UNIVERSIDADE D
COIMBRA

Carolina Pagaimo Faim

PRODUCTION OF THE RECOMBINANT RETROPEPSIN-LIKE PROTEASE FROM *RICKETTSIA CONORII* (APRC)

Dissertation carried out at the Centre for Neuroscience and Cell Biology of the University of Coimbra in the context of the Integrated Master's in Chemical Engineering, oriented by Dr. Isaura Simões and Dr. Jorge F. B. Pereira and presented to the Department of Chemical Engineering, Faculty of Sciences and Technology of University of Coimbra

Coimbra, September 2023

Carolina Pagaimo Faim

PRODUCTION OF THE RECOMBINANT RETROPEPSIN-LIKE PROTEASE FROM *RICKETTSIA CONORII* (APRC)

Dissertation in the scope of the Integrated Master's in Chemical Engineering submitted to the Department of Chemical Engineering of the Faculty of Sciences and Technology of the University of Coimbra.

Supervisors:

Prof. Dr. Isaura Simões

Prof. Dr. Jorge F. B. Pereira

Coimbra, September 2023

UNIVERSIDADE D COIMBRA

Acknowledgements

First, I would like to express gratitude to my supervisor, Dr. Isaura Simões, for her support, patience, advice, and guidance, fundamental for the acquired knowledge during the realization of this project.

I also would like to extend my gratitude to investigator Pedro Curto, for all the help, advice and guidance provided, and for teaching me a lot about his work in the lab.

A special thanks to Dr. Jorge Pereira, for sharing his knowledge and orienting my work with his experience, patience, and wisdom.

To the people who have the patience to put up with me every day, I have a lot to thank them for all the support, friendship, and all the good and epic moments we spent. Their friendship, support and encouragement were essential.

Last, I would like to express my immense appreciation towards my family, specially to my mom, that throughout my life always supported me.

Resumo

Há uma preocupação crescente com as *rickettsioses* e o seu impacto na saúde global, sendo apontadas como agentes patogénicos emergentes/reemergentes. A protease aspártica da *Rickettsia conorii* (APRc) surgiu como um candidato alvo para a intervenção contra as *rickettsioses*, a qual pode ser produzida por microrganismos geneticamente modificados. Neste projeto, a APRc foi produzida utilizando a *E. coli (Escherichia coli)*. O desafio deste projeto objetivou melhorar os rendimentos de dímero *vs.* monómero, devido às dificuldades para obter o dímero canónico da APRc e prosseguir estudos sobre a relação estrutura-função da APRc, impedindo o desenho de inibidores específicos. Para ultrapassar estas limitações, este projeto estudou métodos de lise celular alternativos e modificações ao processo atualmente utilizado no processo de produção da APRc a partir de *E. coli*. Foram realizados ensaios de lise celular mecânicos, utilizando um homogeneizador de alta pressão (*HPH, high-pressure homegenizer)* e físico-químicos (aplicação de surfactantes, temperatura, ciclos de congelação-descongelação) para otimizar os rendimentos de produção da protease.

Inicialmente, avaliou-se a lise mecânica e, dependendo do ensaio realizado, as condições a que o *pellet* celular foi submetido foram modificadas. Primeiro replicou-se o método padrão para comparação com as modificações introduzidas nos ensaios seguintes, a fim de otimizar o método de lise celular utilizando o Emulsiflex^{R}-C3 ligado a uma pressão de 4 bar de azoto, com uma pressão sobre a amostra de 1700 bar, utilizando um volume mínimo de 30 mL e a passagem do *pellet* no mínimo de 3 ciclos de compressão. Em seguida, procedeu-se à otimização do processo padrão, variando o número de ciclos e a pressão aplicada na amostra. Em relação ao rendimento, o método padrão apresentou um rendimento de 0,45 de dímero, enquanto uma alteração a este método com a aplicação de uma pressão mínima na amostra e um aumento do número de ciclos, permitiu obter 1,74 da forma dimérica da protease.

