 nm. The eluted proteins of the 200 mM imidazole gradient step were independently pooled [pool 1 (A); pool 3 (B)], as well as the protein fractions eluted at 150 mM imidazole [pool 2 (B)], and dialyzed overnight against 20 mM HEPES pH 7.4 buffer. **(C)** After dialysis, pools 1 and 3 were purified together in a Mono-S 5/50 GL column, previously equilibrated with 20 mM HEPES pH 7.4 buffer. Protein elution was carried out by implementation of a linear gradient of NaCl (0-1 M) in 20 mM HEPES pH 7.4 buffer, at a flow rate of 0.75 mL/min and monitored at an absorbance of 280 nm. The conductivity is represented by the grey dashed line. The selected eluted fraction (1) is outlined by dashed lines. **(D)** After dialysis, pool 2 was purified individually in a Mono-S 5/50 GL column, under the same conditions described in (C). The selected eluted fraction (2) is outlined by dashed lines. **(E)** SDS-PAGE analysis followed by Coomassie Blue staining of the pooled eluted fractions (pool 1, 2, and 3) resultant from the purifications by IMAC Ni^{2+}, and the selected eluted fractions (1 and 2) from the purification by cation-exchange chromatography. 20 µL of each sample were denatured with 6x loading buffer with DTT and then applied to a 12.5% polyacrylamide gel.

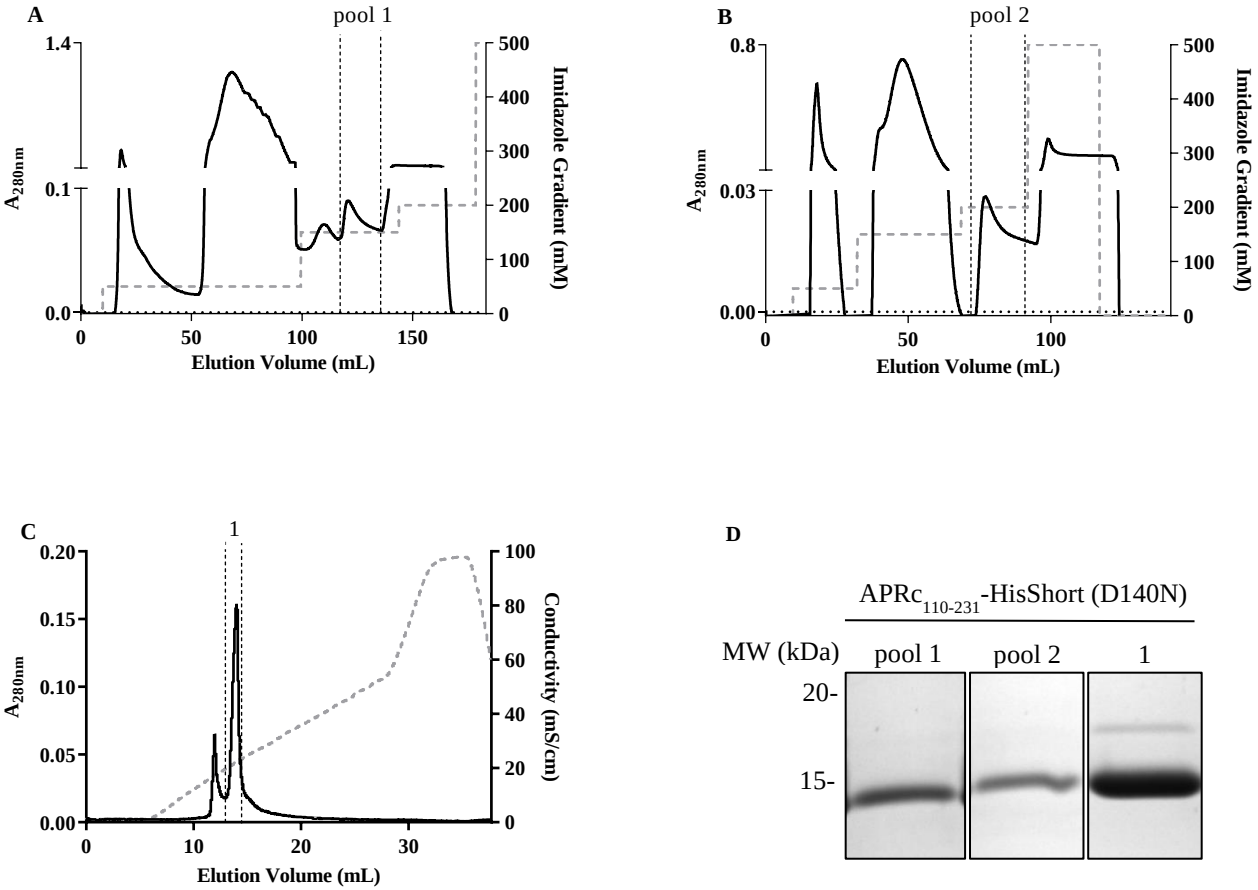


Figure 3.2 (previous page)- Purification of $APRc_{110-231}$-HisShort(D140N) by IMAC Ni^{2+} and cation-exchange chromatography. (A) (B) The recombinant protein $APRc_{110-231}$-HisShort (D140N) was purified by IMAC Ni^{2+} on a HisTrap HP 5 mL column, previously equilibrated in 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl. Protein elution was carried out by implementation of a four-step gradient of imidazole (50 mM, 150 mM, 200 mM, and 500 mM) in 20 mM sodium phosphate pH 7.4 buffer containing 500 mM NaCl (represented by the grey dashed line) at a flow rate of 5 mL/min and monitored at an absorbance of 280 nm. The eluted fractions of the 200 mM imidazole gradient step were independently pooled [pool 1 (A); pool 2 (B)] and dialyzed overnight against 20 mM HEPES pH 7.4 buffer. **(C)** After dialysis, pool 1 and pool 2 were purified together in a Mono-S 5/50 GL column, previously equilibrated with 20 mM HEPES pH 7.4 buffer. Protein elution was carried out by implementation of a linear gradient of NaCl (0-1 M) in 20 mM HEPES pH 7.4 buffer, at a flow rate of 0.75 mL/min and monitored at an absorbance of 280 nm The conductivity (represented by the grey dashed line) was also monitored. The selected eluted fraction (1) is outlined by dashed lines. **(D)** SDS-PAGE analysis followed by Coomassie blue staining of the pooled eluted fractions (pools 1, 2) resultant from the purifications by IMAC Ni^{2+}, and the selected eluted fraction 1 from the purification by cation-exchange chromatography. 20 µL of each sample were denatured with 6x loading buffer with DTT and then applied to a 12.5% polyacrylamide gel.

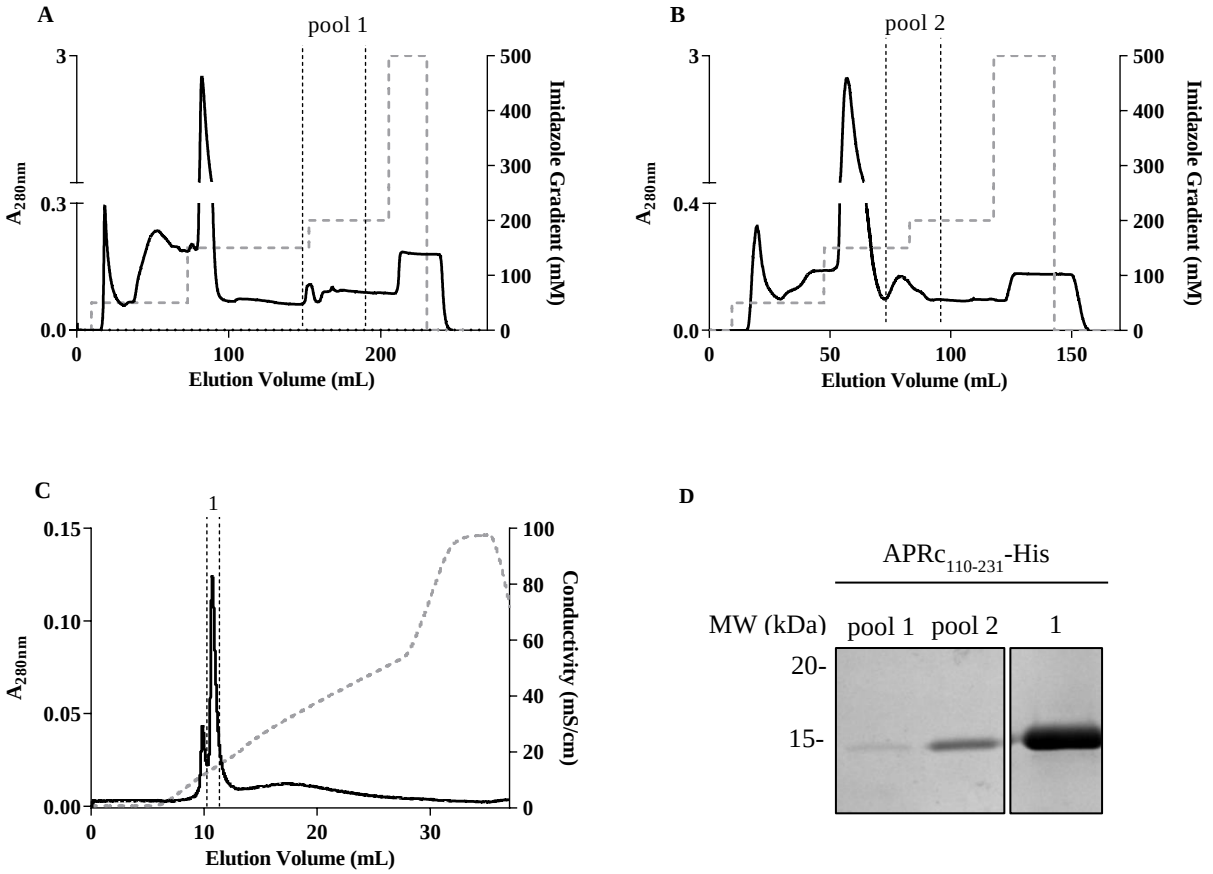


Figure 3.3 (previous page)- Purification of $APRc_{110-231}$-His by IMAC Ni^{2+} and cation-exchange chromatography. (A) (B) The recombinant protein $APRc_{110-231}$-His was purified by IMAC Ni^{2+} on a HisTrap HP 5 mL column, previously equilibrated in 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl. Protein elution was carried out by implementation of a four-step gradient of imidazole (50 mM, 150 mM, 200 mM, and 500 mM) in 20 mM sodium phosphate pH 7.4 buffer containing 500 mM NaCl (represented by the grey dashed line) at a flow rate of 5 mL/min and monitored at an absorbance of 280 nm. The eluted fractions of the 200 mM imidazole gradient step were pooled [pool 1 (A), pool 2 (B)] and dialyzed overnight against 20 mM HEPES pH 7.4 buffer. **(C)** After dialysis, pool 1 and pool 2 were purified together in a Mono-S 5/50 GL column, previously equilibrated with 20 mM HEPES pH 7.4 buffer. Protein elution was carried out by implementation of a linear gradient of NaCl (0-1 M) in 20 mM HEPES pH 7.4 buffer, at a flow rate of 0.75 mL/min and monitored at an absorbance of 280 nm. The conductivity is represented by the grey dashed line. The selected eluted fraction (1) is outlined by dashed lines. **(D)** SDS-PAGE analysis followed by Coomassie Blue staining of the pooled eluted fractions (pool 1, 2) resultant from the purifications by IMAC Ni^{2+} and the selected eluted fraction 1 from the purification by cation-exchange chromatography. 20 µL of each sample were denatured with 6x loading buffer with DTT and then applied to a 12.5% polyacrylamide gel.

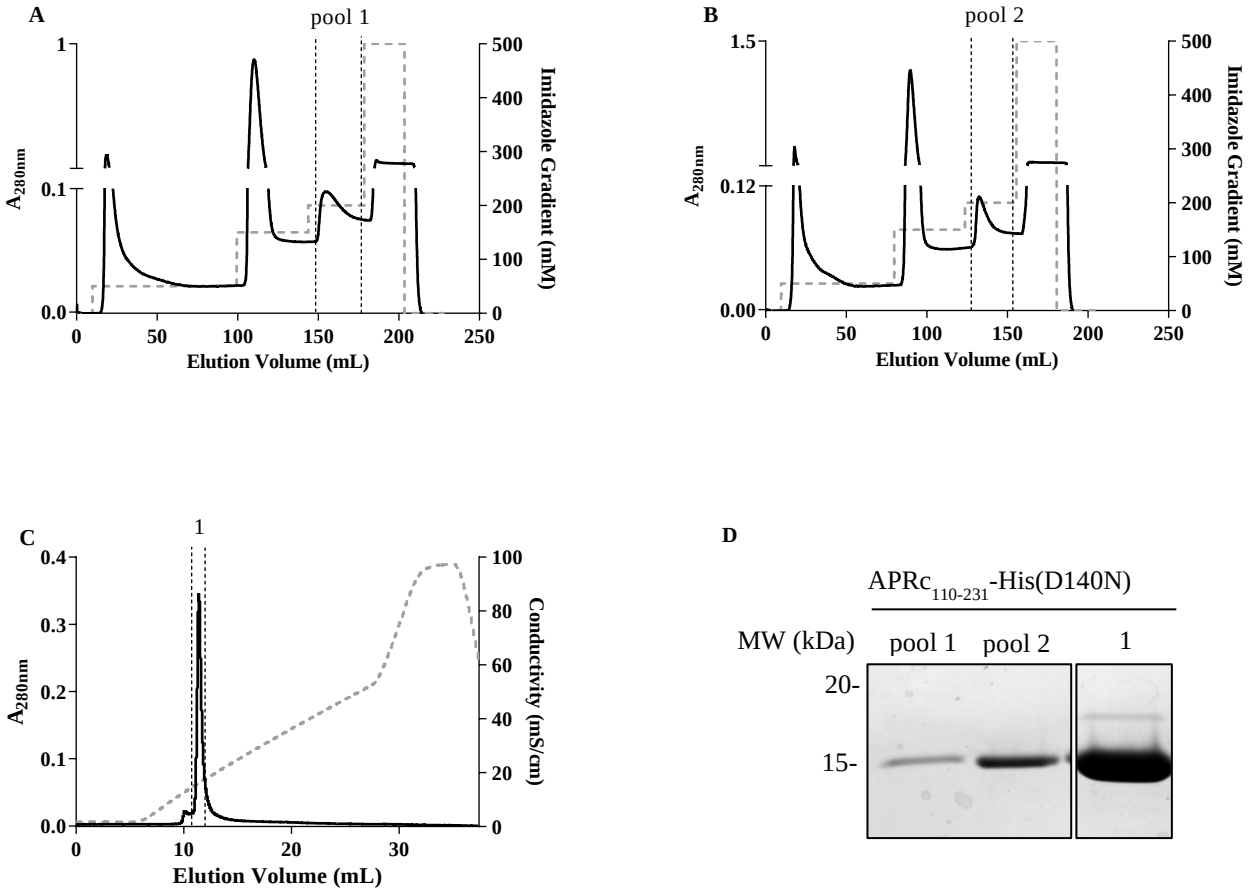


Figure 3.4 (previous page)- Purification of $APRc_{110-231}$-His(D140N) by IMAC Ni^{2+} and cation-exchange chromatography. (A) (B) The recombinant protein $APRc_{110-231}$-His(D140N) was purified by IMAC Ni^{2+} on a HisTrap HP 5 mL column, previously equilibrated in 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl. Protein elution was carried out by implementation of a four-step gradient of imidazole (50 mM, 150 mM, 200 mM, and 500 mM) in 20 mM sodium phosphate pH 7.4 buffer containing 500 mM NaCl (represented by the grey dashed line) at a flow rate of 5 mL/min and monitored at an absorbance of 280 nm. The eluted fractions of the 200 mM imidazole gradient step were pooled [pool 1 (A), pool 2 (B)] and dialyzed overnight against 20 mM HEPES pH 7.4 buffer. **(C)** After dialysis, pool 1 and pool 2 were purified together in a Mono-S 5/50 GL column, previously equilibrated with 20 mM HEPES pH 7.4 buffer. Protein elution was carried out by implementation of a linear gradient of NaCl (0-1 M) in 20 mM HEPES pH 7.4 buffer, at a flow rate of 0.75 mL/min and monitored at an absorbance of 280 nm. The conductivity is represented by the grey dashed line. The selected eluted fraction (1) is outlined by dashed lines. **(D)** SDS-PAGE analysis followed by Coomassie Blue staining of the pooled eluted fractions (pool 1, 2) resultant from the purifications by IMAC Ni $^{2+}$ and the selected eluted fraction 1 from the purification by cation-exchange chromatography. 20 µL of each sample were denatured with 6x loading buffer with DTT and then applied to a 12.5% polyacrylamide gel.

3.1.2 Evaluation of the oligomerization state of recombinant APRc

In retropepsins, the dimerization of the enzyme is essential for proteolytic activity[64]. Therefore, for accurate characterization of the non-immune Ig-APRc binding activity, dimeric recombinant forms of APRc (with activity) are required. To this end, the oligomerization state of the all purified fractions from cation-exchange chromatography was evaluated. The purified proteins were evaluated by analytical size-exclusion chromatography in a Superdex 200 5/150 GL column previously equilibrated in 20 mM HEPES buffer pH 7.4 containing 100 mM NaCl. In size-exclusion chromatography, the molecules are separated by molecular size in a packed column with porous beads. The molecular weight of proteins was estimated by calibration of the column with standard molecular weight proteins. The representative chromatogram of the analytical size-exclusion chromatography of the purified recombinant proteins $APRc_{110-231}$-HisShort (eluted fractions from 150 mM and 200 mM imidazole gradient steps of IMAC Ni^{2+}), $APRc_{110-231}$-HisShort(D140N), $APRc_{110-231}$-His, and $APRc_{110-231}$-His(D140N) is shown in Figure 3.5. The chromatogram shows the monitoring of protein elution at an absorbance of 220 nm and the respective elution volume. The elution volumes for the standard molecular weight proteins are also shown.

The recombinant proteins $APRc_{110-231}$-HisShort (eluted fractions from 200 mM imidazole gradient steps of IMAC Ni^{2+}), $APRc_{110-231}$-HisShort(D140N), $APRc_{110-231}$-His, and $APRc_{110-231}$-His(D140N) displayed an estimated molecular weight of approximately 29 kDa. This confirms that these recombinant forms of APRc are dimeric. The recombinant protein $APRc_{110-231}$-HisShort eluted at 150 mM imidazole in the IMAC displayed an estimated molecular weight of approximately 14 kDa, indicating that this recombinant form of APRc is monomeric.

Even though these recombinant forms of APRc were shown to be mostly dimers with a molecular weight of approximately 29 kDa, in the SDS-PAGE analysis [Figure 3.1 (E), Figure 3.2 (D), Figure 3.3 (D) and Figure 3.4 (D)] they run with a molecular size of approximately 14 kDa. APRc has a characteristic weak homodimerization, only present under certain conditions, which makes it difficult to visualize APRc dimers by SDS-PAGE analysis.

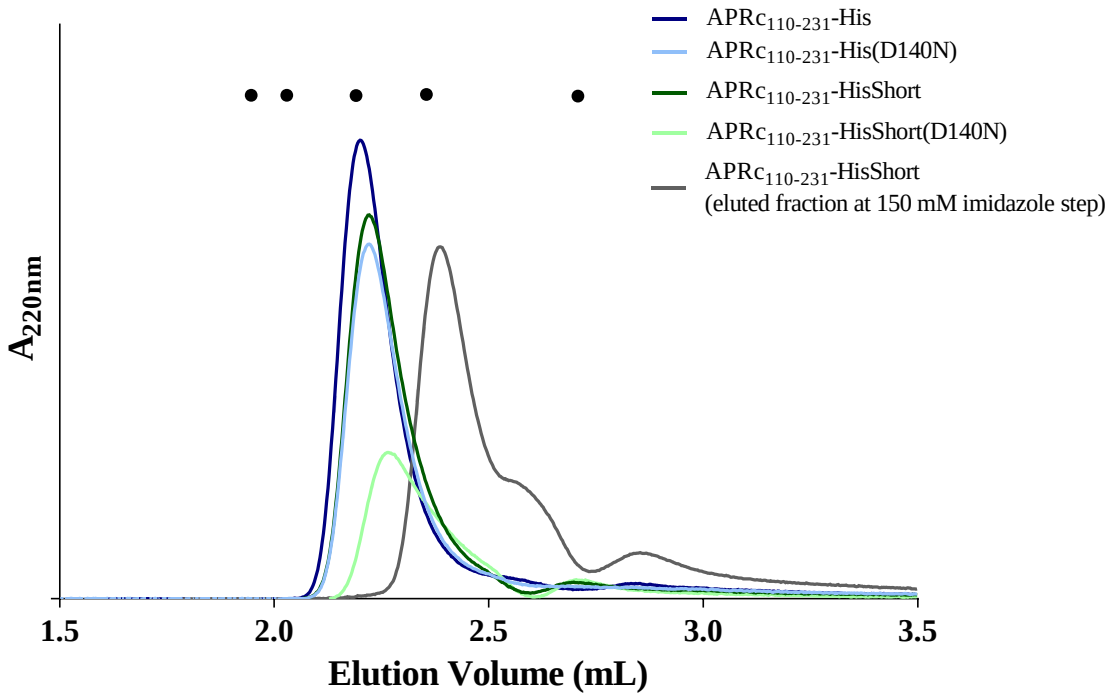


Figure 3.5- Evaluation of the oligomerization state of APRc. The purified samples of the recombinant proteins APRc$_{110-231}$-HisShort (eluted fractions from 150 mM and 200 mM imidazole gradient steps of IMAC Ni^{2+}), APRc$_{110-231}$-HisShort(D140N), APRc$_{110-231}$-His, and APRc$_{110-231}$-His(D140N) were analysed by analytical size-exclusion chromatography in a Superdex 200 5/150 GL column previously equilibrated in 20 mM HEPES buffer pH 7.4 containing 100 mM NaCl. Protein elution was carried out at a flow rate of 0.25 mL/min and monitored at an absorbance of 220 nm. The black dots represent the elution volumes for the standard molecular weight proteins, from left to right: conalbumin (75 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa), ribonuclease A (13.7 kDa), and aprotinin (6.5 kDa).