Os métodos de rompimento celular físico-químicos avaliaram o efeito de diferentes ciclos (1 ou 4) de congelação/descongelação e temperaturas de descongelação (20 °C e 37 °C), com a adição de diferentes soluções tampão com pH 7.4, e adição de surfactantes (*SDS* – Dodecil sulfato de sódio). Nestes ensaios, o rendimento de dímero obtido foi de 0,54, apresentando uma ligeira melhoria de 17% relativamente ao método padrão. Os resultados mostraram ainda que a adição de surfactantes dificulta o processamento dos *pellets* celulares, isto é, de difícil ressolubilização. O aumento da temperatura levou a uma perda quase completa da proteína de interesse.

O desenvolvimento deste projeto demonstrou que o método mais eficaz de lise celular das células de *E. coli* é a lise mecânica utilizando o *HPH* aplicando uma pressão mínima na amostra - uma alteração ao método padrão previamente aplicado na obtenção da APRc. As alterações feitas ao procedimento padrão permitiram otimizar o processo de rompimento e aumentar o rendimento da forma dimérica de APRc de 0,45 para 1,74 (74%).

Palavras-chave: *Rickettsia* | APRc | lise celular | otimização | rendimento

Abstract

There is growing concern about rickettsioses and their impact on global health, being identified as emerging pathogens. The aspartic protease of *Rickettsia conorii* (APRc) has emerged as a candidate target for intervention against rickettsioses, which can be produced using genetically modified microorganisms. In this project, APRc will be produced using *E. coli*. The project challenge aimed to enhance the yields of dimer *vs.* monomer, due to the limitation for obtaining the canonical dimer of APRc and pursuing studies on the structure-function relationship of APRc, not allowing the design of specific inhibitors. To overcome these limitations, this project studied alternative cell lysis methods and changes to the method currently used in the APRc production process from *E. coli*. Mechanical cell lysis tests were carried out using a high-pressure homogenizer (HPH) and physicochemical tests (application of surfactants, temperature, freezing-thawing cycles) to optimize the protease production yields.

Initially, mechanical lysis was evaluated and, depending on the test carried out, the conditions to which the cell pellet was subjected were modified. First, the standard method was replicated for comparison with the modifications introduced in the following tests, to optimize the cell lysis method with Emulsiflex^{R}-C3 connected to a pressure of 4 bar of nitrogen, with a pressure on the sample of 1700 bar, and using a minimum volume of 30 mL and passing the pellet through at least 3 compression cycles. Next, the standard APRc production process was optimized, varying the number of cycles and the pressure applied to the sample. Concerning the recovery yields of dimer, the standard method yielded 0.45 of dimer, while a change to this method applying minimal pressure to the sample and increasing the number of cycles yielded 1.74 mg of the dimeric form of the protease.

The physicochemical cell disruption methods evaluated different freezing/thawing cycles (1 or 4) and thawing temperatures (20 °C and 37 °C), with the addition of different buffer solutions with pH 7.4, and the addition of surfactants (SDS - Sodium Dodecyl Sulfate). In these tests, the yield of dimer was 0.54, showing an improvement of 17% than the standard method. The results also showed that the addition of surfactants difficult the processing of cell pellets, i.e., difficult the re-solubilization. The temperature increase led to an almost complete loss of the target-protein.

The development of this project demonstrated that the most effective method for cell lysis of *E. coli* cells is mechanical lysis using HPH by applying minimal pressure to the sample - a change to the standard method previously applied to obtain APRc. The changes made to the

standard procedure made it possible to optimize the rupture process and increase the yield of the dimeric form of APRc from 0.45 to 1.74 (74%).

Keywords: *Rickettsia* | APRc | cell lysis | optimization | yield

Table of contents

List of Figures

List of Tables

List of abbreviations

ABS	Aqueous biphasic systems
AG	Ancestral group
APRc	Aspartic protease from *Rickettsia conorii*
$CaCl_2$	Calcium chloride
E. coli	*Escherichia coli*
EGFP	Enhanced green fluorescent protein
GFP	Green fluorescent protein
HPH	High pressure homogenizer (Emulsiflex^R-C3)
IgG	Immunoglobulin G
IMAC	Immobilized Metal Affinity Chromatography
Min	Minute(s)
MW	Molecular weight
OD_{280nm}	Optical Density at 280 nm
OD_{600nm}	Optical Density at 600 nm
P_{sample}	Pressure applied on the sample.
RMSF	Rocky Mountain spotted fever
Rpm	Rotations per minute
SDS	Sodium Dodecyl Sulfate
Sec	Second (s)
SEC	Size-exclusion chromatography
SFG	Spotted fever group
TG	Typhus group
TRG	Transitional group
WT	Wild type

HPLC High performance liquid chromatography

1. Introduction

1.1. *Rickettsia* and rickettsioses

Rickettsia includes the causative agents of some of the most historically significant and severe bacterial diseases of humans, first described by Ricketts in 1909 (Abdad et al., 2018; Wood, D. O., Azad, A. F., 2000). They are Gram-negative and obligate intracellular bacteria, that require eukaryotic cells in which to replicate and are transmitted to mammals by arthropod vectors such as ticks, lice, and fleas (Cruz et al., 2014).

Rickettsia species have been classified into several distinct genetic groups including the ancestral group (AG), spotted fever group (SFG), typhus group (TG), and transitional group (TRG). Several rickettsial species belonging to the TG and SFG are pathogenic to humans causing serious illnesses like epidemic typhus (*Rickettsia prowazekii*), Rocky Mountain spotted fever (RMSF) (*Rickettsia rickettsii*), and Mediterranean spotted fever (MSF) (*Rickettsia conorii*) (Angelakis, E., Raoult, D., 2017; Cruz et al., 2014).

So, there is a growing concern about rickettsial diseases (rickettsioses) and their impact on global health, with members of the genus *Rickettsia* being identified, together with other bacteria, as emerging/re-emerging pathogens, responsible for most of emerging infectious diseases events between 1940 and 2004 (Angelakis, E., Raoult, D., 2017; Cruz et al., 2014; Renesto et al., 2005).

1.1.1. *Epidemiology*

Rickettsioses are the pathologies caused by *Rickettsia*, associated with arthropod vectors (e.g., ticks, lice, mites, and fleas) and are transmissible from animals to humans and can originate moderate to severe illnesses. Most SFG rickettsiae are held by ixodid ticks and transmitted by their bites during feeding (from saliva). Two of the most severe rickettsioses in the world are Rocky Mountain spotted fever (RMSF) and Mediterranean spotted fever (MSF). MSF is the most important tick–borne disease in Portugal. In fact, from 1989 to 2000 the incidence rate of MSF in Portugal was one of the highest in the Mediterranean area (9.8 cases per million) (Abdad et al., 2018; Cruz et al., 2014; De Sousa et al., 2003).

Rickettsial infections occur worldwide (rickettsial organisms have been found on all continents (except Antarctica) with the geographic distribution of each species dependent on vector constraints and climatic conditions. As the international travels and tourism increase, associated with an expansion in geographic distribution of vectors due to climate change, tick-borne rickettsioses are expected to increase in the next years, increasing even more the need for

identifying and studying the different rickettsiaes, particularly the pathogenic ones (Parola et al., 2013; Stewart, A. G., Stewart, A. G. A., 2021; Wood, D. O., Azad, A. F., 2000).