3.1.3 Evaluation of the activity of APRc

In Cruz, R. *et al* (2014)[67], APRc was described as an active enzyme with properties similar to those of retropepsins and demonstrated to have activity towards oxidized insulin β-chain, a polypeptide usually cleaved by aspartic proteases. Also, the mutation of the active site is described to impair its proteolytic activity[67]. To further characterize the different recombinant forms of APRc herein purified and enable their use in the following studies, the enzymatic activity of all purified fractions was

evaluated. APRc samples were incubated with the oxidized insulin β-chain overnight at 37ºC in 0.1 M sodium acetate buffer, pH 6 (APRc displays optimal activity *in vitro* at pH 6[67]). Their activity towards this substrate was then evaluated by reverse-phase HPLC in KROMASIL 100 C18 250 4.6 mm column, previously equilibrated in 0.1% TFA. As a negative control, oxidized insulin β-chain was incubated under similar conditions. The elution of peptides was carried out with a linear gradient of acetonitrile (0-80%) in 0.1% TFA. Representative chromatograms are shown in Figure 3.6, and represent the monitoring of the elution of peptides at an absorbance of 220 nm. In Figure 3.6 (A), the activity of the dimeric form of $APRc_{110-231}$-HisShort (Figure 3.5) towards the substrate is compared to the activity of its monomeric form, and to the elution profile of oxidized insulin β-chain. In Figure 3.6 (B), the activity of the dimeric form of $APRc_{110-231}$-HisShort is compared to the activity of its active site mutant ($APRc_{110-231}$-HisShort(D140)). In Figure 3.6 (C), the activity of dimeric $APRc_{110-231}$-His (Figure 3.5) is compared to the elution profile of oxidized insulin β-chain. Finally, in Figure 3.6 (D) the activity of the dimer $APRc_{110-231}$-His towards this substrate is compared to that of its corresponding active site mutant ($APRc_{110-231}$-His(D140)).

The chromatograms show that the wild-type recombinant $APRc_{110-231}$-HisShort (dimer) displayed activity towards the oxidized insulin β-chain substrate, as peaks corresponding to peptides originated from the digestion of the substrate are detected. For the monomeric form of $APRc_{110-231}$-HisShort, no activity towards this substrate was noticed. The wild-type recombinant $APRc_{110-231}$-His (dimer) also showed activity towards the oxidized insulin β-chain. As expected, no digestion peaks were observed upon incubation with the dimeric active site mutants of APRc [$APRc_{110-231}$-HisShort(D140N) and $APRc_{110-231}$-His(D140N)].

These results confirm that the dimeric recombinant wild-type forms of APRc are active, whereas the monomeric form is not. As described in retropepsins, the dimerization of the enzyme is essential for proteolytic activity. These results further suggest that not even the presence of the substrate helps to favour dimer formation. Moreover, these data also confirm that removing the linker sequence in the construct $APRc_{110-231}$-HisShort did not impact its proteolytic activity, corroborating its use (and of the corresponding active site mutant) in the follow-up APRc-Ig binding studies.

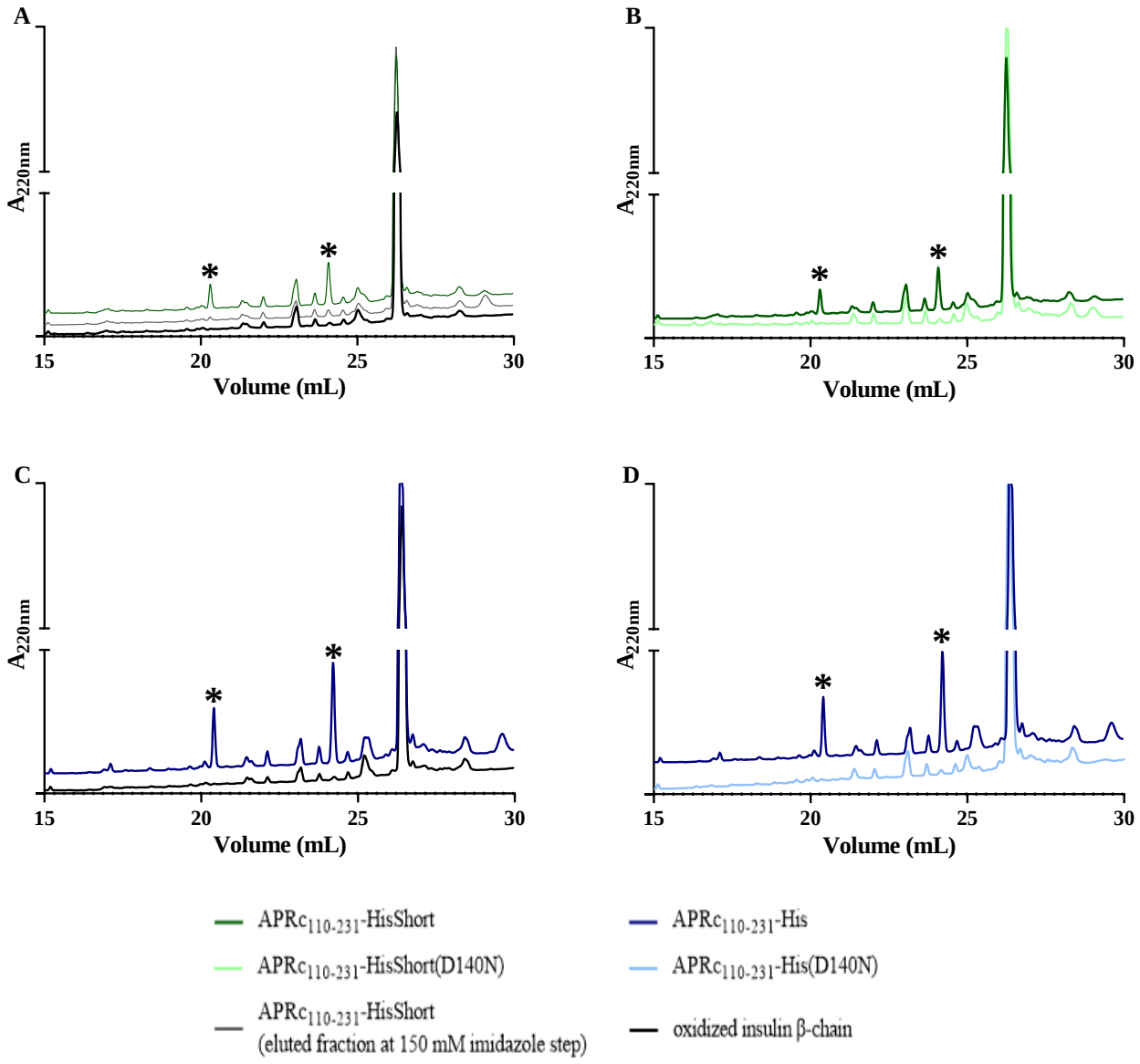


Figure 3.6- Evaluation of the proteolytic activity of APRc. (A) Samples of the eluted fractions of recombinant APRc$_{110-231}$-HisShort (eluted fractions from 200 mM imidazole gradient step, dimer) and of recombinant APRc$_{110-231}$-HisShort (eluted fractions from 150 mM imidazole gradient step, monomer) were incubated with oxidized insulin β-chain overnight at 37ºC, in 0.1 M sodium acetate buffer, pH 6, and then their activity was evaluated by Reverse-Phase HPLC in a KROMASIL 100 C18 250 4.6 mm column, previously equilibrated in 0.1% TFA. As a control, oxidized insulin β-chain was incubated under the same conditions and also analysed by Reverse-Phase HPLC. The elution of peptides was carried out with a linear gradient of acetonitrile (0-80%) in 0.1% TFA, at a flow rate of 1 mL/min, and monitored at an absorbance of 220 nm**.** The peaks corresponding to the peptides originated from the digestion of oxidized insulin β-chain are marked with an asterisk. **(B)** Samples of the eluted fractions of recombinant APRc$_{110-231}$-HisShort (dimer) and of recombinant APRc$_{110-231}$-HisShort(D140) (dimer) were incubated with oxidized insulin β-chain overnight at 37ºC, in 0.1 M sodium acetate buffer, pH 6, and then activity evaluated by Reverse-Phase HPLC, as described in (A). Digestion products are marked with an asterisk. **(C)** A sample of the eluted fraction of recombinant APRc$_{110-231}$-His (dimer) was incubated with oxidized insulin β-chain overnight at 37ºC in 0.1 M sodium acetate buffer, pH 6, and their activity evaluated by Reverse-Phase HPLC, as in (A). As a control, oxidized insulin β-chain was incubated under the same conditions. Digestion products are marked with an asterisk. **(D)** Samples of the eluted fractions of recombinant APRc$_{110-231}$-His (dimer) and of recombinant APRc$_{110-231}$-His(D140) (dimer) were incubated with oxidized insulin β-chain substrate overnight at 37ºC in 0.1 M sodium acetate buffer, pH 6 and their activity evaluated by Reverse-Phase HPLC, as in (A). Digestion products are marked with an asterisk.

3.2 Characterization of the non-immune Ig-APRc binding activity

Many bacteria present proteins that, without the implication of the antigen-binding sites, are able to bind to immunoglobulins (non-immune IgG-binding proteins) with the purpose of scavenging opsonizing antibodies, impairing the recognition by the opsonic receptors of the host organism and subsequently evading phagocytosis, or impairing the deposition of complement components[23,24].

Rickettsia presents a retropepsin-like protease (APRc) that is predicted to participate in the spreading of infection and dissemination of bacteria into deeper tissue by shedding adhesion molecules at the cell surface or by disabling components of the host immune system[67]. Previous results in the laboratory identified APRc-IgG binding activity, anticipating this protease as a potentially novel non-immune Ig-binding protein that may act as a novel evasin, thereby contributing to the rickettsial immune evasion toolbox.

Here we characterized this non-immune Ig-APRc binding activity, starting by evaluating Ig-APRc binding and the potential APRc-Ig-cleavage activity. Indeed, since APRc is an aspartic protease with proteolytic activity[67], there is the possibility that besides binding to immunoglobulins, it is also able to cleave them. This is followed by assessing the interaction of APRc with different immunoglobulins (immunoglobulins from different classes and species) and mapping of the regions in APRc and Ig responsible for binding. The non-immune Ig-APRc binding activity was also evaluated in a more complex context (normal human serum). Since *Rickettsia* usually establishes infection in humans, being in contact with serum, the interaction of APRc with immunoglobulins in the context of human serum was also evaluated. Moreover, since APRc is described to accumulate in the outer membrane and to have a soluble catalytic domain extracellular oriented[67], the non-immune Ig-binding at the surface of *Rickettsia* species was also assessed.

3.2.1 Evaluation of APRc's Ig-binding and Ig-cleavage activity

Since APRc is an aspartic protease with proteolytic activity[67], there is the possibility that besides binding to immunoglobulins, it is also able to cleave them. To evaluate if APRc combines Ig-binding with Ig-cleavage activity, several incubations of recombinant APRc with immunoglobulins were performed under different conditions. Here are presented some of those assays that are representative of the features observed under the different tested conditions.

Recombinant APRc$_{110-231}$-HisShort (dimer/active form) was incubated with polyclonal rabbit IgG specific for a plant-protease epitope (produced by Genscript) (with unrelated specificity) in an equimolar ratio (529.35 ρmole) in PBS buffer pH 7.4, for 2 hours at 37ºC. A control with only rabbit IgG or only the recombinant APRc$_{110-231}$-HisShort, incubated under the same conditions, was performed. The samples were denatured with 6x loading buffer without DTT and then analysed by Western-blot with quail anti-APRc antibody followed by rabbit anti-Chicken IgY (IgG) [HRP] antibody (Figure 3.7).

The Western-blot analysis shows differences in APRc migration to higher molecular weights when incubated with rabbit IgG. Although some unspecific binding of the secondary antibody is detected, the increased signal when both APRc and IgG are present suggests an interaction of APRc with rabbit IgG, which appears to interfere with the oligomeric state of APRc.

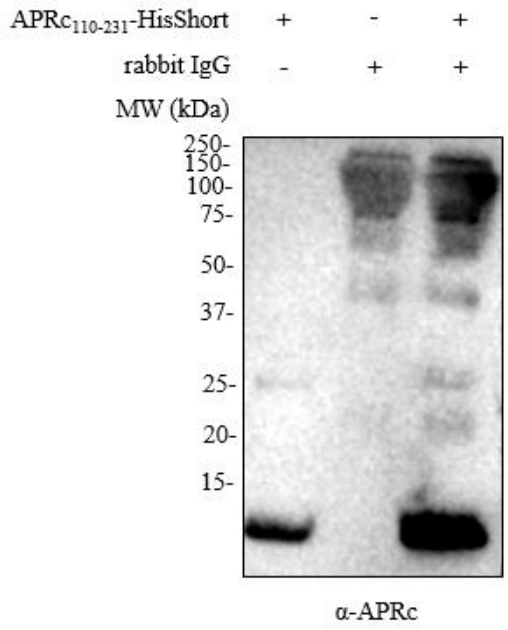


Figure 3.7-Evaluation of rabbit IgG-APRc binding. Recombinant active APRc$_{110-231}$-HisShort was incubated with polyclonal rabbit IgG specific for a plant-protease epitope (produced by Genscript) in an equimolar ratio (529.35 ρmole) in PBS buffer pH 7.4, for 2 hours at 37ºC. As a control, an incubation under the same conditions, only with the polyclonal rabbit IgG or with the recombinant APRc$_{110-231}$-HisShort was performed. 15 µL of each sample were denatured with 6x loading buffer without DTT, applied in a 12.5% polyacrylamide gel, and then evaluated by Western-blot analysis with the polyclonal quail antibody anti-APRc (raised against the soluble domain of APRc (recombinant APRc$_{110-231}$) (1:2 000 in TBS-T 2% BSA) followed by the polyclonal rabbit anti-Chicken IgY (IgG) (whole molecule)-peroxidase (A9046, Sigma-Aldrich) (1:2 000 in TBS-T 2% BSA).

Recombinant APRc$_{110-231}$-HisShort and the respective active site mutant [APRc$_{110-231}$-HisShort(D140N)] (dimeric forms), were then incubated with human IgG (#I2511, Sigma-Aldrich) at different molar ratios [equimolar (333.7 ρmole) /1:0.1/1:0.05 (S:E)], in PBS buffer pH 7.4, for 4h at 37ºC. Controls under the same conditions with only human IgG or with the recombinant APRc$_{110-231}$-

HisShort WT and mutant forms were performed in parallel. The samples were denatured with 6x loading buffer without DTT and then analysed by Western-blot with a rabbit anti-APRc antibody followed by the secondary monoclonal mouse anti-rabbit IgG (M205) [HRP], with secondary mouse anti-rabbit IgG antibody only, and with the monoclonal mouse anti-human IgG Fc [HRP]. In this way, it would be possible to evaluate the contribution of non-specific interactions of the secondary antibody and any changes in the migration pattern of human IgG. In Figure 3.8 (A), (B), and (C) are represented the Western-blot analyses of the incubations of human IgG and APRc at an equimolar ratio, at a 1:0.1 (S:E) molar ratio, and at 1:0.05 (S:E) molar ratio, respectively.

The Western-blot analysis of the incubations at an equimolar ratio of human IgG and APRc [Figure 3.8 (A)], showed differences in APRc migration when incubated with human IgG. It is possible to observe a decrease in the dimeric (~25 kDa) and monomeric (~14 kDa) forms of APRc, and an increase in signal at higher molecular weights for both wild-type and mutant forms. The Western-blot analysis with the secondary antibody only (mouse anti-rabbit IgG antibody), despite presenting unspecific binding to the human IgG, did not show differences in the profile of the different incubations with human IgG, thereby, suggesting that the higher molecular weight gains observed with the anti-APRc antibody are indeed related to APRc. These observations also confirmed that there is an interaction of APRc with human IgG and that this interaction appears to stabilize APRc's oligomeric state in high molecular weight complexes. From the Western-blot with anti-Human IgG Fc antibody, no apparent differences were observed between the incubations of human IgG with the recombinant $APRc_{110-231}$-HisShort (the dimeric and active form of APRc), with recombinant $APRc_{110-231}$-HisShort(D140N) (the dimeric and inactive form of APRc) to the control only with human IgG. Therefore, under these conditions, there is no evidence that APRc entails IgG cleavage**.**

From the Western-blot analyses of the incubations with lower molar ratios of APRc, 1:0.1 (S:E) and 1:0.05 (S:E) [Figure 3.8 (B) and Figure 3.8 (C), respectively], a reduction in the intensity of the signal at high molecular weights was observed in the membranes probed with anti-APRc antibody, which could be related with the fact that a lower amount of APRc was used in these assays. Although it is not possible to rule out unspecific binding of the secondary antibody, one would expect a signal with similar intensity in all incubations where human IgG is present, and that is not the case. Therefore, these results further corroborate APRc-human IgG binding. Again, from the Western-blot with anti-Human IgG, no differences were observed between the incubations of human IgG with the recombinant $APRc_{110-231}$-HisShort WT and mutant forms, to the control only with human IgG.