1.1.2. *Pathogenicity and virulence factors*

To resist the rickettsioses, it is important to know in detail the strategies used by bacteria to spread infection and cause disease in hosts. Pathogenic bacteria produce different types of molecules recognized as virulence factors, which help them to invade and colonize the host. Virulence factors are unique for specific species (Sharma et al., 2017). The comparison of the multiple available genomes of *Rickettsia* is demonstrating to be the most practical method to identify new factors that may play a role in pathogenicity. As obligate intracellular bacteria, rickettsiae need to adhere to and invade target cells to gain access to a nutrient-rich cytoplasmic niche, which facilitates their proliferation, replication, and dissemination, thus leading to the successful establishment of infection in the host (A. Sahni et al., 2019).

Life cycle

The life cycle of these bacteria starts when they infect the arthropod vector and start to grow and multiply in almost all organs. The arthropod vectors may transmit the rickettsial species though vertical/transovarial transmission or horizontal transmission. In the transovarial transmission, the ovaries of adult female ticks are the infected organs, and consequently the oocytes are also infected, and so the eggs. Thus, the infection persists in all the following stages, spreading to the next generations. However, the horizontal transmission occurs during the feeding of the arthropod vector, for example most of SFG rickettsiae are in ioxid ticks and are transmitted by their bites during feeding (saliva) (Abdad et al., 2018). In this type of transmission, the salivary glands are infected by bacteria so when the vectors are feeding, they can bite and infect humans. When feeding, they salivate into the wound and release immunomodulatory and anti-inflammatory responses, increasing the blood flow and delaying the host's mechanisms to eliminate parasites. Therefore, *Rickettsia*'s are secreted into the bite site and benefit of this area to establish infection. (Azad and Beard, 1998; Kazimírová and Štibrániová, 2013; Sahni et al., 2019; Walker et al., 2003). After *Rickettsia* are in the host cells, they have to escape from the phagosome before replication because then, in the cytosol, bacteria start the binary fission and redistribute inside the cell, completing the life cycle when the host cell lysates or the infections spreads through neighbour cells (Mcclure et al., 2017). The infection of human hosts can also occur by the feces of infected lice or flea released in the skin during the feeding. By scratching this area, the rickettsial organisms can enter the skin

throughout the bite site, making humans considered as accidental hosts for *Rickettsia* (Abdad et al., 2018; Blanton, 2019).

Pathogenicity

As mentioned before, rickettsiae are obligate intracellular bacteria, meaning that they need to adhere to and invade target host cells to survive, multiply, and successfully establish the infection. Also, they can take advantage of the nutrients and energy available in the cytosol of the host (Cruz et al., 2014; Li and D. H. Walker, 1992). The main target of the infection through *Rickettsia* is the vascular endothelium, where these pathogens can proliferate, injuring the endothelial cells (Stewart and Stewart, 2021; Walker et al., 2003). Furthermore, rickettsiae can infect other types of cells, like dendritic cells and macrophages, being important the identification and characterization of different molecular components involved in host-*Rickettsia* interactions, enabling the application of new discoveries in prevention and new targeted therapies for rickettsioses (Chan et al., 2011; Curto et al., 2021; Sahni et al., 2019; Wood, D. O., Azad, A. F., 2000).

During some periods of infection process, *Rickettsia´*s spend time outside the host, before gaining intracellular access, not having protection from the host cell, and getting exposed to antibodies and other components of the host's innate immunity machinery (Chan et al., 2011; S. K. Sahni et al., 2013). Therefore, the survival of *Rickettsia*'s depends on their capacity of evade killing until being able to invade a proper cellular host. These bacteria are resistant to serum bactericidal effects and can evade complement-mediated killing, like it was demonstrated by several studies (Curto et al., 2021; Barro, A. M., 2020). Until now, three rickettsial surface proteins were identified as relevant participants in the serum resistance, and they illustrate two different mechanisms mediating partial survival of *Rickettsia* in human serum through enrolment of regulators of complement activation, suggesting that rickettsiae may have grown multiple mechanisms to inhibit recognition by host serum components (Curto et al., 2021; Barro, A. M., 2020). So, a deep and better understanding of immune evasion mechanisms and pathogenicity in rickettsiae is fundamental for the development of new approaches to treating rickettsial infections.