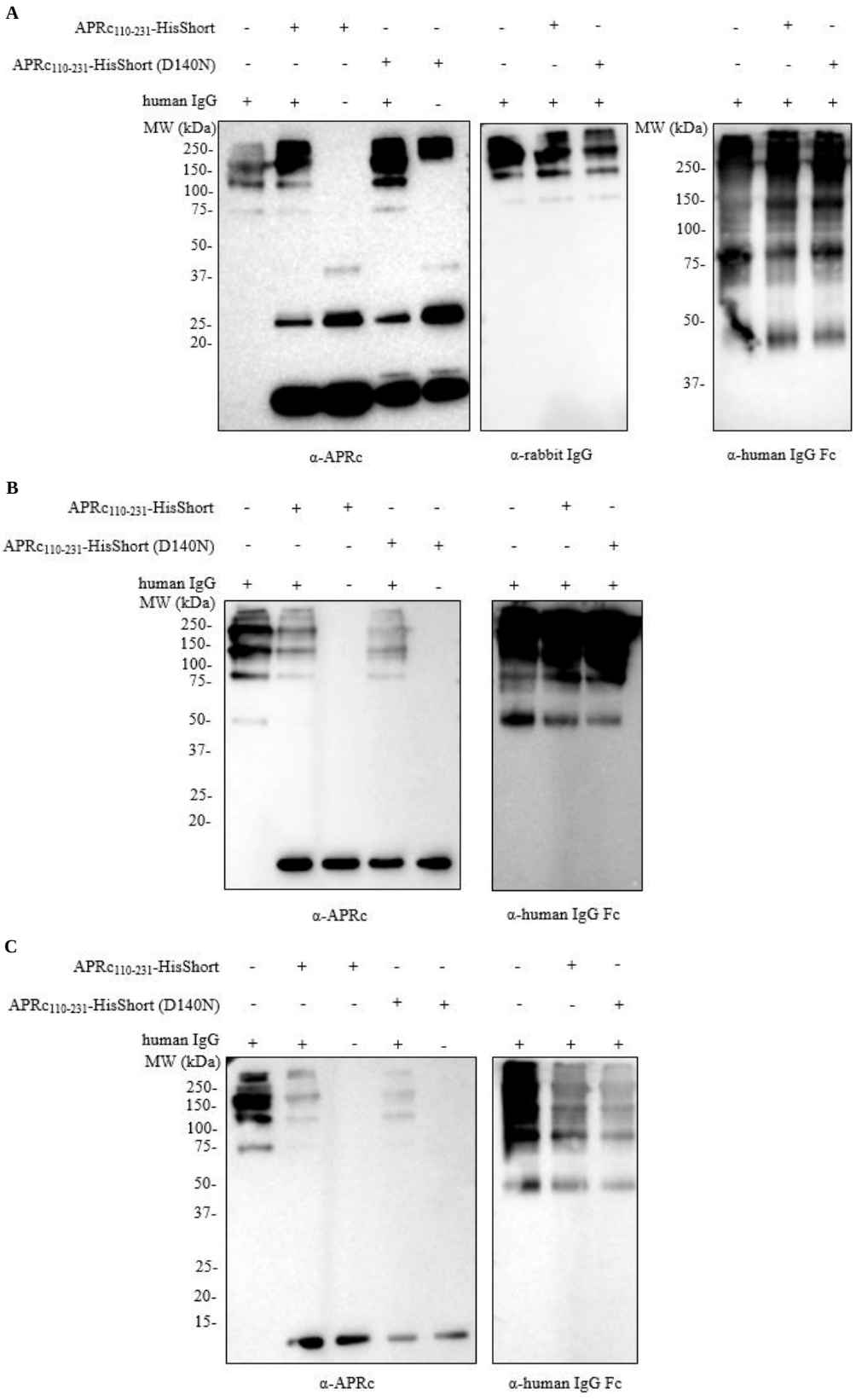

A
APRc110-231-HisShort
APRc110-231-HisShort (D140N)
human IgG
MW (kDa)
250-
150-
100-
75-
50-
37-
25-
20-
α-APRc
α-rabbit IgG
α-human IgG Fc
B
APRc110-231-HisShort
APRc110-231-HisShort (D140N)
human IgG
MW (kDa)
250-
150-
100-
75-
50-
37-
25-
20-
α-APRc
α-human IgG Fc
C
APRc110-231-HisShort
APRc110-231-HisShort (D140N)
human IgG
MW (kDa)
250-
150-
100-
75-
50-
37-
25-
20-
15-
α-APRc
α-human IgG Fc

Figure 3.8 (previous page)- Evaluation of human IgG-APRc binding at different molar ratios. Recombinant $APRc_{110-231}$-HisShort and $APRc_{110-231}$-HisShort(D140N) (dimeric forms), were incubated with human IgG (#I2511, Sigma-Aldrich) at equimolar ratio (333.7 ρmole) **(A)**, at a molar ratio 1:0.1 (S:E) **(B)**, and at a molar ratio 1:0.05 (S:E) **(C)**, in PBS buffer pH 7.4, for 4h at 37ºC. As a control, incubations under the same conditions only with human IgG or only with the recombinant forms of APRc were performed. 3 µL of each sample were denatured with 6x loading buffer without DTT, applied in a 10% polyacrylamide gel, and then evaluated by Western-blot analysis. Membranes were probed with the polyclonal rabbit antibody anti-APRc (raised towards the sequence Cys-Tyr-Thr-Arg-Thr-Tyr-Leu-Thr-Ala-Asn-Gly-Glu-Asn-Lys-Ala), produced by GenScript (1:500 in TBS-T 2% BSA) followed by the monoclonal mouse anti-rabbit IgG (M205) [HRP] (A01827-200, GenScript) (1:20 000 in TBS-T 2% BSA) or with the monoclonal mouse anti-rabbit IgG (M205) [HRP] (A01827-200, GenScript) (1:20 000 in TBS-T 2% BSA) or with the monoclonal mouse anti-Human IgG Fc [HRP] (A01854-200, GenScript) (1:20 000 in TBS-T 2% BSA).

To further evaluate APRc-IgG cleavage, several incubations were performed at different pH values and incubation times, but no apparent cleavage was ever observed (data not shown). As an additional strategy, recombinant $APRc_{110-231}$-HisShort in its dimeric (active) and monomeric (inactive) form were independently incubated with human IgG (#I2511, Sigma-Aldrich) at a 50 µg:6 µg ratio (S:E), in PBS buffer pH 7.4, overnight at 37ºC. Controls for each condition with only human IgG were performed under the same incubation conditions. The samples were denatured with 6x loading buffer without DTT and then analysed by Western-blot with a rabbit anti-Mouse IgG [HRP] antibody (Figure 3.9).

When compared with the respective controls, it was possible to observe a decrease in IgG in the sample incubated with the dimeric form of APRc, which is not as apparent in the incubation with the monomer. However, since this reduction was not accompanied by the detection of any cleavage products, this assay could not confirm the Ig-cleavage activity of APRc.

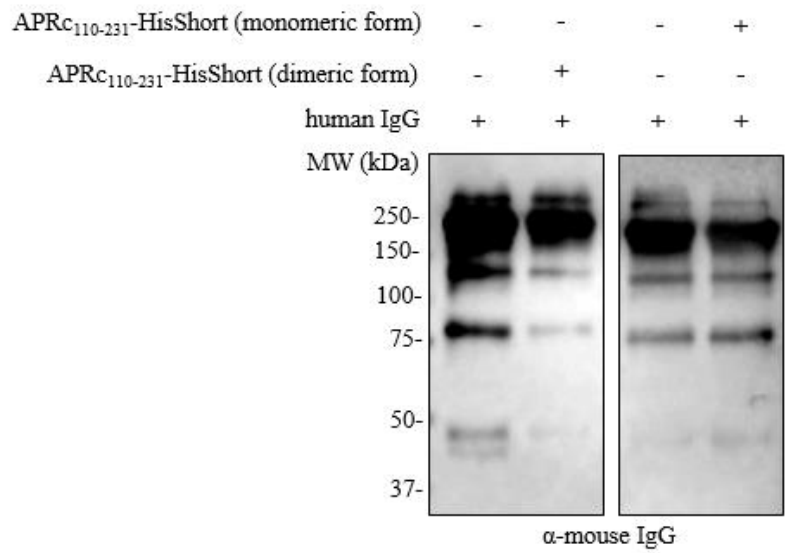


Figure 3.9 (previous page) - Evaluation of human IgG-APRc binding with dimeric and monomeric forms of APRc. Recombinant $APRc_{110-231}$-HisShort, dimeric and monomeric forms, were incubated with human IgG (#I2511, Sigma-Aldrich) at a 50 µg:6 µg ratio (S:E), in PBS buffer pH 7.4, overnight at 37ºC. As a control, incubations in the same conditions, only with human IgG were performed. 3 µL of each sample were denatured with 6x loading buffer without DTT, applied in a 10% polyacrylamide gel, and then evaluated by Western-blot analysis with the polyclonal rabbit anti-Mouse IgG (whole molecule)-peroxidase (A9044, Sigma-Aldrich) (1:5 000 in TBS-T 2% BSA).

In summary, under all conditions here described, it can be concluded that the interaction of APRc with human or rabbit IgG promotes the oligomeric stabilization of APRc. However, there is no apparent evidence that APRc cleaves IgG.

3.2.2 Evaluation of APRc binding to immunoglobulins from different species and different classes

Many bacteria present non-immune Ig-binding proteins that are able to interact with immunoglobulins from different species and from different classes and subclasses[24]. Therefore, it was evaluated if APRc is also able to interact with other immunoglobulins, besides the above reported interaction between APRc and human/rabbit IgG. These interactions were evaluated by enzyme-linked immunosorbent assay (ELISA) with streptavidin-biotin detection, and for that biotinylated APRc was recombinantly produced in E. *coli* and purified.

Production of biotinylated APRc

Biotinylated APRc was generated by enzymatic biotinylation, where the protein fused with a biotin accepting peptide (avi-tag) was co-expressed with biotin protein ligase (birA). The Avi-tag functions as a recognition site for birA, enabling biotinylation to occur in the presence of biotin. The expression construct pCoofy_$HisAPRc_{110-231}$-avi - which includes APRc amino acids 110-231 fused at the N terminus with a His-tag followed by an HRV 3C cleavage site, and at the C terminus with a 15 amino acid Avi-tag peptide - was transformed in BL21 Star cells, containing the pDW363ΔMBP plasmid. Protein expression was induced at an OD_{600nm} of 0.6 for 3 hours at 30ºC, and APRc biotinylation was confirmed by Western blot analysis with streptavidin-HRP (data not shown).The cells were harvested by centrifugation and frozen.

In order to separate the recombinant protein from the other components of the cell lysates, the total soluble protein extract was purified by immobilized metal affinity chromatography with Ni^{2+}. As mentioned, recombinant biotinylated APRc has a C-terminal His-tag, enabling its purification using an IMAC column with coupled nickel ions (Histrap HP 5 mL column). Protein elution was carried out by implementing a four-step gradient of imidazole (50 mM, 150 mM, 200 mM, and 500 mM) in 20 mM sodium phosphate pH 7.4 buffer containing 500 mM NaCl. The chromatogram corresponding to the IMAC-Ni^{2+} purification of biotinylated APRc is represented in Figure 3.10 (A) and shows the monitoring of protein elution at an absorbance of 280 nm and the respective gradient of imidazole.

The protein of interest was eluted at 150 mM imidazole. Those fractions were pooled, dialyzed overnight against 20 mM Tris/HCl pH 8 buffer, and further purified by anion-exchange chromatography in a Mono-Q 5/50 GL column, previously equilibrated with 20 mM Tris/HCl pH 8 buffer. The recombinant biotinylated APRc is negatively charged at that pH, enabling the binding to the positively charged column. Protein elution was carried out by implementing a linear gradient of NaCl (0-1 M) in Tris/HCl pH 8 buffer. The chromatogram corresponding to the anion-exchange chromatography of biotinylated APRc is represented in Figure 3.10 (B), and shows the monitoring of protein elution at an absorbance of 280 nm, and the respective NaCl concentration given by monitoring buffer conductivity.

Protein samples of the eluted fractions from the IMAC-Ni^{2+} and the anion-exchange chromatography were analysed by SDS-PAGE followed by Coomassie blue staining. The SDS-PAGE images are shown in Figure 3.10 (C).

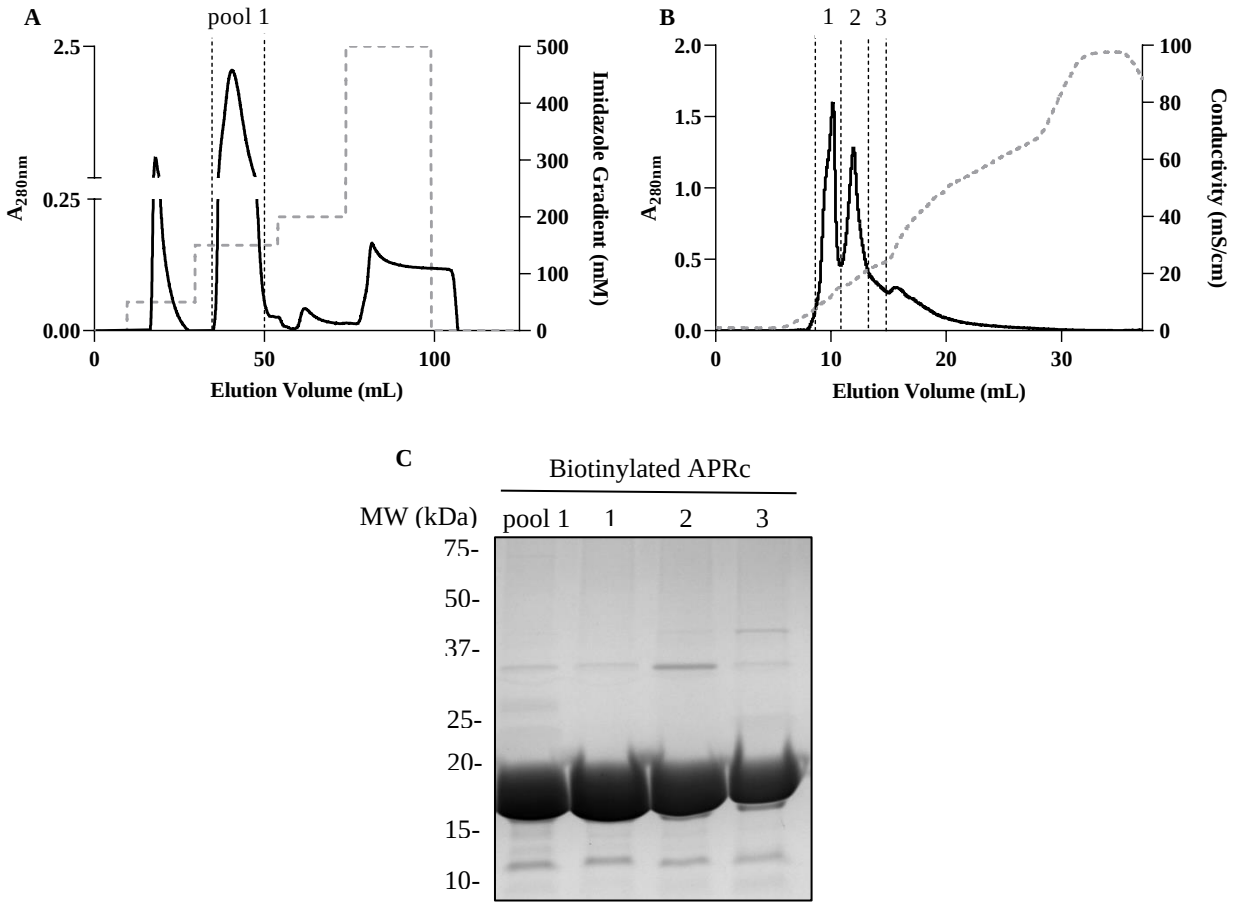


Figure 3.10- Purification of biotinylated APRc by IMAC Ni^{2+} and anion-exchange chromatography. (A) The recombinant biotinylated APRc was purified by IMAC Ni^{2+} on a Histrap HP 5 mL column, previously equilibrated in 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl. Protein elution was carried out by implementation of a four-step gradient of imidazole (50 mM, 150 mM, 200 mM, and 500 mM) in 20 mM sodium phosphate pH 7.4 buffer containing 500 mM NaCl (represented by the grey dashed line) at a flow rate of 5 mL/min and monitored at an absorbance of 280 nm. The eluted fractions of the 150 mM imidazole gradient step were pooled (pool 1) and dialyzed overnight against 20 mM Tris/HCl pH 8 buffer. **(B)** After dialysis, the protein was purified by anion-exchange chromatography in a Mono-Q 5/50 GL column, previously equilibrated with 20 mM Tris/HCl pH 8 buffer. Protein elution was carried out by implementation of a linear gradient of NaCl (0-1 M) in 20 mM Tris/HCl pH 8 buffer, at a flow rate of 0.75 mL/min and monitored at an absorbance of 280 nm. The conductivity (represented by the grey dashed line) was also monitored. The selected eluted fractions (1, 2 and 3) are outlined by dashed lines. **(C)** SDS-PAGE analysis followed by Coomassie blue staining of the pooled eluted fractions (pool 1) resultant from the purification of the recombinant biotinylated APRc by IMAC Ni^{2+} and the selected eluted fractions 1, 2 and 3 from the purification by anion-exchange chromatography. 20 µL of each sample were denatured with 6x loading buffer with DTT and then applied to a 12.5% polyacrylamide gel.

After purification, the His-tag at the N terminus of recombinant APRc was removed by digestion at the HRV 3C consensus cleavage site downstream of the His-tag sequence. The collected fractions 1, 2, and 3 were independently incubated with HRV 3C protease in 50 mM Tris-HCl pH 7.5 buffer

containing 150 mM NaCl overnight at 4ºC. Afterwards, to clean-up the samples, the digestion products were incubated with Ni Sepharose High Performance beads for 15 minutes at room temperature. The Ni beads were removed, and dialysis was performed against PBS for 3 hours at 4ºC. Samples before and after digestion, and after purification with Ni Sepharose High Performance beads were collected, and analysed by SDS-PAGE followed by Coomassie blue staining. Figure 3.11 (A) shows that tag removal and sample clean-up were successfully accomplished for all APRc fractions.

Since this purification resulted in three different fractions of biotinylated APRc, the interaction between each fraction and human IgG was initially evaluated by sandwich enzyme-linked immunosorbent assay (ELISA) to select the fraction with the best performance for follow-up assays. To this end, ELISA plates were firstly coated with human IgG (#I2511, Sigma-Aldrich) and then biotinylated APRc diluted in PBS buffer at different concentrations (0 µg/mL, 20 µg/mL, 50 µg/mL, 200 µg/mL, 450 µg/mL, 600 µg/mL, and 800 µg/mL) was added per well. APRc binding was probed with HRP-conjugated Streptavidin and detected with 1-Strep™ Ultra TMB ELISA substrate. The titration results of the binding of each eluted fraction of biotinylated APRc to human IgG are represented in Figure 3.11 (B).

From these results, it was possible to observe some differences in binding between these biotinylated forms, which may likely result from some variation in folding (charge distribution) among APRc molecules, as anticipated by the elution profile observed in the Mono Q (Figure 3.10 (B)). Therefore, based on its highest binding capacity, fraction 3 was selected to be used in the following assays.

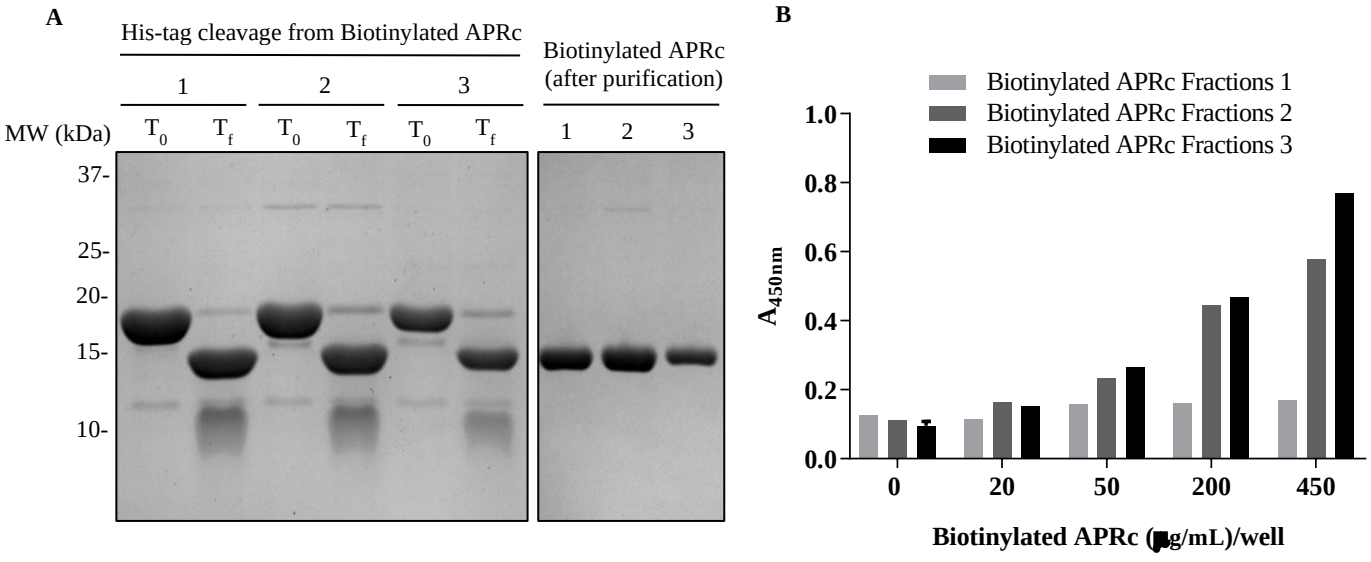


Figure 3.11 (previous page)- His-tag removal from biotinylated APRc and evaluation of APRc fraction with higher interaction with IgG through sandwich enzyme-linked immunosorbent assay (ELISA) with streptavidin-biotin detection. (A) The collected fractions 1, 2 and 3 of biotinylated APRc (from Figure 3.10) were incubated with HRV 3C protease in 50 mM Tris-HCl pH 7.5 buffer containing 150 mM NaCl overnight at 4ºC followed by incubation with Ni Sepharose High Performance beads for 15 minutes at room temperature. The Ni Sepharose High Performance beads were removed, and a dialysis was performed against PBS for 3 hours at 4ºC. Samples before (T_0) and after (T_f) digestion, and after purification with Ni Sepharose High Performance beads were collected and analysed by SDS-PAGE followed by Coomassie blue staining. 5 µL of each sample were applied in a 12.5% polyacrylamide gel**. (B)** 96 well ELISA plates were coated with 1 µg of human IgG (#I2511, Sigma-Aldrich) per well followed by incubation at 37ºC for 2 hours. The remaining protein-binding sites in the coated wells were blocked by incubation at 37ºC for 2 hours with PBS-T buffer containing 3% of BSA and then biotinylated APRc diluted in PBS buffer at different concentrations (0 µg/mL, 20 µg/mL, 50 µg/mL, 200 µg/mL, 450 µg/mL, 600 µg/mL, and 800 µg/mL) was added per well, followed by incubation at 37ºC for 1 hour. HRP-conjugated Streptavidin (Cell Signaling Technology) (1:2000) in PBS-T buffer containing 3% of milk was added to each well and incubated at 37ºC for 1 hour. 1-Strep™ Ultra TMB ELISA substrate was added to each well and incubated at room temperature for 15/20 minutes followed by the addition of 2 M sulfuric acid per well. The evaluation of biotinylated APRc binding to human IgG was performed through detection of HRP activity towards the TMB substrate at an absorbance of 450nm.

Interaction of biotinylated APRc to immunoglobulins from different species and different classes

The interaction between biotinylated APRc and immunoglobulins from different species and different classes was then evaluated by ELISA with streptavidin-biotin detection. ELISA plates were firstly coated with human IgG, mouse IgG, polyclonal rabbit anti-shewasin D antibody[71], human IgA, and human IgM and then biotinylated APRc diluted in PBS buffer at different concentrations (0 µg/mL, 20 µg/mL, 50 µg/mL, 200 µg/mL, 450 µg/mL, 600 µg/mL, and 800 µg/mL) was added per well. APRc binding was probed with HRP-conjugated Streptavidin and detected with 1-Strep™ Ultra TMB ELISA substrate. Three replicates of each condition were performed. The graphs representing ELISA results on the titration of APRc binding to the different immunoglobulins are shown in Figure 3.12 (A) and (B).

These results showed that recombinant biotinylated APRc can bind to IgG from different species in a concentration-dependent manner since interaction with rabbit, human, and mouse IgG was detected. The recombinant biotinylated APRc shows higher interaction with rabbit IgG, followed by human IgG, and mouse IgG. Moreover, recombinant biotinylated APRc displays the ability to bind to different classes of human immunoglobulins, since interaction with human IgG, IgM, and IgA was also observed. The recombinant biotinylated APRc shows higher interaction with human IgG, followed by human IgM and human IgA.

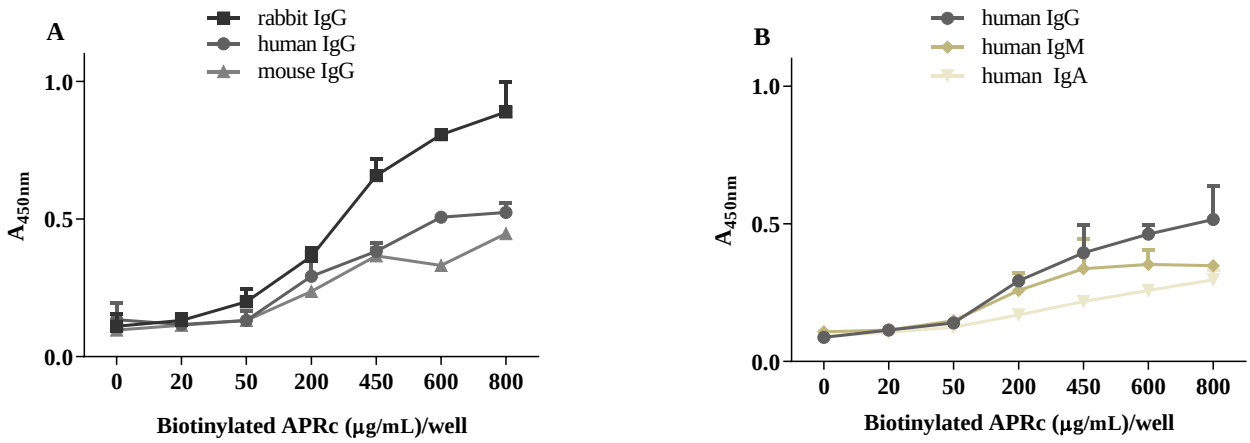


Figure 3.12 - Evaluation of binding of biotinylated APRc to immunoglobulins from different species and different classes. (A) 96-well ELISA plates were coated with 1 µg of human IgG (#I2511, Sigma-Aldrich), mouse IgG (#I5381, Sigma-Aldrich), polyclonal rabbit anti-shewasin D antibody (produced by Genscript)[71] per well followed by incubation at 37ºC for 2 hours. **(B)** 96 well ELISA plates were coated with 1 µg of human IgG (#I2511, Sigma-Aldrich), human IgA (#I4036, Sigma-Aldrich), and human IgM (#I8260, Sigma-Aldrich) per well followed by incubation at 37ºC for 2 hours. All plates were blocked by incubation at 37ºC for 2 hours with PBS-T buffer containing 3% of BSA and then biotinylated APRc diluted in PBS buffer at different concentrations (0 µg/mL, 20 µg/mL, 50 µg/mL, 200 µg/mL, 450 µg/mL, 600 µg/mL, and 800 µg/mL) were added per well, followed by incubation at 37ºC for 1 hour. HRP-conjugated Streptavidin (Cell Signaling Technology) (1:2000) in PBS-T buffer containing 3% of milk was added to each well and incubated at 37ºC for 1 hour. 1-Strep™ Ultra TMB ELISA substrate was added to each well and incubated at room temperature for 15/20 minutes followed by the addition of 2 M sulfuric acid per well. The evaluation of biotinylated APRc binding to immunoglobulins was performed through detection of HRP activity towards the TMB substrate at an absorbance of 450nm. Data represent mean values +/- standard deviations from three replicates per condition.

3.2.3 Mapping of the regions in APRc and IgG responsible for binding

Region in APRc responsible for binding to immunoglobulins

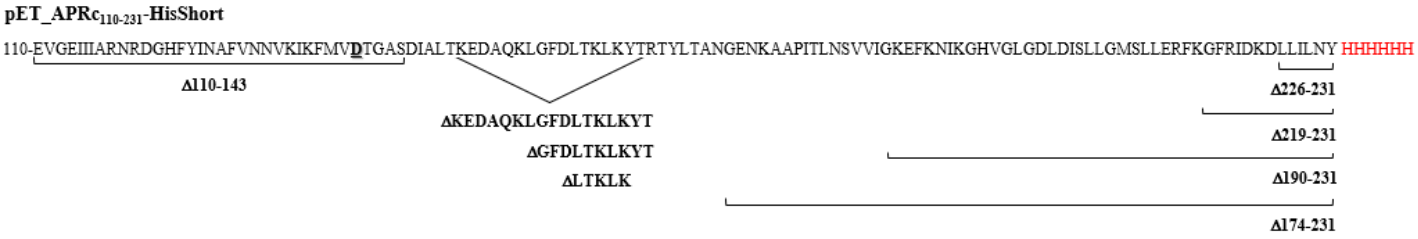
For the evaluation of the region of APRc responsible for binding to immunoglobulins, several truncated forms of the soluble domain were generated, comprising deletions at each terminal and internal deletions [Figure 3.13 (A) and (B)]. The latter corresponding to various-size deletions of a wide loop in APRc structure, which displays a very distinct conformation from all the other structures of retropepsins[68]. The soluble domain pET_APRc$_{110-231}$-HisShort and the different truncated forms of APRc were then recombinantly expressed in *E. coli* BL21 Star and evaluated for binding to IgG by Far-Western. All cultures were inoculated at a starting OD_{600nm} of 0.05 and protein expression induced at an OD_{600nm} of 0.6, and left overnight at 20ºC. Before cell lysis, E. *coli* growth was assessed (OD_{600nm}) [Figure 3.14 (A)]. A cell density corresponding to an OD_{600nm} of 3, per condition, was harvested by

centrifugation and the pellets resuspended in BugBuster Protein Extraction Reagent. The total fractions were denatured with 6xloading buffer with DTT and then evaluated by SDS-PAGE followed by Coomassie blue staining [Figure 3.14 (B)] and by Western blot with the rabbit anti-Mouse IgG [HRP] antibody [Figure 3.14 (C)]. To evaluate the differences in binding to IgG, from the different truncated forms of APRc compared to $APRc_{110-231}$-HisShort, a quantification of the relative levels of signal intensity in Western blot normalized to total protein levels in Coomassie stained gels was performed [Figure 3.14 (D)].

A slightly lower *E. coli* growth was observed for the recombinant forms $APRc_{110-231}$-HisShort and $APRc_{144-231}$-HisShort, when compared with the other truncated forms of APRc [Figure 3.14 (A)], which may suggest some toxicity associated with these protein products. From the Coomassie staining, it is possible to confirm the expression of all APRc constructs with molecular weights ranging between 14 and 12 kDa, with a lower accumulation of $APRc_{144-231}$-HisShort compared with the others [Figure 3.14 (B)].

When evaluated by Western-blot analysis with the rabbit anti-mouse IgG [HRP] antibody, all recombinant forms of APRc displayed binding to the antibody. Although these results suggest that more than one region of APRc contributes to IgG binding, a significant decrease in signal was observed for the recombinant forms $APRc(\Delta GFDLTKLKYT)_{110-231}$-HisShort and $APRc(\Delta KEDAQKLGFDLTKLKYT)_{110-231}$-HisShort [Figure 3.14 (C) and (D)], suggesting that this wide loop in APRc (residues 150 to 166) is one of the regions contributing to complex formation. Interestingly, this region is usually among the most variable regions in retropepsins, being reported to grant diverse immunological properties to several retroviral proteases[68], further confirming its surface-exposure. However, further studies are required to identify the other regions in APRc involved in IgG-binding.

A



B

pET_APRc$_{110-231}$-HisShort
110-EVGEIIIARNRDGHFYINAFVNNVKIKFMVDTGASDIALTKEDAQKLGFDLTKLKYTRTYLTANGENKAAPITLNSVVIGKEFKNIKGHVGLGDLDISLLGMSLLERFKGFRIDKDLLILNY HHHHHH

pET_APRc$_{110-225}$-HisShort
110-EVGEIIIARNRDGHFYINAFVNNVKIKFMVDTGASDIALTKEDAQKLGFDLTKLKYTRTYLTANGENKAAPITLNSVVIGKEFKNIKGHVGLGDLDISLLGMSLLERFKGFRIDKD HHHHHH

pET_APRc$_{110-218}$-HisShort
110-EVGEIIIARNRDGHFYINAFVNNVKIKFMVDTGASDIALTKEDAQKLGFDLTKLKYTRTYLTANGENKAAPITLNSVVIGKEFKNIKGHVGLGDLDISLLGMSLLERFK HHHHHH

pET_APRc$_{110-189}$-HisShort
110-EVGEIIIARNRDGHFYINAFVNNVKIKFMVDTGASDIALTKEDAQKLGFDLTKLKYTRTYLTANGENKAAPITLNSVVIG HHHHHH

pET_APRc$_{110-173}$-HisShort
110-EVGEIIIARNRDGHFYINAFVNNVKIKFMVDTGASDIALTKEDAQKLGFDLTKLKYTRTYLTAN HHHHHH

pET_APRc$_{144-231}$-HisShort
144-DIALTKEDAQKLGFDLTKLKYTRTYLTANGENKAAPITLNSVVIGKEFKNIKGHVGLGDLDISLLGMSLLERFKGFRIDKDLLILNY HHHHHH

pET_APRc(ΔLTKLK)$_{110-231}$-HisShort
110-EVGEIIIARNRDGHFYINAFVNNVKIKFMVDTGASDIALTKEDAQKLGFDYTRTYLTANGENKAAPITLNSVVIGKEFKNIKGHVGLGDLDISLLGMSLLERFKGFRIDKDLLILNY HHHHHH

pET_APRc(ΔGFDLTKLKYT)$_{110-231}$-HisShort
110-EVGEIIIARNRDGHFYINAFVNNVKIKFMVDTGASDIALTKEDAQKLRTYLTANGENKAAPITLNSVVIGKEFKNIKGHVGLGDLDISLLGMSLLERFKGFRIDKDLLILNY HHHHHH

pET_APRc(ΔKEDAQKLGFDLTKLKYT)$_{110-231}$-HisShort
110-EVGEIIIARNRDGHFYINAFVNNVKIKFMVDTGASDIALTRTYLTANGENKAAPITLNSVVIGKEFKNIKGHVGLGDLDISLLGMSLLERFKGFRIDKDLLILNY HHHHHH

Figure 3.13- Representation of the amino acid sequence of the soluble domain pET_APRc$_{110-231}$-HisShort and the respective truncated forms of this construct. (A) Schematic representation of the deletions performed for the design of the different truncated forms of the soluble domain of APRc. The truncated forms include deletion at the N terminus of the protein (deletion of the residues 110 to 143), deletions at the C terminus of the protein (deletion of the residues 174 to 231, 190 to 231, 219-231, and 226 to 231) and internal deletions, corresponding to various-size deletions of a wide loop in APRc structure (deletion of the residues LTKLK, GFDLTKLKYT and KEDAQKLGFDLTKLKYT). **(B)** Amino acid sequences of the different constructs of APRc used in this study. All constructs are fused with the C-terminal tag sequence HHHHHH. The catalytic aspartate of APRc and the C-terminal tag sequence HHHHHH are highlighted in each construct.

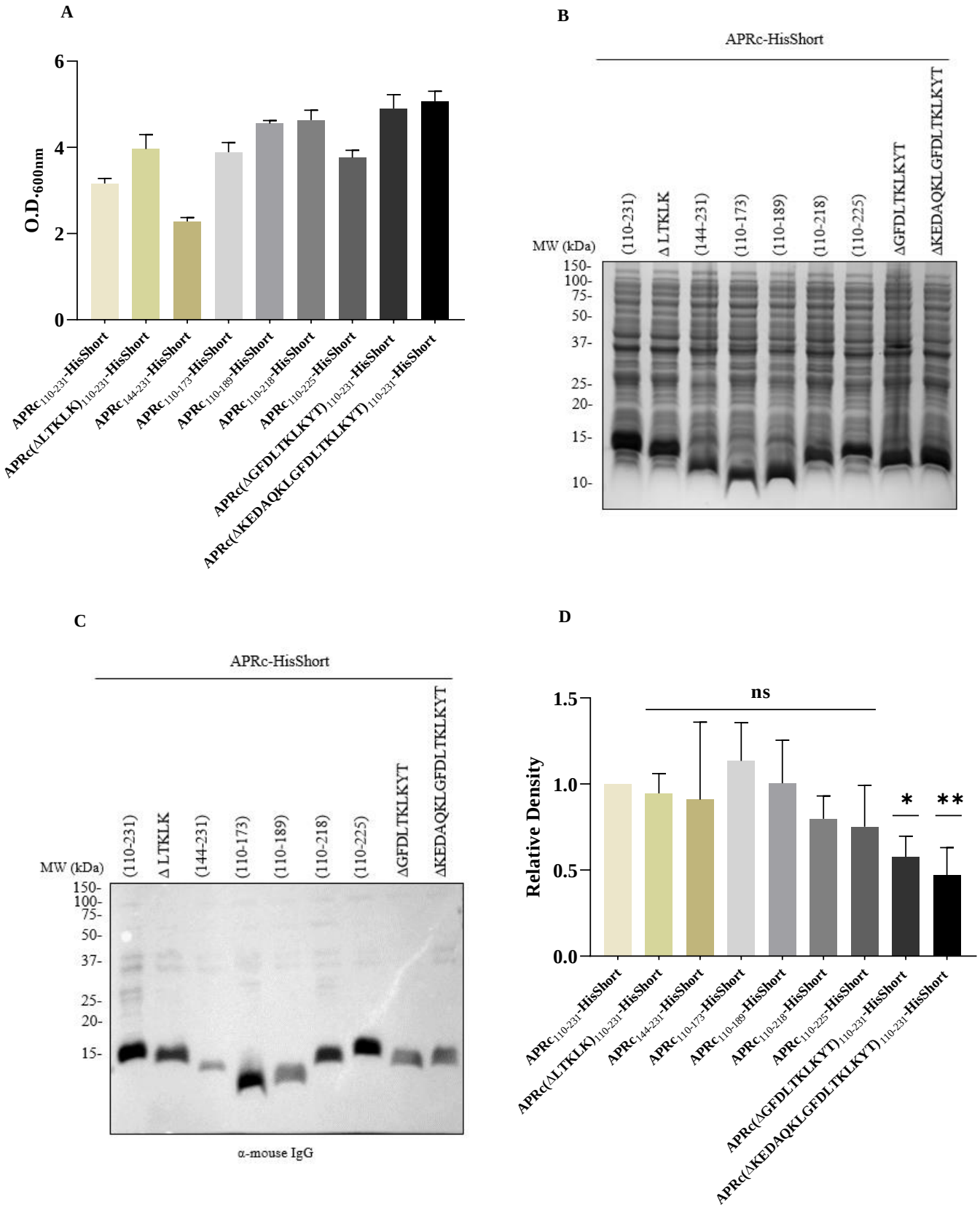

A
O.D.600nm
6
4
2
0
APRc110-231-HisShort
APRc(ΔLTKLK)110-231-HisShort
APRc144-231-HisShort
APRc110-173-HisShort
APRc110-189-HisShort
APRc110-218-HisShort
APRc110-225-HisShort
APRc(ΔGFDLTKLKYT)110-231-HisShort
APRc(ΔKEDAQKLGFDLTKLKYT)110-231-HisShort
B
APRc-HisShort
MW (kDa)
150-
100-
75-
50-
37-
25-
20-
15-
10-
(110-231)
Δ LTKLK
(144-231)
(110-173)
(110-189)
(110-218)
(110-225)
ΔGFDLTKLKYT
ΔKEDAQKLGFDLTKLKYT
C
APRc-HisShort
MW (kDa)
150-
100-
75-
50-
37-
25-
20-
15-
(110-231)
Δ LTKLK
(144-231)
(110-173)
(110-189)
(110-218)
(110-225)
ΔGFDLTKLKYT
ΔKEDAQKLGFDLTKLKYT
α-mouse IgG
D
Relative Density
1.5
1.0
0.5
0.0
ns
*
**
APRc110-231-HisShort
APRc(ΔLTKLK)110-231-HisShort
APRc144-231-HisShort
APRc110-173-HisShort
APRc110-189-HisShort
APRc110-218-HisShort
APRc110-225-HisShort
APRc(ΔGFDLTKLKYT)110-231-HisShort
APRc(ΔKEDAQKLGFDLTKLKYT)110-231-HisShort

Figure 3.14 (previous page)- Evaluation of the region in APRc responsible for binding to immunoglobulins. pET_APRc$_{110-231}$-HisShort, pET_APRc$_{144-231}$-HisShort, pET_APRc(ΔLTKLK)$_{110-231}$-HisShort, pET_APRc(ΔGFDLTKLKYT)$_{110-231}$-HisShort, pET_APRc(ΔKEDAQKLGFDLTKLKYT)$_{110-231}$-HisShort, pET_APRc$_{110-173}$-HisShort, pET_APRc$_{110-189}$-HisShort, pET_APRc$_{110-218}$-HisShort, and pET_APRc$_{110-225}$-HisShort (see Figure 3.13 for sequences) were independently transformed in BL21 Star cells and then grown at 37ºC in 1 mL of LB medium with 100 µg/mL ampicillin. When an OD_{600nm} of 0.6 was reached, the expression was induced overnight with IPTG at 20ºC. Before protein extraction, the E.*coli* growth was assessed at an absorbance of 600 nm. For protein extraction, the cells were harvested by centrifugation and the pellets resuspended in BugBuster Protein Extraction Reagent for 20 minutes at room temperature with agitation. The total fractions were denatured with 6xloading buffer with DTT. **(A)** Measurement of *E. coli* growth at an absorbance of 600 nm after protein expression. **(B)** SDS-PAGE analysis followed by Coomassie blue staining of total protein extracts for each recombinant truncated form of APRc. 20 µL of each sample were applied in a 12.5% polyacrylamide gel. **(C)** Representative Western-blot analysis of total protein extracts for each recombinant truncated form of APRc with the polyclonal rabbit anti-Mouse IgG (whole molecule)-peroxidase (A9044, Sigma-Aldrich) (1:5 000 in TBS-T 2% BSA). 20 µL of each sample were applied in a 12.5% polyacrylamide gel. **(D)** Graph shows quantification of relative levels of signal intensity in Western blot normalized to total protein levels in Coomassie stained gels. Results represent six independent replicates. Data represent mean values +/- standard deviations. Dunnet's one-way ANOVA multiple comparisons test was used to determine the significance between APRc$_{110-231}$-HisShort and truncated forms. *p<0.05 **p<0.01; ns: non-significant.

Region in IgG responsible for binding to APRc

For the evaluation of the region in IgG responsible for binding to APRc, different fragments of human IgG [F(ab')$_2$ and F(ab')] were produced and then used in cross-linking reactions with glutaraldehyde and recombinant dimeric APRc. The cross-linker glutaraldehyde helps to stabilize the weak interactions between two close molecules, enabling the visualization of the complexes formed by APRc and the different human IgG fragments.