Virulence factors

The life-threatening character of many rickettsial species is the consequence of their highly virulent properties and unique biological characteristics (Cruz et al., 2014). Virulence is described as an ability of an organism to infect the host and cause a disease. The mediators of

virulence in the bacterial systems are often moonlighting proteins, which are the primary weapons for the host targeting, needing immediate attention (Sharma et al., 2017).

For enlightenment, moonlighting proteins embrace a class of multifunctional proteins in which a single protein performs multiple physiologically relevant biochemical or biophysical functions. The function of a moonlighting protein can vary because of changes in cellular localization, cell type, oligomeric state, or the cellular concentration of a ligand, substrate, cofactor, or product (Jeffery, 1999). Thus, moonlighting proteins represent a vast group of virulence factors in many pathogenic bacteria.

Back to virulent properties, in the specific case of *Rickettsia*, to survive, the bacteria should enter host cells, adhere, and invade them. The process of adhesion happens because of the connection of rickettsial proteins to host cells receptors. When *Rickettsia* gain access to the host cytosol, bacteria start to spread intra and intercellularly through the host cell actin cytoskeleton. After bacteria invasion, a defence mechanism called autophagy is activated in host cells to eliminate the invading bacteria, owing to cellular stress. Though, several intracellular bacteria have developed mechanisms to avoid autophagy (Cruz et al., 2014; M. Li et al., 2015; Sharma et al., 2017; Teixeira, B. D., 2020).

Several studies have demonstrated that the outer membrane proteins that influence pathogenicity and virulence of rickettsiae can be cleaved *in vitro* by APRc, a protease found in the outer membrane of *Rickettsia*. Thus, this protease might have a role in virulence of rickettsiae and can be an example of a moonlighting protein (Cruz et al., 2014; M. Li et al., 2015; Barro, A. M., 2020; Teixeira, B. D., 2020).

1.1.3. *Disease symptoms and therapeutics*

Rickettsioses are difficult to diagnose because there are no rapid tests available to establish the diagnosis during acute infection. The diagnosis of a rickettsial infection is usually achieved through the gathering of complete clinical history indicating exposure to a potential source of rickettsial disease alongside laboratory testing (Stewart, A. G., Stewart, A. G. A., 2021). Recognition of these diseases as a cause of acute febrile illness is important because, when proper treatment is administered, symptoms can be quickly alleviated. When highly pathogenic species are involved (for example *Rickettsia rickettsii*), delay in treatment is associated with poor outcomes and death, since some infections can be rapidly progressive (Abdad et al., 2018; Blanton, 2019; Parola et al., 2005).

Symptoms of these diseases include high fever, headache, myalgia, and rash. An eschar may develop at the site of inoculation and provide a diagnostic clue. In severe disease,

complications may include renal failure, myocarditis, meningoencephalitis, pneumonitis, and acute respiratory distress syndrome. High fever is the main symptom, that truly characterizes most of rickettsioses. In part, disease severity depends on the patient, on the causative *Rickettsia* species and their associated virulence factors - as a way of example, RMSF and epidemic typhus lead to a more severe disease course, whereas African tick bite fever is typically a mild disease (Cruz et al., 2014; Stewart, A. G., Stewart, A. G. A., 2021).

Although rickettsioses can be treated with antibiotics, more targeted therapeutic approaches against rickettsiae are desired for several reasons. Rickettsioses are highly predominant especially in poor countries, and there are indications of the development of antibiotic resistance. (Botelho-Nevers et al., 2012; Osterloh, 2020). Also, bacterial pathogenicity can be associated with proteolytic enzymes, among other factors (A. Sahni et al., 2019). The emergent and severe character of rickettsioses, together with the lack of targeted therapies, strengthen the importance of identifying new protein factors that may work as potential targets for the development of more efficient therapies against these diseases (Cruz et al., 2014; Parola et al., 2005).