For F(ab')$_2$ production, human IgG was digested with pepsin in 50 mM sodium acetate pH 4.0 buffer containing 100 mM NaCl. Samples from the beginning and the end of the digestion were collected, denatured with 6xloading buffer with and without DTT, and evaluated by SDS-PAGE followed by Coomassie blue staining [Figure 3.15 (A)]. This would allow the confirmation of the digestion of the human IgG. The expected molecular weight of human IgG when denatured without DTT is approximately 150 kDa; under reducing conditions two main protein bands are expected, corresponding to the heavy chain with 50 kDa and the light chain with 25 kDa. Under non-reducing conditions, human F(ab')$_2$ presents a molecular weight of approximately 110 kDa, and under reducing conditions, only a band with approximately 25 kDa is expected. The SDS-PAGE analysis of the samples confirmed digestion of the human IgG, as a band with approximately 110 kDa was detected. Also, in the samples denatured with DTT, only a protein band with approximately 25 kDa was detected in the digestion product [Figure 3.15 (A)].

After confirmation of digestion, 10 µM of pepstatin (inhibitor of pepsin) were added to inhibit the enzyme, and the sample was further concentrated in a Vivaspin 500, 50 kDa PES column. Samples were collected after concentration, denatured with 6xloading buffer with and without DTT, and analysed by Western-blot with mouse anti-human IgG Fc [HRP] antibody [Figure 3.15 (B)]. Since this antibody reacts with the Fc portion of human IgG, it would detect any undigested human IgG or Fc fragments in the samples. In the Western-blot analysis, it was possible to observe the residual presence of undigested human IgG, as a molecule with approximately 110 kDa was still detected in digested samples (denatured without DTT) [Figure 3.15 (B)].

To further clean-up the sample, digestion products were incubated with protein A Mag Sepharose. After the incubation, the resin was separated and the supernatant containing $F(ab')_2$ was collected. Samples collected after purification and intact human IgG, were denatured with 6xloading buffer with and without DTT and evaluated by Western-blot with anti-Mouse IgG [HRP] antibody [Figure 3.15 (C)]. This antibody would allow the detection of both $F(ab')_2$ and Fc fragments. In the Western-blot analysis, the presence of human $F(ab')_2$ and no Fc-fragments was confirmed, as only a protein band with approximately 110 kDa was detected in the samples denatured without DTT. When denatured with DTT, only a protein band with approximately 25 kDa was detected, and no band at 50 kDa was observed [Figure 3.15 (C)].

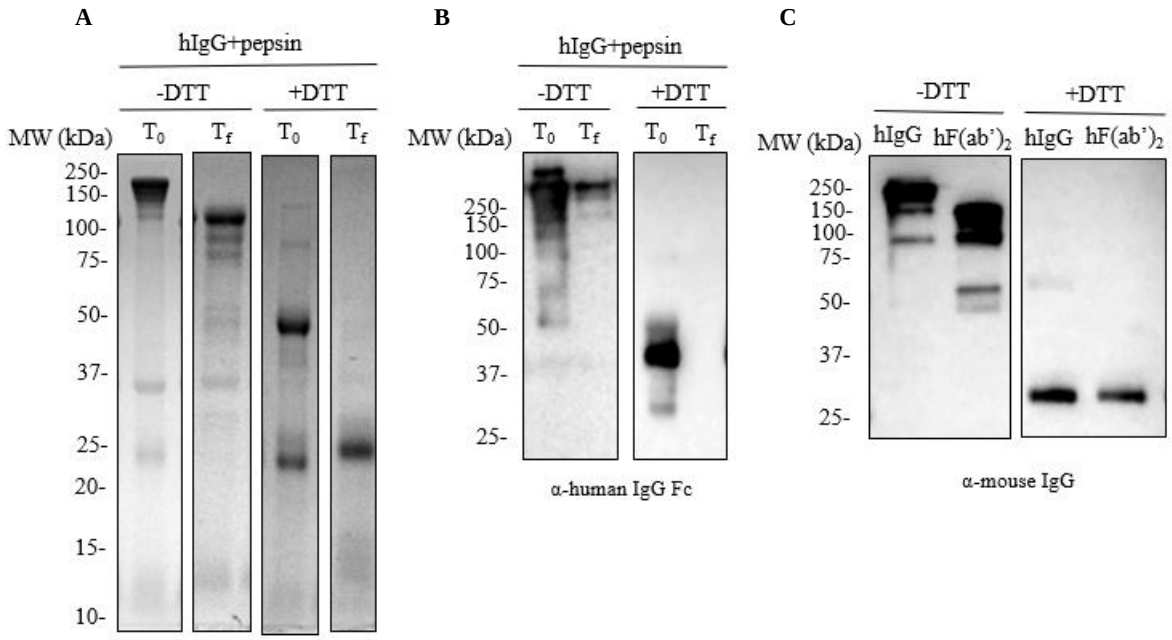


Figure 3.15 (previous page)- F(ab')$_2$ production and purification from the digestion of human IgG with pepsin. (A) Human IgG (#I2511, Sigma-Aldrich) was digested with pepsin in 50 mM sodium acetate pH 4.0 buffer containing 100 mM NaCl, for 8 hours and 30 minutes at 37ºC. Samples from the beginning and the end of the digestion with pepsin were collected, denatured with 6xloading buffer with and without DTT and evaluated by SDS-PAGE followed by Coomassie blue staining. 10 µL of each sample were applied to a 10% polyacrylamide gel. **(B)** After confirmation of the digestion of the human IgG, 10 µM of pepstatin were added to stop the digestion and the sample was diluted in PBS buffer and further concentrated in a Vivaspin 500, 50 kDa PES column. Samples were collected after concentration, denatured with 6xloading buffer with and without DTT. 3 µL of each sample were applied in a 10% polyacrylamide gel, and then evaluated by Western-blot analysis with the monoclonal mouse anti-Human IgG Fc [HRP] (A01854-200, GenScript) (1:20 000 in TBS-T 2% BSA)**. (C)** To remove the Fc fragments resultant from the digestion of IgG with pepsin or the undigested human IgG, the sample was incubated with protein A Mag Sepharose at room temperature for 2 hours. After the incubation, the resin was separated with a magnetic device and the supernatant containing F(ab')$_2$ was collected. Samples were collected after purification with protein A Mag Sepharose and together with samples of intact human IgG, were denatured with 6xloading buffer with and without DTT. 0.5 µg from each sample were applied in a 10% polyacrylamide gel and then evaluated by Western-blot analysis with the polyclonal rabbit anti-Mouse IgG (whole molecule)-peroxidase (A9044, Sigma-Aldrich) (1:5 000 in TBS-T 2% BSA).

For F(ab') production, human F(ab')$_2$ (pepsin-digested samples containing Fc fragments) were digested with papain in PBS pH 7.4 buffer containing 0.02 M EDTA and 0.02 M cysteine. Samples from the beginning and the end of the digestion were collected, denatured with 6xloading buffer without DTT, and evaluated by SDS-PAGE followed by Coomassie blue staining [Figure 3.16 (A)], to confirm the full digestion of the human F(ab')$_2$. Human F(ab'), when denatured without DTT, presents a molecular weight of 50 kDa, and when denatured with DTT, presents a molecular weight of 25 kDa. In the SDS-PAGE analysis, molecules of approximately 50, 37, 25, and 15 kDa were detected at T0. Human F(ab')$_2$ should display here a molecular weight of 110 kDa; therefore its appears that the digestion buffer somehow impacts the migration/reductive state of human F(ab')$_2$. Importantly, no molecules of 50 kDa (or the major band at 25 kDa) are present at the end of digestion, suggesting that it was successful [Figure 3.16 (A)].

To further purify these human F(ab') fragments, the sample was diluted with 20 mL of PBS pH 7.4 buffer, and then concentrated in an Amicon^R Ultra-4 Centrifugal Filter Unit (Ultracel 3000 Da). To further remove any remaining Fc fragments, the sample was incubated protein A Mag Sepharose. After the incubation, the resins were separated, and the supernatant containing F(ab') was collected. To inactivate papain, 10 µM of E-64 protease inhibitor were added. Samples collected after purification and samples of intact human IgG were denatured with 6xloading buffer without DTT and evaluated by Western-blot with the mouse anti-Human Ig Kappa Light Chain antibody followed by the rabbit anti-Mouse IgG [HRP] antibody, or with the mouse anti-human IgG Fc [HRP] antibody [Figure 3.16 (B)], in order to confirm the presence of human F(ab') and no presence of Fc-fragments, respectively. In the Western-blot analysis with the mouse anti-Human Ig Kappa Light Chain antibody, the presence of human F(ab') in the purified sample was confirmed (~50 kDa band). In the Western-blot analysis with

the mouse anti-human IgG Fc [HRP] antibody, no signal was detected in the F(ab') sample [Figure 3.16 (B)].

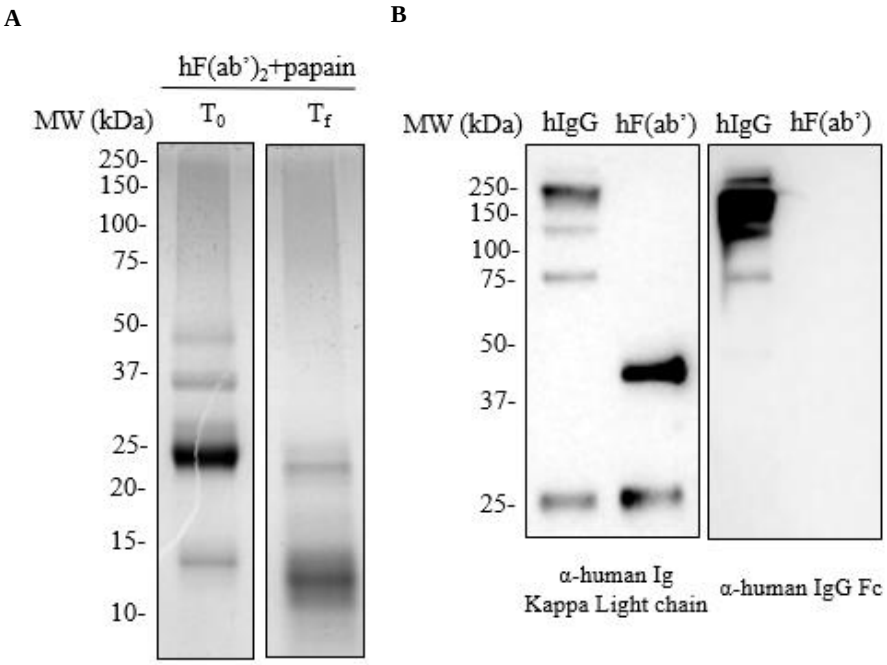


Figure 3.16- F(ab') production and purification from the digestion of human F(ab')$_2$ with papain. (A) human F(ab')$_2$ (pepsin-digested samples containing Fc fragments) were digested with papain in PBS pH 7.4 buffer containing 0.02 M EDTA and 0.02 M cysteine for 1 hour and 30 minutes at 37ºC. Samples from the beginning and the end of the digestion with papain were collected, denatured with 6xloading buffer without DTT and evaluated by SDS-PAGE followed by Coomassie blue staining. 10 µL of each sample were applied to a 10% polyacrylamide gel. **(B)** After confirmation of the digestion of the human F(ab')$_2$, the sample was diluted in 20 mL of PBS pH 7.4 buffer, and then concentrated in an Amicon^R Ultra-4 Centrifugal Filter Unit (Ultracel 3000 Da). To remove the Fc fragments, the sample was further incubated with protein A Mag Sepharose at room temperature for 2 hours. After the incubation, the resins were separated with a magnetic device and the supernatant containing F(ab') was collected. To stop the digestion, 10 µM of E-64 protease inhibitor were added. Samples were collected after purification with protein A Mag Sepharose and together with samples of intact human IgG, were denatured with 6xloading buffer without DTT. 0.5 µg of human IgG or purified F(ab') were applied in a 10% polyacrylamide gel and then evaluated by Western-blot analysis with the monoclonal mouse anti-Human Ig Kappa Ligth Chain [MAB10050, R&D Systems, Bio-techne] (1:259 in TBS-T 2% BSA) followed by the polyclonal rabbit anti-Mouse IgG (whole molecule)-peroxidase (A9044, Sigma-Aldrich) (1:10 000 in TBS-T 2% BSA) or with the monoclonal mouse anti-Human IgG Fc [HRP] (A01854-200, GenScript) (1:20 000 in TBS-T 2% BSA).

To evaluate the region in IgG responsible for binding to APRc, cross-linking reactions with glutaraldehyde between APRc and the different fragments of human IgG were performed. Recombinant dimeric APRc$_{110-231}$-HisShort and APRc$_{110-231}$-HisShort(D140N) were incubated with intact human IgG, purified human F(ab')$_2$, and purified human F(ab') fragments, in an equimolar ratio (166.85 ρmole), for 4 hours at 37ºC. Afterwards, the samples were treated with glutaraldehyde. As a control, similar incubations were treated with H_2O. All the samples were incubated for 4 min at 37ºC, and the reaction

was stopped by adding 1 M Tris-HCl pH 8.0. The samples were denatured with 6x loading buffer with DTT and analysed by Western blot. In Figure 3.17 are represented the Western blot analyses of the cross-linking reactions between recombinant APRc and human IgG with the following antibodies: i) rabbit anti-APRc antibody followed by the mouse anti-Rabbit IgG (M205) [HRP] antibody; ii) mouse anti-Human IgG Fc [HRP] antibody; and iii) mouse anti-Human Ig Kappa Light Chain antibody followed by the rabbit anti-Mouse IgG [HRP] antibody. In Figure 3.18 are represented the Western blot analyses of the cross-linking reactions between recombinant APRc and human $F(ab')_2$ with the antibodies: i) rabbit anti-APRc antibody followed by the mouse anti-Rabbit IgG (M205) [HRP] antibody; and ii) anti-Human Ig Kappa Light Chain antibody followed by the rabbit anti-Mouse IgG [HRP] antibody. In Figure 3.19, are represented the Western-blot analyses of the cross-linking reactions between recombinant APRc and human F(ab'), with the antibodies: rabbit anti-APRc antibody followed by the mouse anti-Rabbit IgG (M205) [HRP] antibody; anti-Human Ig Kappa Light Chain antibody followed by the rabbit anti-Mouse IgG [HRP] antibody.

From the Western-blot analysis of the cross-linking reaction between APRc (recombinant $APRc_{110-231}$-HisShort and $APRc_{110-231}$-HisShort(D140N)) and human IgG, it was clear a difference in APRc migration pattern, with a decrease in the signal of the monomeric and oligomeric forms of the protease and a shift to higher molecular weights. This was visible for both the wild-type and mutant forms, again confirming APRc and IgG interaction. When probed with the mouse anti-Human IgG Fc [HRP] antibody, it was impossible to assess any differences in human IgG Fc-fragments since, with the cross-linker, larger complexes are formed that are not visible by Western-blot analysis. From the Western-blot analysis with the mouse anti-Human Ig Kappa Light Chain antibody, slight differences in the migration of the light chain to higher molecular weights are noticed in the presence of APRc. This suggests that APRc might interact with the Fab region of human IgG (Figure 3.17).

This was further confirmed in the Western-blot analysis of the cross-linking reaction between APRc (recombinant $APRc_{110-231}$-HisShort and $APRc_{110-231}$-HisShort(D140N)) and human $F(ab')_2$ fragments. For both cross-linking incubations with the wild-type and mutant forms of APRc, evident differences in APRc migration were observed in the presence of human $F(ab')_2$, as a pattern of APRc shifted to higher molecular weights was observed. From the Western-blot analysis with the mouse anti-Human Ig Kappa Light Chain antibody, differences in the migration of human $F(ab')_2$ fragments were also noticed, as migration to higher molecular weights was also detected. This suggests that there is a similar stabilization of APRc (WT and mutant) with human $F(ab')_2$, corroborating the binding site of IgG to APRc in the Fab region (Figure 3.18).

To evaluate if the interaction would depend on an intact $F(ab')_2$ domain, purified F(ab') fragments were then tested in these cross-linking studies. From the Western-blot analysis shown in Figure 3.19, it

was possible to confirm again differences in the migration pattern of APRc wild-type and mutant forms in the presence of F(ab'), further corroborating the interaction of APRc and F(ab') and suggesting that APRc can still bind to this monovalent fragment of the antibody. Since the expected cross-linked APRc-F(ab') products are smaller (compared with an intact IgG or $F(ab')_2$)), it was possible to detect the appearance of these products with the mouse anti-Human Ig Kappa Light Chain. These results further suggest that APRc-Ig binding does not require an intact $F(ab')_2$ fragment.

In summary, these crosslinking assays confirmed APRc-human IgG interaction through the Fab region, independent of APRc catalytic activity. However, for a complete characterization of the binding site in human IgG, additional cross-linking reactions with Fc fragments should be performed to assess if APRc also interacts in the Fc region. Moreover, the interaction of APRc to a Single Chain Variable Fragment (ScFv) should also be tested, to determine the minimal binding unit in IgG responsible for the interaction with APRc, as it remains to be clarified if the interaction in the Fab region occurs through the variable or constant regions of the heavy or the light chain.

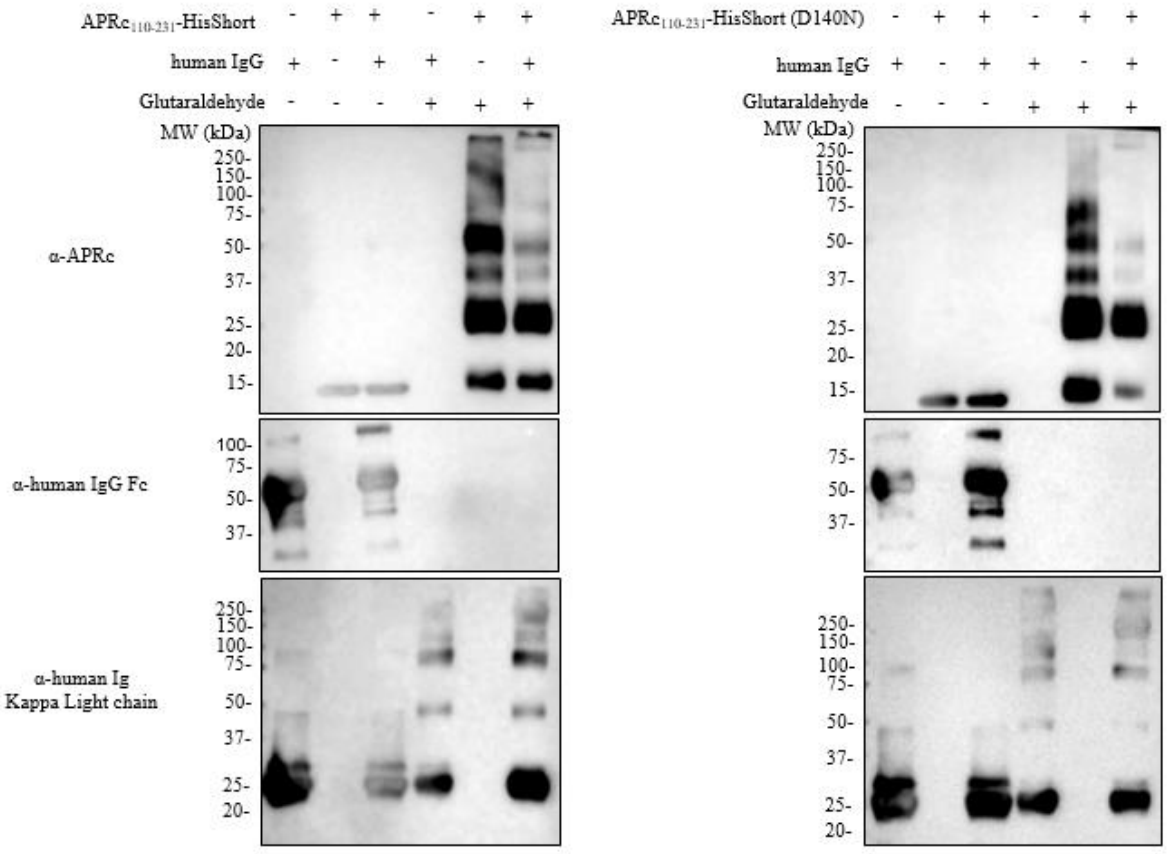


Figure 3.17 (previous page)- Cross-linking reaction with glutaraldehyde between recombinant APRc$_{110-231}$-HisShort or APRc$_{110-231}$-HisShort(D140N) and human IgG. Recombinant APRc$_{110-231}$-HisShort and APRc$_{110-231}$-HisShort(D140N) were incubated with intact human IgG (#I2511, Sigma-Aldrich) in an equimolar ratio (166.85 ρmole), for 4 hours at 37ºC. Afterwards, the samples were treated with glutaraldehyde. As a control, similar incubations were treated with H_2O. All the samples were incubated for 4 min at 37ºC, and the reaction was stopped by adding 1 M Tris-HCl pH 8.0. The samples were denatured with 6x loading buffer with DTT. 3 µl of each sample were applied in a 10% polyacrylamide gel and then evaluated by Western-blot analysis with the polyclonal rabbit antibody anti-APRc (raised towards the sequence Cys-Tyr-Thr-Arg-Thr-Tyr-Leu-Thr-Ala-Asn-Gly-Glu-Asn-Lys-Ala), produced by GenScript (1:500 in TBS-T 2% BSA) followed by the monoclonal mouse anti-rabbit IgG (M205) [HRP] (A01827-200, GenScript) (1:20 000 in TBS-T 2% BSA), with the monoclonal mouse anti-Human IgG Fc [HRP] (A01854-200, GenScript) (1:20 000 in TBS-T 2% BSA), or with the monoclonal mouse anti-Human Ig Kappa Ligth Chain [MAB10050, R&D Systems, Bio-techne] (1:259 in TBS-T 2% BSA) followed by the polyclonal rabbit anti-Mouse IgG (whole molecule)-peroxidase (A9044, Sigma-Aldrich) (1:10 000 in TBS-T 2% BSA.