1.2. *Rickettsia*'s aspartic protease of the retropepsin type, APRc

The emergent character of rickettsioses and the lack of reliable targeted therapeutics and specific treatment against them, strengthens the importance of identifying new protein factors for the potential development of innovative therapeutic tools (Cruz et al., 2014).

Therefore, proteases play critical functions in evasion of host immune defences, acquisition of nutrients for growth and proliferation, and facilitation of dissemination (tissue damage) during infection. The retropepsins were first identified with the discovery of the Human Immunodeficiency Virus (HIV-1) protease in 1980's. These types of proteases require homodimerization of two monomeric units to form a functional enzyme. So, 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 (Konvalinka et al., 2015; M. Li et al., 2015; Wlodawer and Gustchina, 2000).

Cruz *et al* (2014), identified and characterized a novel retropepsin-like enzyme, APRc (Aspartic Protease from *Rickettsia conorii*), that is expressed by several rickettsial species. This protease shares several enzymatic properties with other aspartic proteases, including autolytic activity, activity dependent on dimer formation, optimum pH, and sensitivity to specific HIV-1 protease inhibitors. Furthermore, this novel rickettsial protease is accumulated into the outer membrane of *R. conorii* and is predicted to comprise three transmembrane domains, plus a

soluble catalytic domain, extracellular oriented. It was also demonstrated *in vitro* processing of two conserved autotransporter proteins by APRc, located at the outer membrane, suggesting its potential role as a modulator of other rickettsial surface proteins involved in adhesion to and invasion of host cells. Thus, this protease may also assist the spread of the infection and dissemination of bacteria into deeper tissue through the shedding of cell surface adhesion molecules or the inactivation of the components of the host immune system (Cruz et al., 2014; Curto et al., 2021; M. Li et al., 2015; Barro, A. M., 2020).

In recent studies, Curto *et al* (2021) proved and confirmed that expression of the full-length APRc in *E. coli* is sufficient to promote serum resistance. They provide evidence that expression of the full-length protease in *E. coli* is sufficient to promote resistance to complement-mediated killing and that interaction with IgG contributes to serum resistance, since the APRc binds to IgG of distinct species as well as to different classes of human immunoglobulins. Also, it was demonstrated that APRc is a novel moonlighting protein from *Rickettsia*, acting as a novel evasin and contributing to *Rickettsia*'s immune evasion toolbox (Curto et al., 2021).

The autoprocessing activity of APRc in mammalian cells was visualized by Teixeira, B. (2020), by identifying differences in the accumulation patterns between the WT and mutant forms of the shorter fusion proteins. When they compared the different truncated forms of APRc, the results suggest that the autoprocessing of the longer form ($APRc_{87-231}$) is slower when compared with the shorter constructs ($APRc_{99-231}$ and $APRc_{110-231}$). Moreover, the generation of different truncated constructs and the analysis of the different forms by analytical size-exclusion chromatography revealed that only the form $APRc_{110-231}$ accumulated as a dimer, whereas the remaining forms accumulated mostly as monomers (Teixeira, B. D., 2020). That is why in this work we use one of the shorter constructs, $APRc_{110-231}$-HisShort, because as previously demonstrated, this is the form accumulated as a dimer upon expression in *E. coli*. This dimeric form was also used to determine the crystal structure, by Dra Isaura Simões, and showed to be stable.

Although the topology of the dimer observed in the crystals appears to reflect a crystal-packing artifact, and not the biologically relevant interface, the structure of the APRc monomer is closely related to those of the other retropepsins, since the APRc monomer shows the canonical fold observed in all retropepsins, either of viral or of eukaryotic origin. The dimer of APRc observed in the crystals is entirely different from the canonical dimer of retroviral enzymes and is incapable of forming a proper active site. Though the topology of the dimer is the same in all three crystal forms, the angle between the two molecules is considerably

different, suggesting that the observed dimer must be an artefact of the expression and crystallization conditions. In addition, it was discovered that each one of the retroviral monomers had a molecular weight in the range of 10 to 15 kDa (Cruz et al., 2014; Curto et al., 2021; M. Li et al., 2015).