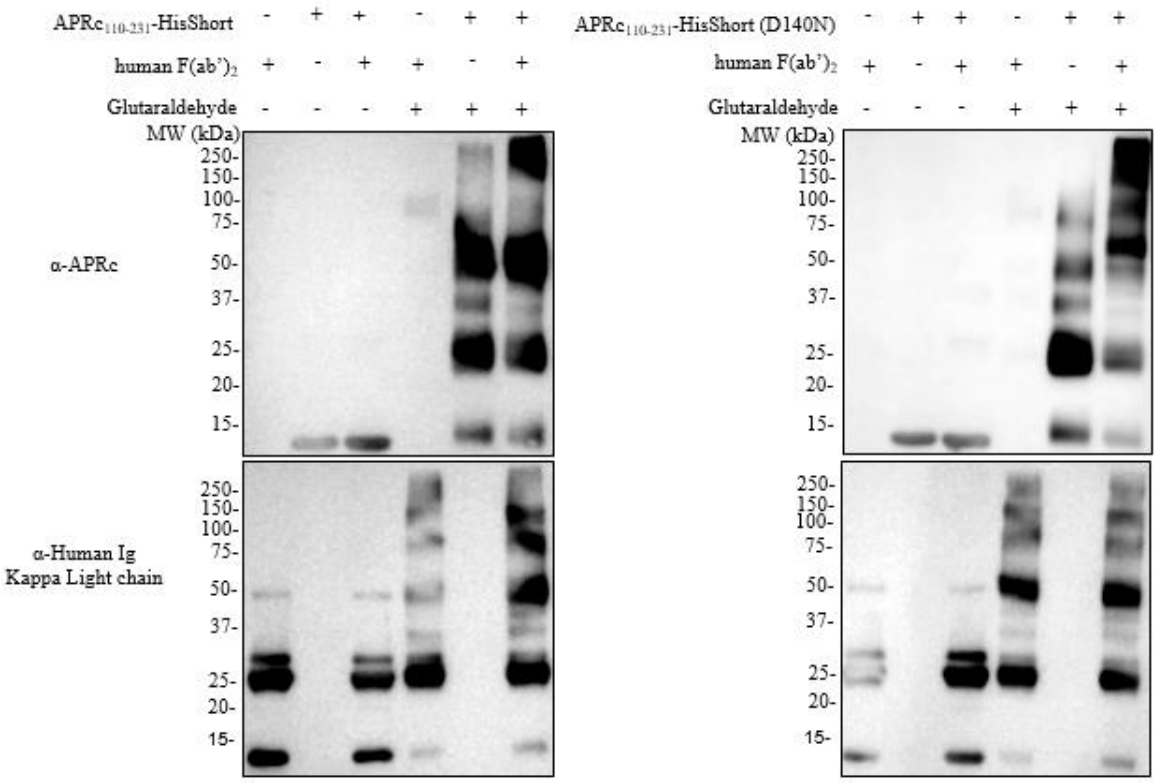


Figure 3.18- Cross-linking reaction with glutaraldehyde between recombinant APRc$_{110-231}$-HisShort or APRc$_{110-231}$-HisShort(D140N) and human F(ab')$_2$. Recombinant APRc$_{110-231}$-HisShort and APRc$_{110-231}$-HisShort(D140N) were incubated with human F(ab')$_2$ in an equimolar ratio (166.85 ρmole), for 4 hours at 37ºC. Afterwards, the samples were treated with glutaraldehyde. As a control, similar incubations were treated with H_2O. All the samples were incubated for 4 min at 37ºC, and the reaction was stopped by adding 1 M Tris-HCl pH 8.0. The samples were denatured with 6x loading buffer with DTT. 3 µl of each sample were applied in a 10% polyacrylamide gel and then evaluated by Western-blot analysis with the polyclonal rabbit antibody anti-APRc (raised towards the sequence Cys-Tyr-Thr-Arg-Thr-Tyr-Leu-Thr-Ala-Asn-Gly-Glu-Asn-Lys-Ala), produced by GenScript (1:500 in TBS-T 2% BSA) followed by the monoclonal mouse anti-rabbit IgG (M205) [HRP] (A01827-200, GenScript) (1:20 000 in TBS-T 2% BSA), or with the monoclonal mouse anti-Human Ig Kappa Ligth Chain [MAB10050, R&D Systems, Bio-techne] (1:259 in TBS-T 2% BSA) followed by the polyclonal rabbit anti-Mouse IgG (whole molecule)-peroxidase (A9044, Sigma-Aldrich) (1:10 000 in TBS-T 2% BSA.

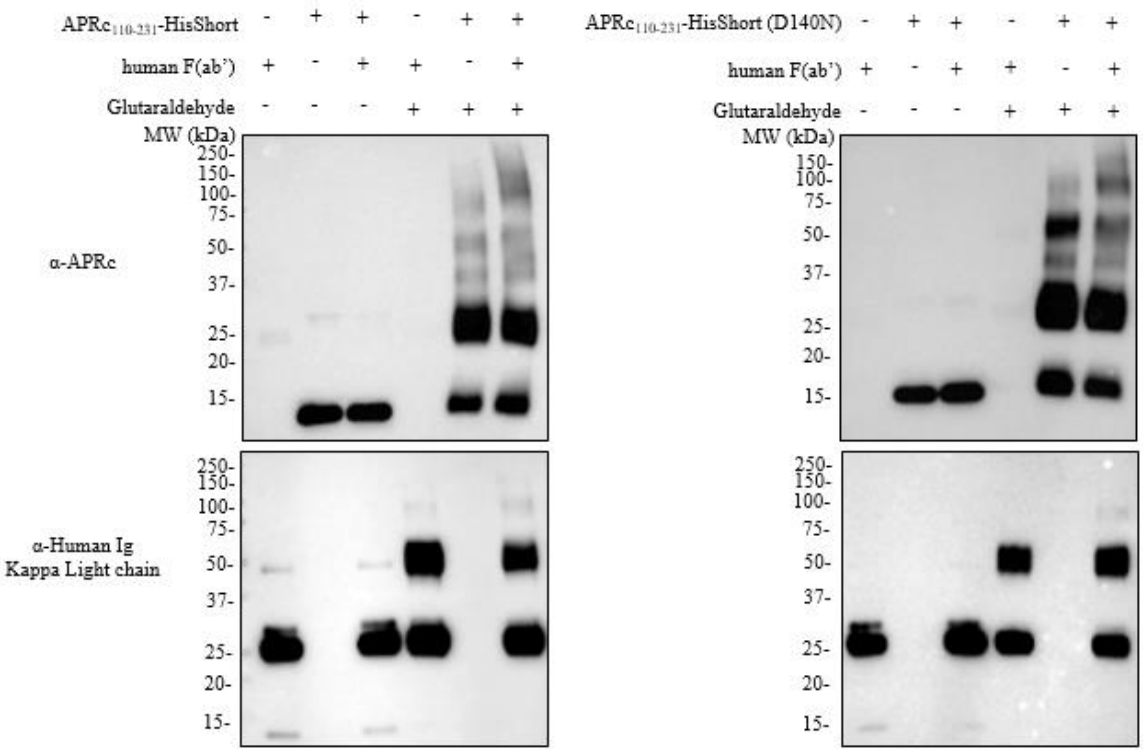


Figure 3.19- Cross-linking reaction with glutaraldehyde between recombinant APRc$_{110-231}$-HisShort or APRc$_{110-231}$-HisShort(D140N) and human F(ab'). Recombinant APRc$_{110-231}$-HisShort and APRc$_{110-231}$-HisShort(D140N) were incubated with human F(ab') in an equimolar ratio (166.85 ρmole), for 4 hours at 37ºC. Afterwards, the samples were treated with glutaraldehyde. As a control, similar incubations were treated with H_2O. All the samples were incubated for 4 min at 37ºC, and the reaction was stopped by adding 1 M Tris-HCl pH 8.0. The samples were denatured with 6x loading buffer with DTT. 3 µl of each sample were applied in a 10% polyacrylamide gel and then evaluated by Western-blot analysis with the polyclonal rabbit antibody anti-APRc (raised towards the sequence Cys-Tyr-Thr-Arg-Thr-Tyr-Leu-Thr-Ala-Asn-Gly-Glu-Asn-Lys-Ala), produced by GenScript (1:500 in TBS-T 2% BSA) followed by the monoclonal mouse anti-rabbit IgG (M205) [HRP] (A01827-200, GenScript) (1:20 000 in TBS-T 2% BSA), or with the monoclonal mouse anti-Human Ig Kappa Ligth Chain [MAB10050, R&D Systems, Bio-techne] (1:259 in TBS-T 2% BSA) followed by the polyclonal rabbit anti-Mouse IgG (whole molecule)-peroxidase (A9044, Sigma-Aldrich) (1:10 000 in TBS-T 2% BSA.

3.2.4 Evaluation of non-immune Ig-APRc binding in serum samples

The non-immune Ig-APRc interaction using purified immunoglobulins has been characterized in the previous sections. Since *Rickettsia* establishes infection in humans, being in contact with serum, the interaction of APRc with immunoglobulins in the complex context of normal human serum (NHS) was next evaluated. For the validation of the interaction between APRc and immunoglobulins in NHS, a pull-down of APRc and an immunoprecipitation of immunoglobulins from incubations of the recombinant APRc$_{110-231}$-HisShort and the respective active site mutant [APRc$_{110-231}$-HisShort(D140N)] with NHS were performed, followed by evaluation of the eluted samples by Western-blot analysis with anti-Human IgG Fc [HRP] and anti-APRc, respectively.

The pull-down assay of APRc was performed using His Mag Sepharose Ni Beads. His Mag Sepharose Ni Beads are magnetic beads with immobilized nickel ions. The recombinant forms of APRc utilized in this assay have a C-terminal His-tag with affinity for nickel ions, enabling its capture. After several optimizations, the pull-down was performed with the recombinant forms of APRc, $APRc_{110-231}$-HisShort and $APRc_{110-231}$-HisShort(D140N), that were individually incubated with His Mag Sepharose Ni Beads. After incubation, the magnetic beads were collected and incubated with NHS for 4 hours at 37º C. Protein elution was performed by incubation and denaturation with loading buffer with DTT diluted 6x in PBS. The beads were separated and the eluted fractions were collected. These fractions, together with input samples, were analysed by SDS-PAGE followed by Coomassie blue staining [Figure 3.20 (A)] and analysed by Western blot with mouse anti-Human IgG Fc [HRP] antibody [Figure 3.20 (B)].

The immunoprecipitation of immunoglobulins was performed using protein A Mag Sepharose. These resins with immobilized protein A ligands have a high affinity for monoclonal or polyclonal antibodies. The recombinant forms of APRc, $APRc_{110-231}$-HisShort and $APRc_{110-231}$-HisShort(D140N) were individually incubated with NHS for 4h at 37ºC. Afterwards, these suspensions were incubated with protein A Mag Sepharose and then the magnetic resins were collected using a magnetic device and washed with TBS buffer containing 250 mM NaCl and TBS buffer containing 1 % Octyl-β-Glucoside. Protein elution was then carried out by incubation and denaturation with loading buffer with DTT diluted 6x in PBS. The resins were separated and the eluted fractions were collected. Eluted fractions, together with input samples, were analysed by SDS-PAGE followed by Coomassie blue staining [Figure 3.21 (A)] and analysed by Western blot with rabbit anti-APRc antibody followed by the mouse anti-Rabbit IgG (M205) [HRP] antibody [Figure 3.21 (B)].

The Commassie blue staining of input and eluted fractions confirmed the pull-down of APRc [Figure 3.20 (A)] and immunoprecipitation of IgG [Figure 3.21 (A)]. Analysis of the pull-down samples by Western-blot with mouse anti-Human IgG Fc [HRP] antibody confirmed the detection of IgG in eluted fractions, further corroborating IgG binding to the immobilized APRc during its incubation with serum [Figure 3.20 (B)]. The same was confirmed when the reverse analysis was performed by immunoprecipitating human IgG from NHS. Both the wild-type and mutant forms of APRc were co-immunoprecipitated with IgG, as shown by the Western-blot analysis with rabbit anti-APRc antibody [Figure 3.21 (B)]. Therefore, these results showed that APRc interacts with human IgG in the context of NHS. Again, this interaction appears to be independent of the catalytic activity of the protease.

As an additional control to the evaluation of the non-immune Ig-APRc interaction in the context of NHS, the immunoprecipitation assay here described should be repeated in samples of NHS depleted from IgG, to confirm that no APRc is detected.

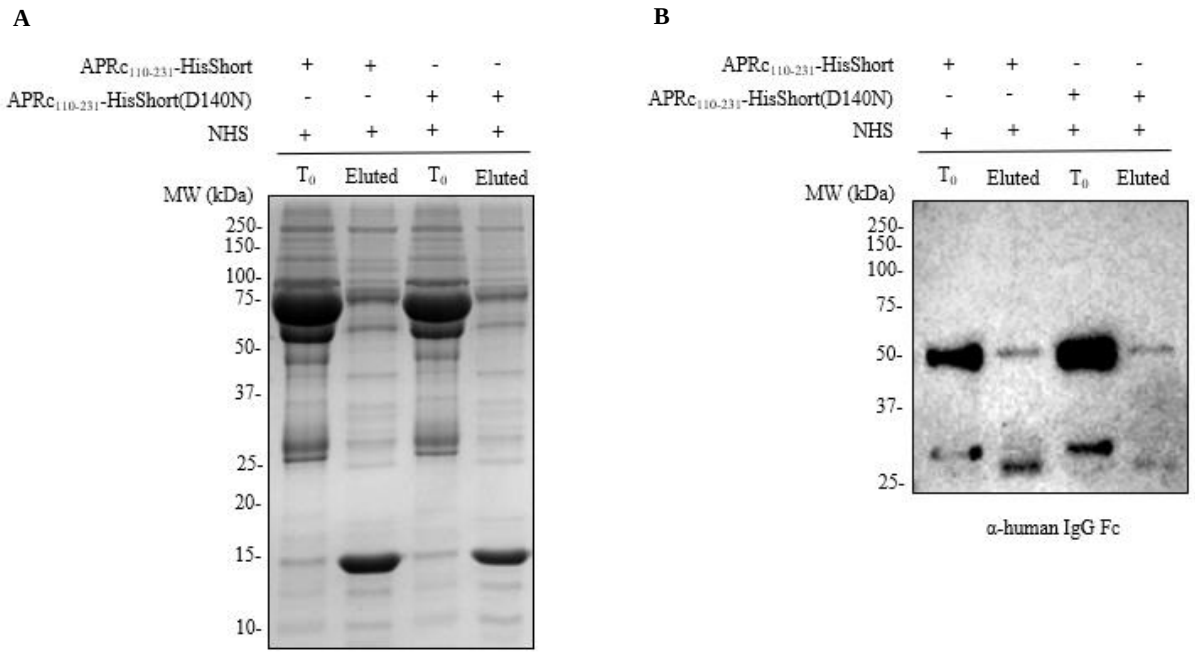


Figure 3.20- Evaluation of non-immune IgG-APRc binding in normal human serum by pull-down assay with His Mag Sepharose Ni Beads. The recombinant forms of APRc, $APRc_{110-231}$-HisShort and $APRc_{110-231}$-HisShort(D140N) were individually incubated with His Mag Sepharose Ni Beads for 1 hour and 30 minutes, at room temperature followed by incubated with NHS for 4 hours at 37º C. Magnetic beads were then washed with PBS buffer containing 200mM NaCl and PBS containing 0.1 % Tween. Protein elution was carried out by incubation and denaturation of the magnetic beads with loading buffer with DTT diluted 6x in PBS. **(A)** SDS-PAGE analysis followed by Coomassie blue staining of the eluted fractions (eluted) together with input samples (T_0). 10 µL of each sample were applied in a 12.5% polyacrylamide gel **(B)** Western-blot analysis of the eluted fractions (eluted) together with input samples (T_0). 3 µL of each sample were applied in a 10% polyacrylamide gel, and then evaluated by Western-blot analysis with the monoclonal mouse anti-Human IgG Fc [HRP] (A01854-200, GenScript) (1:20 000 in TBS-T 2% BSA.

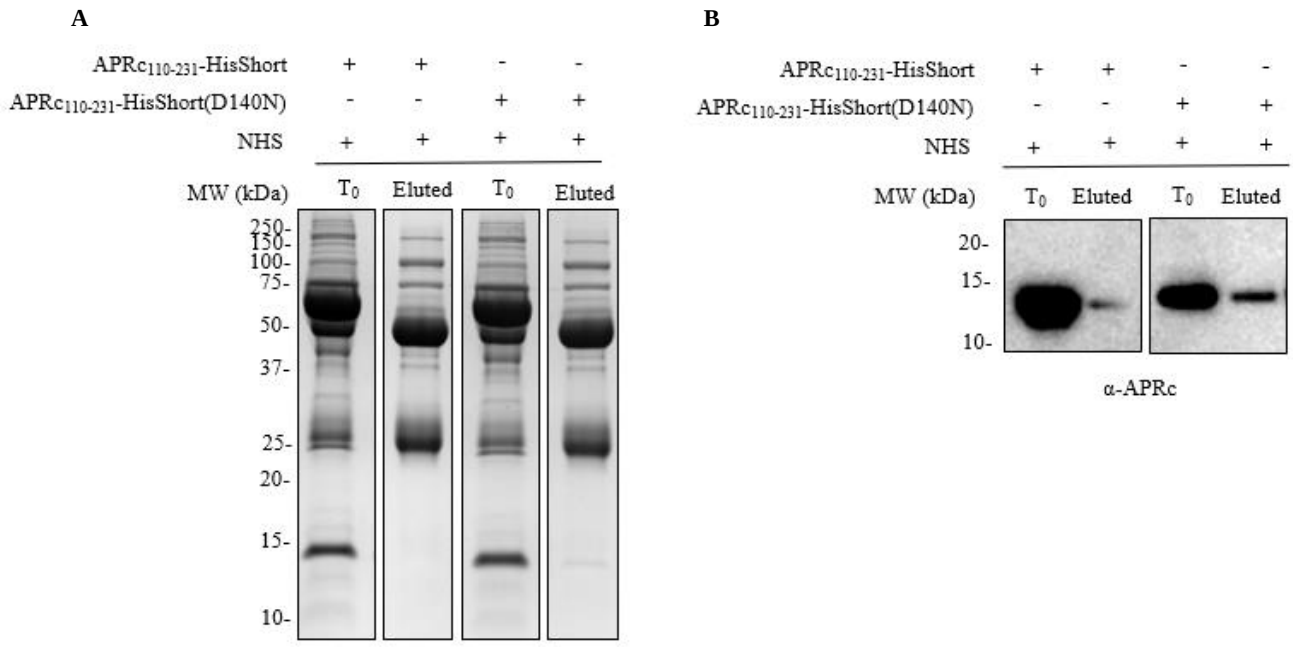


Figure 3.21- Evaluation of non-immune IgG-APRc binding in normal human serum by immunoprecipitation assay with protein A Mag Sepharose. The recombinant forms of APRc, APRc$_{110-231}$-HisShort and APRc$_{110-231}$-HisShort(D140N) were individually incubated with NHS for 4h at 37ºC, followed by incubation with protein A Mag Sepharose for 35 min at room temperature. Magnetic resins were then washed with TBS buffer containing 250 mM NaCl and TBS buffer containing 1 % Octyl-β-Glucoside. Protein elution was carried out by incubation and denaturation of the magnetic resins with loading buffer with DTT diluted 6x in PBS. **(A)** SDS-PAGE analysis followed by Coomassie blue staining of the eluted fractions (eluted) together with input samples (T_0). 10 µL of each sample were applied in a 12.5% polyacrylamide gel **(B)** Western-blot analysis of the eluted fractions (eluted) together with input samples (T_0). 3 µL of each sample were applied in a 10% polyacrylamide gel, and then evaluated by Western-blot analysis with the polyclonal rabbit antibody anti-APRc (raised towards the sequence Cys-Tyr-Thr-Arg-Thr-Tyr-Leu-Thr-Ala-Asn-Gly-Glu-Asn-Lys-Ala), produced by GenScript (1:500 in TBS-T 2% BSA) followed by the monoclonal mouse anti-rabbit IgG (M205) [HRP] (A01827-200, GenScript) (1:20 000 in TBS-T 2% BSA).

Evaluation of APRc Ig-cleavage activity in serum samples

In parallel with the pull-down assays with NHS from an adult donor (previous section), a pull-down assay with APRc (wild-type and mutant form) using human normal umbilical cord serum was also performed to compare differences in the migration pattern of eluted proteins. Samples of serum before incubation, samples corresponding to unbound fractions collected after sedimentation of the His Mag Sepharose Ni Beads (not-linked fractions), and eluted fractions were analysed by SDS-PAGE followed by Coomassie blue staining [Figure 3.22 (A)] and by Western blot with mouse anti-Human IgG Fc [HRP] antibody [Figure 3.22 (B)].

From the Western-blot analysis with mouse anti-Human IgG Fc [HRP] antibody, it was possible to detect IgG signal in the eluted fractions when using both APRc wild-type or the catalytic mutant. These results further corroborate IgG binding to the immobilized APRc during its incubation with serum [Figure 3.22 (B)] . Moreover, this provides additional confirmation of APRc-IgG binding in human serum samples from different sources/compositions.

Interestingly, from the SDS-PAGE analysis [Figure 3.22 (A)], it was possible to observe apparent differences in the profiles of the eluted proteins between incubations with the wild-type and the mutant form of APRc. Namely, in the region between 75 kDa and 50 kDa, and above 15 kDa. These results suggest that APRc might cleave some serum components. Supported by this evidence in complex samples, the potential cleavage of IgG by APRc was again evaluated in this context.

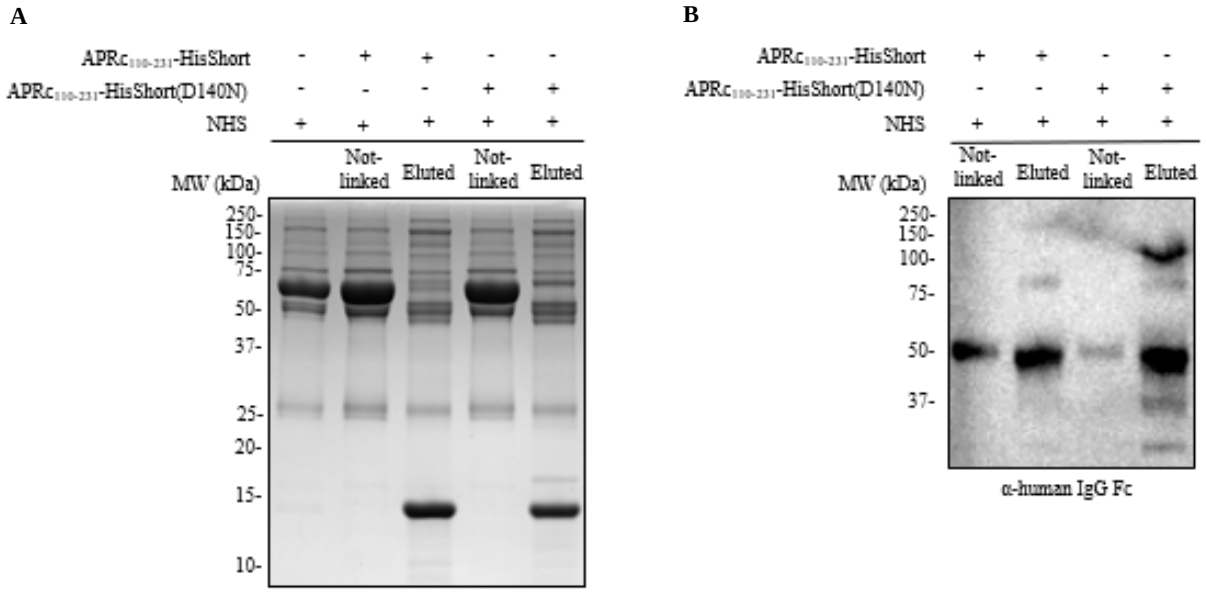


Figure 3.22- Evaluation of non-immune Ig-APRc binding in human normal umbilical cord serum by pull-down assay with His Mag Sepharose Ni Beads. The recombinant forms of APRc, APRc$_{110-231}$-HisShort and APRc$_{110-231}$-HisShort(D140N) were individually incubated with His Mag Sepharose Ni Beads for 1 hour and 30 minutes, at room temperature followed by incubation with normal umbilical cord serum for 4 hours at 37º C. Magnetic beads were then washed with PBS buffer containing 200 mM NaCl and PBS containing 0.1 % Tween. Protein elution was carried out by incubation and denaturation of the magnetic beads with loading buffer with DTT diluted 6x in PBS. **(A)** SDS-PAGE analysis followed by Coomassie blue staining from the samples of normal umbilical cord human serum (before incubation), the samples collected after sedimentation of the His Mag Sepharose Ni Beads (not-linked), and the samples of the eluted fractions (eluted). 10 µL of each sample were applied in a 12.5% polyacrylamide gel **(B)** Western-blot analysis of samples collected after sedimentation of the His Mag Sepharose Ni Beads (not-linked), and the eluted samples (eluted). 3 µL of each sample were applied in a 10% polyacrylamide gel, and then evaluated by Western-blot analysis with the monoclonal mouse anti-Human IgG Fc [HRP] (A01854-200, GenScript) (1:20 000 in TBS-T 2% BSA).

To further evaluate APRc IgG-cleavage activity in normal umbilical cord human serum, depletion of human serum albumin (HSA) was initially performed. The goal was to remove this highly abundant protein that could, otherwise, mask the analysis of the protein profiles between 75 and 50 kDa (as anticipated by the profile of not-linked fractions in Figure 3.22 (A)). To this end, human normal umbilical cord serum diluted in digestion buffer (50 mM sodium phosphate pH 6.5 buffer, containing 100 mM NaCl) was incubated with Cibacron blue agarose. After incubation, the supernatant was collected [normal umbilical cord human serum depleted from human serum albumin (NHSΔHSA)]. Samples before (NHS) and after depletion (NHSΔHSA) were analysed by SDS-PAGE followed by Coomassie blue staining [Figure 3.23 (A)]. The NHSΔHSA was then individually incubated with recombinant $APRc_{110-231}$-HisShort and $APRc_{110-231}$-HisShort(D140N) in digestion buffer. As a control, the NHSΔHSA was incubated with APRc's dilution buffer (20 mM HEPES buffer pH 7.4 containing 100 mM NaCl). The samples were denatured with 6x loading buffer with DTT and analysed by SDS-PAGE followed by Coomassie blue staining [Figure 3.23 (B)] and evaluated by Western-blot analysis with the mouse anti-Human IgG Fc [HRP] antibody [Figure 3.23 (C)].

From the SDS-PAGE analysis of serum samples before and after Cibacron blue incubation, it was possible to conclude that the depletion of HSA was only partial [Figure 3.23 (A)]. However, the incubation with APRc was still performed. From the SDS-PAGE analysis of the digestion samples [Figure 3.23 (B)], no significant differences in protein profiles were observed between the incubations with the wild-type and the mutant forms of APRc. Furthermore, these samples were analysed by Western blot with anti-human IgG Fc [HRP] antibody [Figure 3.23 (C)], and the differences in signal could not be related to APRc activity. Therefore, consistent with the previous observations with isolated IgG, no apparent cleavage of IgG by APRc was observed in the context of normal human serum.

In the future, the depletion of HSA from serum samples should be improved, and the digestion assays optimized to gain additional insights on the nature of the other serum proteins that appear to be cleaved by APRc, as suggested by the pull-down assays. Although preliminary, these results suggest that APRc function in serum samples may go beyond IgG binding.

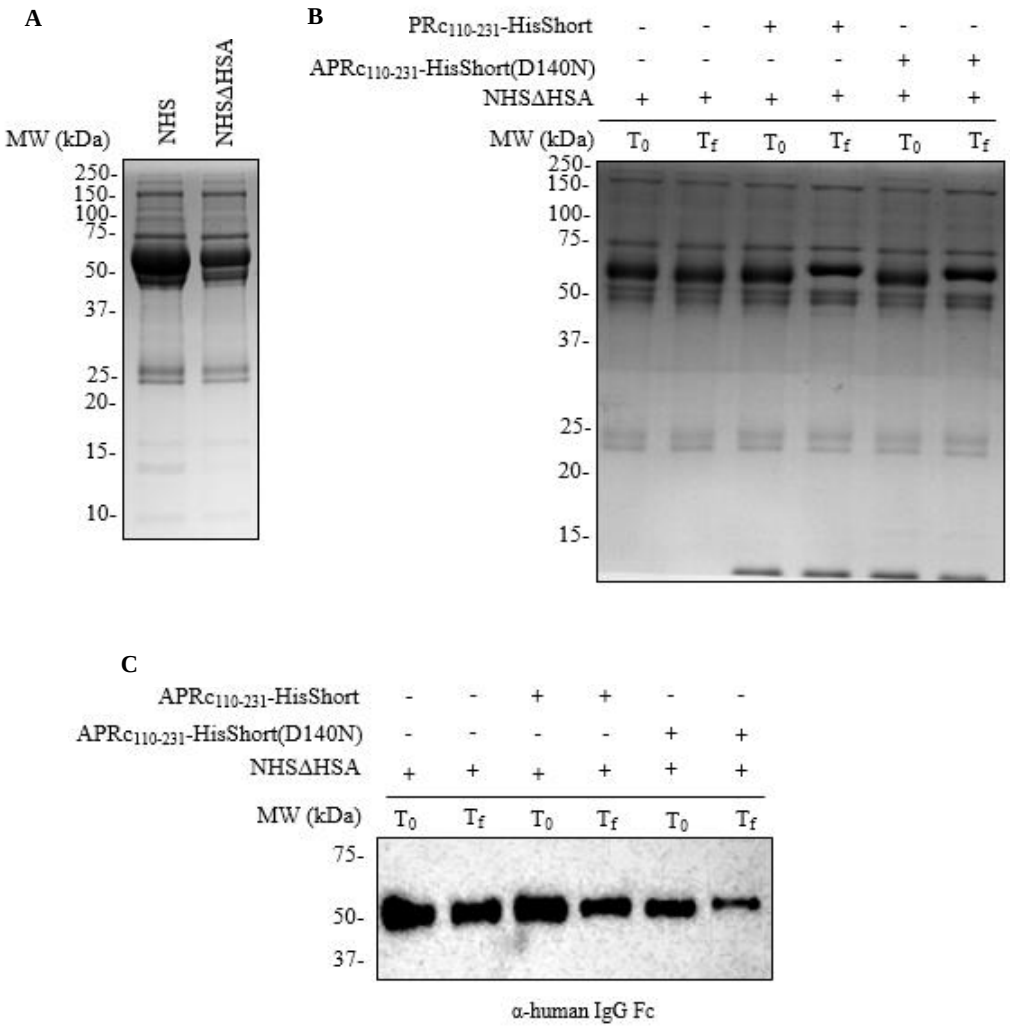


Figure 3.23- Evaluation of APRc-IgG cleavage activity in human normal umbilical cord serum. (A) Depletion of HSA from normal umbilical cord serum. Human normal umbilical cord serum diluted in digestion buffer (50 mM sodium phosphate pH 6.5 buffer, containing 100 mM NaCl) was incubated with Cibacron blue agarose for 35 minutes at room temperature. After incubation, the suspension was sedimented by centrifugation and the supernatant was collected [normal umbilical cord human serum depleted from human serum albumin (NHSΔHSA)]. Samples before (NHS) and after (NHSΔHSA) cibacron blue treatment were analysed by SDS-PAGE followed by Coomassie blue staining. 10 µL of each sample were applied in a 12.5% polyacrylamide gel. **(B)** Digestion reaction of NHSΔHSA with APRc. NHSΔHSA was incubated with APRc for 5 hours at 37ºC in digestion buffer. As a control, the NHSΔHSA was incubated with APRc's dilution buffer (20 mM HEPES buffer pH 7.4 containing 100 mM NaCl). Samples from the beginning (T_0) and the end (T_f) of the digestion were collected, denatured with 6x loading buffer with DTT and then analysed by SDS-PAGE followed by Coomassie blue staining. 10 µL of each sample were applied in a 12.5% polyacrylamide gel. **(C)** Western blot analysis of human IgG in digestion samples of NHSΔHSA with APRc. 3 µL of each sample were applied in a 10% polyacrylamide gel, and then evaluated by Western-blot analysis with the monoclonal mouse anti-Human IgG Fc [HRP] (A01854-200, GenScript) (1:20 000 in TBS-T 2% BSA).

3.2.5 Evaluation of non-immune Ig-binding at the surface of *Rickettsia* species

APRc is expressed in *Rickettsia* and present at the outer membrane of these bacteria, with the catalytic domain extracellular oriented[67]. Moreover, when expressed in *E. coli* a similar localization has been reported for the full-length APRc[67]. Therefore, assuming the presence of APRc at the surface of *Rickettsia*, we next sought to evaluate the non-immune Ig-binding in intact *Rickettsia*. To this end, PFA-treated *R. africae* was used as a representative species of SFG *Rickettsia*. The PFA-fixed *Rickettsia* were independently incubated with human IgG, NHS, or PBS buffer, for 2 hours at 37ºC. After washing with PBS, the samples were treated with PBS buffer containing 1 M NaCl to elute interacting proteins, and the remaining bacteria were removed by centrifugation. The cell-free eluted fractions were analysed by Western-blot with anti-Human IgG Fc [HRP] antibody (Figure 3.24).

The Western-blot analysis revealed binding of human IgG to intact (PFA-treated) *R. africae*, both when using isolated IgG and in the context of NHS (Figure 3.24). This suggests the presence of protein(s) at the surface of Rickettsia capable of non-immune Ig-binding. With the outer membrane localization of APRc and the above characterization of its Ig-binding activity, we anticipate APRc as one of those proteins.

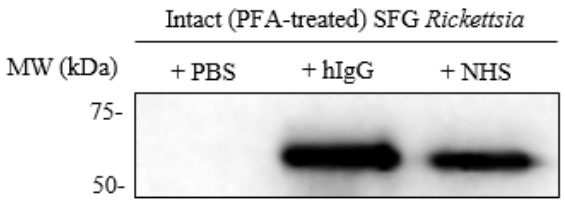


Figure 3.24- Evaluation on non-immune IgG-binding at the surface of *Rickettsia*. Intact (PFA-treated) *R. africae* (6.45 x 10^8 PFU) were independently incubated with 275 µg of human IgG (#I2511, Sigma-Aldrich), 50 µL of NHS diluted 2x in PBS buffer (contains approximately 275 µg of IgG), or 50 µL of PBS buffer, for 2 hours at 37ºC. The samples were treated with PBS buffer containing 1 M NaCl and the remaining bacteria were removed by centrifugation. 35 µL of each cell-free elution fraction were denatured with 6x loading buffer with DTT, applied in a 12.5% polyacrylamide gel, and then evaluated by Western-blot analysis with the monoclonal mouse anti-Human IgG Fc [HRP] (A01854-200, GenScript) (1:20 000 in TBS-T 2% BSA).

4 Discussion and Conclusions

Many pathogens have established different tactics to evade the host's immune responses and establish infection[23]. *Rickettsia* has been reported to be resistant to the serum bactericidal effects and to be able to evade complement-mediated killing[21], suggesting that rickettsial species have developed different mechanisms to inhibit recognition by host serum components. To date, at least three rickettsial surface proteins have been identified as mediators of these evasion mechanisms. The rickettsial autotransporter protein (rOmpB), which specifically interacts with factor H (a soluble host complement inhibitor)[20], and the rickettsial outer membrane proteins Adr1 and Adr2, that interact with the terminal complement complex inhibitor vitronectin (Vn) were shown to, individually, mediate partial survival of *Rickettsia* in human serum[18,19,22]. Interestingly, many pathogenic bacteria have also been shown to use non-immune Ig-binding proteins to avoid recognition by innate immune serum components[23,24]. These proteins (*e.g.*, the staphylococcal proteins A[39] and Sbi[44], the group A streptococci M-proteins[49,50], the group A and B streptococci ARP, Sir and β proteins[29,31,32,33], the protein L from *F. magna*[38], the MID protein from *B. catarrhalis*[34] and the protein D from *H. influenzae*[36]) have been implicated in different functions related to pathogenic evasion mechanisms, such as inhibition of opsonophagocytic killing, complement killing, B-cell depletion, mast cell activation, activation of B-cells, histamine release, and degranulation of mast cells[25,30,32,33,35,37,43,45,46,49,51,53]. Indeed, Ig-binding proteins are important virulence factors in pathogenic bacteria and have been identified in both Gram-positive (the majority) and Gram-negative bacteria[24,25]. With this work, we provide evidence for a novel Ig-binding protein from *Rickettsia*, anticipating additional immune-evasion tactics for this obligate pathogen.

We demonstrate that APRc, the rickettsial retropepsin, can bind to immunoglobulins from different species and different classes. APRc shows interaction with rabbit IgG, human IgG, and mouse IgG, with an apparent higher interaction to rabbit IgG. Furthermore, APRc also shows interaction with different classes of human immunoglobulins (IgG, IgM, and IgA), showing a higher binding activity to the IgG class. Compared with other non-immune Ig-binding proteins, APRc shares this capacity to interact with immunoglobulins from different species. The Type I of IgG Fc-binding proteins from Gram-positive bacteria (*e.g.*, staphylococcal protein A), Type III (*e.g.*, streptococcal protein G), Type V, and Type VI present higher reactivities to human and rabbit IgG, compared to mouse IgG, as does APRc. However, Type II (*e.g.*, M-proteins) and type IV show reactivity towards human and rabbit IgG, but not towards mouse IgG[26,27]. Moreover, APRc's ability to interact with immunoglobulins from different classes has also been demonstrated for several Ig-binding proteins. In Gram-positive bacteria, the staphylococcal protein A, an IgG Fc-binding protein that can also interact through the Fab fragment, interacts with IgG, IgM, IgA, and IgE[24,42] and the protein L from *F. magna*, an Ig-binding protein that interacts with the κ-type light chains of IgG shows reactivity towards all classes of immunoglobulins[38]. In Gram-negative bacteria, the Fab-fragment Ig-binding bacteria from *H. pylori* (Hsp60) shows reactivity towards human IgG subclasses IgG1 and IgG3, and IgM. The Ig-binding protein from *H. somni* presents reactivity towards bovine IgG2, IgA, and IgM[24,55], and the Ig-binding protein from *S.*

maltophilia interacts with the four subclasses of human IgG and with rabbit IgG, and also with mouse IgG and IgA[24,59]. Since the Ig-binding proteins present a great variety in reactivity with immunoglobulins, in the future, it would also be important to evaluate the interaction of APRc with immunoglobulins from other species, with other classes of immunoglobulins, and also with different subclasses of IgG.

When comparing the different Ig classes, it is known that they differ in the nature of their heavy chains, which are related to their different functions and involvement in different types of immune responses, with most of these differences primarily found in the Fc-fragment[72]. As described above, different IgG Fab-binding proteins generally show reactivity to other Ig classes. Our cross-linking studies showed that APRc binds to $F(ab')_2$ as well as to the monovalent fragment F(ab') of human IgG. Although further studies are required to validate the interaction through this region for the other immunoglobulin classes, we hypothesize that APRc might be interacting through the light chain. There are only two main types of light chains, which would explain the observed binding to the different classes of immunoglobulins in our ELISA assays.

In several experiments, under different conditions, the interaction of APRc with IgG promoted the oligomeric stabilization of APRc, but no apparent proteolytic activity towards IgG was observed. This is not totally unexpected, as other proteins bind to immunoglobulins but do not cleave them, such as the staphylococcal proteins A[39] and Sbi[44], the group A streptococci M-proteins[49,50], the group A and B streptococci ARP, Sir, and β proteins[29,31,32,33], the protein L from *F. magna*[38], the MID protein from *B. catarrhalis*[34], and the protein D from *H. influenzae*[36], among others[25]. Therefore, APRc appears to be integrated into this group. In contrast, there are some Ig-binding proteins that present Ig-cleavage activity, for example IgA_1 proteases from *Neisseria meningitidis, Streptococcus pneumoniae and H. influenzae*, and the IgG protease from *S. pyogenes*[25,73].

When mapping the region of APRc involved in Ig-binding, our results suggest that this interaction occurs through a very distinct wide loop of APRc (comprising residues 157-166) that presents the most distinct conformation from all the other structures of retropepsins[68]. However, since we could not identify a truncated form of APRc for which binding was completely abolished, this suggests that the interaction depends on more than one region in the protease. For most Ig-binding proteins, for which the interaction domains have been identified, it has been reported that these domains comprise polypeptide repeats of one of several types that can vary in number and that the interaction can depend on more than one region for binding[24]. For example, the staphylococcal protein A presents five highly homologous IgG-binding domains that are highly hydrophilic and resistant to proteases, lack cysteine residues, and are constituted by tightly packed antiparallel α-helices, stabilized by hydrophobic interactions between the α-helices. The interaction between SpA and IgG occurs through the residues of

the second and first α-helices of the Ig-binding domain[24,39]. The streptococci protein G presents an Ig-binding domain composed of two antiparallel and two parallel β-strands with the α-helix situated along its diagonal. The interaction between SpG and IgG occurs through the C-terminal part of the α-helix, its N-terminal part of the third β-strand, and the loop region connecting these two structural elements[24,52]. Moreover, the protein L from *F. magna* presents an Ig-binding domain with four or five highly homologous repeats and composed of a β-sheet formed by four β-strands and a central α-helix. PpL also has two independent binding sites (with different affinities)[24,38,42]. Therefore, it is not unexpected that we could not fully map the interacting region in APRc. Interestingly, since APRc has no sequence or structural similarity with other reported Ig-binding proteins/domains, our results anticipate a new binding interface that requires additional structural characterization for a deeper characterization (either to understand its functional relevance in the context of rickettsial infection or as an alternative reagent for immunoglobulin purification).

This non-immune Ig-binding activity of APRc was confirmed with purified immunoglobulins, as well as in serum samples. As for the purified IgG, no apparent IgG-cleavage activity was observed in these more complex samples. However, our preliminary results point towards APRc cleavage of other serum components, clearly suggesting that APRc may interfere with the function of other serum proteins. In fact, many known Ig-binding proteins can also interact with other proteins of the host's blood serum, sometimes in order to evade complement-mediated killing more effectively[24]. For example, the IgA Fc-binding proteins from *S. pyogenes*, protein ARP and protein Sir (that also shows binding activity towards IgG) interact with the complement regulator C4BP, inhibiting complement at the bacterial surface[30] and the staphylococcal IgG-binding protein Sbi also interacts with other serum proteins, such as the complement component C3 and the complement regulator Factor H contributing to immune evasion by stimulating inhibition of opsonization and complement inhibition[45,46].

Moreover, we have demonstrated Ig-binding at the surface of *R. africae*, confirming the existence of Ig-binding proteins in *Rickettsia*. Given the localization of APRc in the outer membrane of *Rickettsia*, we hypothesize that APRc might be one of the proteins contributing to this activity. Although further studies are required to assess the functional significance of this non-immune Ig-binding activity, one can speculate that it could lead to the protection of *Rickettsia* from complement-mediated killing (serum bactericidal activity or opsonization/phagocytosis) by inhibiting/reducing deposition of complement factors. Therefore, we hypothesize that APRc may act as a novel evasin contributing to rickettsial immune evasion toolbox. Indeed, preliminary data from the laboratory show that *E.coli*-expressing APRc at the surface protects *E. coli* from complement-mediated killing (serum bactericidal activity). However, further characterization of this mechanism is still necessary. This characterization will provide essential information for the future development of new therapeutics for rickettsial diseases.

5 References

1. Ricketts, H. T. A micro-organism which apparently has a specific relationship to rocky mountain spotted fever: a preliminary report. *JAMA* **52**, 379–380 (1909).
2. Hurst, L. Chapter 2: Morphology and Ultrastructure of Bacteria. in *Bacteriology* 82–83 (ED-Tech Press, 2019).
3. Giménez, D. F. Staining Rickettsiae in Yolk-Sac Cultures. *Stain Technol.* **3**, 135–140 (1964).
4. Diop, A., Karkouri, K. El, Raoult, D. & Fournier, P. E. Genome sequence- - based criteria for demarcation and definition of species in the genus *Rickettsia*. *Int J Syst Evol Microbiol* **70**, 1738–1750 (2020).
5. Gillespie, J. J. *et al.* *Rickettsia* Phylogenomics : Unwinding the Intricacies of Obligate Intracellular Life. *PLoS One* **3**, 21–27 (2008).
6. Merhej, V. & Raoult, D. Rickettsial evolution in the light of comparative genomics. *Biol Rev Camb Philos Soc* **86**, 379–405 (2011).
7. Abdad, M. Y., Abdallah, R. A., Fournier, P.-E., Stenos, J. & Vasoo, S. A Concise Review of the Epidemiology and Diagnostics of Rickettsioses: *Rickettsia* and *Orientia spp*. *J Clin Microbiol* **56**, 1–10 (2018).
8. Jones, K. E. *et al.* Global trends in emerging infectious diseases. *Nat. Publ. Gr.* **451**, 990–994 (2008).
9. Blanton, L. S. The Rickettsioses: A Practical Update. in *Infectious Disease Clinics of North America* **33**, 213–229 (2019).
10. Walker, D. H. & Ismail, N. Emerging and re-emerging rickettsioses : endothelial cell infection and early disease events. *Nat Rev Microbiol* **6**, 375–386 (2008).
11. Morand, S., Chaisiri, K., Kritiyakan, A. & Kumlert, R. Disease Ecology of Rickettsial Species : A Data Science Approach. *Trop. Med. Infect. Dis* **5**, (2020).
12. Tomassone, L., Portillo, A., Nováková, M., Sousa, R. De & Oteo, J. A. Neglected aspects of tick-borne rickettsioses. *Parasit. Vectors* **11**, (2018).
13. Osterloh, A. The neglected challenge : Vaccination against rickettsiae. *PLoS Negl Trop Dis* **14**, 1–35 (2020).
14. Walker, D. H., Valbuena, G. A. & Olano, J. P. Pathogenic Mechanisms of Diseases Caused by *Rickettsia*. *New York Acad. Sci.* **990**, 1–11 (2003).
15. Ceraul, S. M. Chapter 13: Transmission and the Determinants of Transmission Efficiency. in *Intracellular Pathogens II: Rickettsiales* (eds. Palmer, G. H. & Azad, A. F.) 391–415 (ASM Press, 2012).
16. Sahni, S. K., Narra, H. P., Sahni, A. & Walker, D. H. Recent molecular insights into rickettsial pathogenesis and immunity. *Future Microbiol.* **8**, 1265–1288 (2013).
17. Curto, P., Simões, I., Riley, S. P. & Martinez, J. J. Differences in Intracellular Fate of Two Spotted Fever Group *Rickettsia* in Macrophage-Like Cells. *Front. Cell. Infect. Microbiol.* **6**, 1–14 (2016).
18. Riley, S. P., Patterson, J. L., Nava, S. & Martinez, J. J. Pathogenic *Rickettsia* species acquire vitronectin from human serum to promote resistance to complement-mediated killing. *Cell. Microbiol.* **16**, 849–861 (2014).
19. Fish, A. I., Riley, S. P., Singh, B., Riesbeck, K. & Martinez, J. J. The *Rickettsia conorii* Adr1 Interacts with the C-Terminus of Human Vitronectin in a Salt-Sensitive Manner. *Front. Cell. Infect. Microbiol.* **7**, 1–12 (2017).
20. Riley, S. P., Patterson, J. L. & Martinez, J. J. The Rickettsial OmpB beta-Peptide of *Rickettsia conorii* Is Sufficient To Facilitate Factor H-Mediated Serum Resistance. *Infect. Immun.* **80**, 2735–2743 (2012).
21. Chan, Y. G. *et al.* Molecular Basis of Immunity to Rickettsial Infection Conferred through Outer Membrane Protein B □ †. *Infect. Immun.* **79**, 2303–2313 (2011).
22. Garza, D. A., Riley, S. P. & Martinez, J. J. Expression of *Rickettsia* Adr2 protein in *E . coli* is sufficient to promote resistance to complement-mediated killing , but not adherence to mammalian cells. *PLoS One* **12**, 1–15 (2017).
23. Sarantis, H. & Grinstein, S. Subversion of Phagocytosis for Pathogen Survival. *Cell Host Microbe* **12**, 419–431 (2012).
24. Sidorin, E. V. & Solov'eva, T. F. IgG-Binding Proteins of Bacteria. *Biochem.* **76**, 295–308 (2011).

25. Collin, M. & Kilian, M. Chapter 18. Bacterial Modulation of Fc Effector Functions. in *Antibody Fc: Linking Adaptive and Innate Immunity* 317–332 (Elsevier Inc., 2014).

26. Boyle, M. D. P. Chapter 1: Introduction to Bacterial Immunoglobulin–Binding Proteins. in *Bacterial Immunoglobulin–Binding Proteins: Applications in Immunotechnology* **2**, 1–21 (Academic Press, 1990).

27. Boyle, M. D. P. *Bacterial Immunoglobulin-Binding Proteins Volume 1*. (Academic Press, 1989).

28. Boyle, M. D. Bacterial Immunoglobulin-Binding Proteins. in *Encyclopedia of Immunology* (ed. Delves, P. J.) 323–327 (Academic Press, 1998).

29. Lindahl, G. & Akerstrom, B. Receptor for IgA in group A streptococci : cloning of the gene and characterization of the protein expressed in *Escherichia coli*. *Mol. Immunol.* **3**, 239–247 (1989).

30. Thern, A., Stenberg, L., Dahlbäck, B. & Lindahl, G. Ig-binding surface proteins of *Streptococcus pyogenes* also bind human C4b-binding protein (C4BP), a regulatory component of the complement system. *J Immunol* **154**, 375–386 (1995).

31. Stenberg, L., O'Toole, P., Mestecky, J. & Lindahl, G. Molecular characterization of protein Sir, a streptococcal cell surface protein that binds both immunoglobulin A and immunoglobulin G. *J. Biol. Chem.* **269**, 13458–64 (1994).

32. Jerlstrom, P. G., Talay, S. R., Valentin-Weigand, P., Timmis, K. N. & Chhatwal, G. S. Identification of an Immunoglobulin A Binding Motif Located in the beta-Antigen of the c Protein Complex of Group B Streptococci. *Infect. Immun.* **64**, 2787–2793 (1996).

33. Russell-Jones, G. J., Gotschlich, E. C. & Blake, M. S. A surface receptor specific for human IgA on Group B streptococci possessing the Ibc protein antigen. *J. Exp. Med.* **160**, 1467–1475 (1984).

34. Forsgren, A. *et al.* Isolation and Characterization of a Novel IgD-Binding Protein from *Moraxella catarrhalis*. *J. Immunol.* **167**, 2112–2120 (2001).

35. Wingren, A. G. *et al.* The Novel IgD Binding Protein from *Moraxella catarrhalis* Induces Human B Lymphocyte Activation and Ig Secretion in the Presence of Th2 Cytokines. *J. Immunol.* **168**, 5582–5588 (2002).

36. Ruan, M. R., Akkoyunlu, M., Grubb, A. & Forsgren, A. Protein D of *Haemophilus influenzae*. A novel bacterial surface protein with affinity for human IgD. *J Immunol* **145**, 3379–84 (1990).

37. Akkoyunlu, M., Ruan, M. & Forsgren, A. Distribution of Protein D , an Immunoglobulin D-Binding Protein, in *Haemophilus Strains*. *Infect. Immun.* **59**, 1231–1238 (1991).

38. Akerstroms, B. & Bjorck, L. Protein L : An Immunoglobulin Light Chain-binding Bacterial Protein. *J. Biol. Chem.* **264**, 19740–19746 (1989).

39. Forsgren, A. & Sjöquist, J. 'Protein A' from *S. aureus.* I. Pseudo-immune reaction with human gamma-globulin. *J Immunol* **97**, 822–7 (1966).

40. Lin, I., Reyn, A. & Birch-Anderson, A. Electron microscopy of staphylococcal protein a reactivity and specific antigen-antibody reactions. *Acta Path. Microbiol. Scand.* **80**, 281–291 (1972).

41. Cheung, A. L., Bayer, A. S., Peters, J. & Ward, J. I. Analysis by Gel Electrophoresis, Western Blot, and Peptide Mapping of Protein A Heterogeneity in *Staphylococcus aureus* Strains. *Infect. Immun.* **55**, 843–847 (1987).

42. Choe, W., Durgannavar, T. A. & Chung, S. J. Fc-Binding Ligands of Immunoglobulin G: An Overview of High Affinity Proteins and Peptides. *Materials (Basel).* **9**, (2016).

43. Kim, H. K., Thammavongsa, V., Schneewind, O. & Missiakas, D. Recurrent infections and immune evasion strategies of *Staphylococcus aureus*. *Curr. Opin. Microbiol.* **15**, 92–99 (2012).

44. Zhang, L., Jacobson, K., Vasi, J., Lindberg, M. & Frykberg, L. A second IgG-binding protein in *Staphylococcus aureus*. *Microbiology* **408**, 985–991 (1995).

45. Smith, E. J., Visai, L., Kerrigan, S. W., Speziale, P. & Foster, T. J. The Sbi Protein Is a Multifunctional Immune Evasion Factor of *Staphylococcus aureus*. *Infect. Immun.* **79**, 3801–3809 (2011).

46. Haupt, K. *et al.* The *Staphylococcus aureus* Protein Sbi Acts as a Complement Inhibitor and Forms a Tripartite Complex with Host Complement Factor H and C3b. *PLoS Pathog.* **4**, (2008).

47. Clark, E., Upadhyay, A., Bagby, S. & van den Elsen, J. IsaB, a new immunoglobulin-binding protein from *Staphylococcus aureus*. *Mol. Immunol.* **46**, 2834–2835 (2009).

48. Itoh, S., Hamada, E., Kamoshida, G., Yokoyama, R. & Takii, T. Staphylococcal

superantigen-like protein 10 (SSL10) binds to human immunoglobulin G (IgG) and inhibits complement activation via the classical pathway. *Mol. Immunol.* **47**, 932–938 (2010).

49. Fischetti, V. A. Streptococcal M Protein : Molecular Design and Biological Behavior. *Am Soc Microbiol* **2**, 285–314 (1989).

50. Frick, I., Akesson, P. & Cooney, J. Protein H — a surface protein of *Streptococcus pyogenes* with separate binding sites for IgG and albumin. *Mol. Microbiol.* **12**, 143–151 (1994).

51. Berge, A., Kihlberg, B., Sjo, A. G. & Bjo, L. Streptococcal Protein H Forms Soluble Complement-activating Complexes with IgG , but Inhibits Complement Activation by IgG-coated Targets. *J. Biol. Chem.* **272**, 20774–20781 (1997).

52. Katol, K. *et al.* Model for the complex between protein G and an antibody Fc fragment in solution. *Structure* **3**, 79–85 (1995).

53. Genovese, A. *et al.* Immunoglobulin Superantigen Protein L Induces IL-4 and IL-13 Secretion from Human FcεRI+ Cells Through Interaction with the κ Light Chains of IgE. *J Immunol* **170**, 1854–1861 (2003).

54. Leo, J. C. & Goldman, A. The immunoglobulin-binding Eib proteins from *Escherichia coli* are receptors for IgG Fc. *Mol. Immunol.* **46**, 1860–1866 (2009).

55. Corbeil, L. B., Bastida-Corcuera, F. L. D. & Beveridge, T. J. *Haemophilus somnus* Immunoglobulin Binding Proteins and Surface Fibrils. *Infect. Immun.* **65**, 4250–4257 (1997).

56. Zav'yalov, V. P. *et al.* Zav'yalov, V. P., Abramov, V. pH6 antigen (PsaA protein) of *Yersinia pestis*, a novel bacterial Fc-receptor. *FEMS Immunol. Med. Microbiol.* **14**, 53–57 (1996).

57. Sidorin, E. V *et al.* Isolation and Characterization of a Low Molecular Weight Immunoglobulin Binding Protein from *Yersinia pseudotuberculosis*. *Biochem.* **71**, 1278–1283 (2006).

58. Sidorin, E. V. *et al.* Chaperone Skp from *Yersinia pseudotuberculosis* Exhibits Immunoglobulin G Binding Ability. *Biochem.* **74**, 406–415 (2009).

59. Grover, S., McGee, Z. A. & Odell, W. D. Isolation of a 30 kDa immunoglobulin binding protein from *Pseudomonas maltophilia*. *J. Immunol. Methods* **141**, 187–197 (1991).

60. Amini, H.-R., Ascencio, F., Cruz-Villacorta, A., Ruiz-Bustos, E. & Wadstriim, To. Immunochemical properties of a 60 kDa cell surface-associated heat shock-protein (Hsp60) from *Helicobacter pylori*. *FEMS Immunol. Med. Microbiol.* **16**, 163–172 (1996).

61. Rawlings, N. D. & Barrett, A. J. Chapter 1-Introduction: Aspartic and Glutamic Peptidases and Their Clans. in *Handbook of Proteolytic Enzymes* (eds. Neil, D. & Salvesen, R.) 3–19 (Academic Press, 2013).

62. Rawlings, N. D. *et al.* The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* **46**, D624–D632 (2018).

63. Wlodawer, A. & Gustchina, A. Structural and biochemical studies of retroviral proteases. *Biochim. Biophys. Acta.* **1477**, 16–34 (2000).

64. Konvalinka, J., Kräusslich, H. & Müller, B. Retroviral proteases and their roles in virion maturation. *Virology* 1–15 (2015).

65. Dunn, B. M., Goodenow, M. M., Gustchina, A. & Wlodawer, A. Retroviral proteases. *Genome Biol.* **3**, pp.reviews3006.1– reviews3006.7 (2002).

66. Lv, Z., Chu, Y. & Wang, Y. HIV protease inhibitors : a review of molecular selectivity and toxicity. *HIV AIDS* **7**, 95–104 (2015).

67. Cruz, R., Huesgen, P., Riley, S. P., Wlodawer, A. & Faro, C. RC1339/APRc from *Rickettsia conorii* Is a Novel Aspartic Protease with Properties of Retropepsin-Like Enzymes. *PLoS Pathog.* **10**, (2014).

68. Li, M. *et al.* Structure of RC1339/APRc from *Rickettsia conorii* , a retropepsin-like aspartic protease. *Acta Crystallogr D Biol Crystallogr* **71**, 2109–2118 (2015).

69. Bechah, Y. *et al.* Genomic, proteomic, and transcriptomic analysis of virulent and avirulent *Rickettsia prowazekii* reveals its adaptive mutation capabilities. *Genome Res* **20**, 655–663 (2010).

70. Scholz, J., Besir, H., Strasser, C. & Suppmann, S. A new method to customize protein expression vectors for fast , efficient and background free parallel cloning. *BMC Biotechnol.* **13**, 12 (2013).

71. Leal, A. R. *et al.* Enzymatic properties , evidence for in vivo expression , and intracellular localization of shewasin D , the pepsin homolog from *Shewanella denitrificans*. *Nat. Publ. Gr.* 1–12

(2016).

72. Schroeder, H. W. & Cavacini, L. Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* **125**, S41–S52 (2010).

73. Pawel-Rammingen, U. von, P.Johansson, B. & Bjorck, L. IdeS , a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. *EMBO J.* **21**, 1607–1615 (2002).

6 Annexes

pET_APRc$_{110-231}$-HisShort

110-EVGEIIIARNRDGHFYINAFVNNVKIKFMV**D**TGASDIALTKEDAQKLGFDLTKLKYTRTYLTANGENKAAPITLNSVVIGKEFKNIKGHVGLGDLDISLLGMSLLERFKGFRIDKDLLILNY HHHHHH

pET_APRc$_{110-231}$-HisShort(D140N)

110-EVGEIIIARNRDGHFYINAFVNNVKIKFMV**N**TGASDIALTKEDAQKLGFDLTKLKYTRTYLTANGENKAAPITLNSVVIGKEFKNIKGHVGLGDLDISLLGMSLLERFKGFRIDKDLLILNY HHHHHH

Figure 6.1- Representation of the amino acid sequence of the constructs pET_APRc$_{110-231}$-HisShort and pET_APRc$_{110-231}$-HisShort(D140N). The construct pET_APRc$_{110-231}$-HisShort includes the coding sequence of APRc amino acids 110-231 fused with the C-terminal tag sequence HHHHHH. The construct pET_APRc$_{110-231}$-HisShort(D140N) is the active site mutant of pET_APRc$_{110-231}$-HisShort, where the catalytic aspartate was replaced by asparagine. These residues are highlighted in the sequences.