APRc was previously produced and purified from *E. coli* in Dra Isaura Simões laboratory, and the challenges found were related to the high yields of monomer and poor yields of active dimer that were obtained until the end of the process, which has been a limitation to get the canonical dimer of APRc and pursue studies on structure-function relationship of APRc, enabling the design of specific inhibitors. It was assumed that these limitations are related to the established method of APRc production, more specifically in the recovery of recombinant APRc from *E. coli* cells (cell lysis). The method currently used is a mechanical method, using a High-Pressure Homogenizer (Emulsiflex^{R}-C3), a mechanical cell lysis method that may be aggressive for the extraction of functional protein, i.e., causing losses in the yield of functional protein. Therefore, to overcome these limitations, this project focused on evaluating the cell lysis step in the process of production of the APRc, in order to pursuit a better understanding of APRc and the development of new targeted therapies.

1.3. *Escherichia coli* as a heterologous protein expression system

The production of recombinant proteins in microbial systems have revolutionized biochemistry. The power to express and purify the desired recombinant protein in a large capacity allows its biochemical characterization, its application in industrial processes and the development of commercial goods (Rosano and Ceccarelli, 2014). The two basic elements for expression of recombinant proteins are the expression vector and the expression host.

An expression vector is created in a way that a gene cloned will be transcribed. Plasmid vectors are used for the cloning and expression of a foreign gene in prokaryotic systems. In addition, a promoter is the key component of an expression system because of its role in controlling the transcription initiation of associated genes (Bernaudat et al., 2011; Kaur et al., 2018).

There are several hosts used to produce recombinant proteins such as bacteria, yeast, plants, and animals. *Escherichia coli* (*E. coli*) is the most common heterologous system for protein expression/production, from not only prokaryotic but also eukaryotic origins, may due to the strong advantages of using *E. coli* as host organism that include rapid growth, high cell density, transformation with exogenous DNA is highly efficient causing expression of

recombinant proteins at extremely high rates, and finally low-cost medium components (Evans and Xu, 2011; Kaur et al., 2018; Kim et al., 2016; Rosano and Ceccarelli, 2014).

The selection of the most appropriate *E. coli* host strain is crucial for the expression of recombinant proteins and depends upon the expression vector used. In the last years, the predominant vector used is BL21 (DE3) and its derivative strains. This vector contains T7 promoters, which turn this *E. coli* BL21 strain the preferential host, specifically because it produces the RNA polymerase for this promoter (Evans and Xu, 2011; Kaur et al., 2018).

The appropriate combination of vector and host maximizes the quantity and quality of protein produced. For example, a lot of life-saving medicines are produced in *E. coli* actually, but until now, heterologous gene expression in *E. coli* is not an easy task, since proteins are very diverse in structure and physicochemical properties, it is impossible to predict if a protein of interest will have a good expression, simple to purify, present desired biological/therapeutic activity or crystallize (Bernaudat et al., 2011; Kaur et al., 2018). Some of the factors affecting expression of heterologous proteins are the size of the heterologous protein (*E. coli* only express proteins with less than 100 kDa), and toxicity. Consequently, can occur expression problems like no or low expression levels, or the formation of inclusion bodies and expression of inactive proteins. The inclusion body formation is a consequence of the overexpression of heterologous proteins in bacteria, that results in the accumulation of protein product in inactive and insoluble deposits inside the cells. Although the conversion of insoluble inclusion body proteins into soluble and bioactive ones can be achieved, the protein refolding is still a processual challenge. Currently, some methods as dilution and dialysis are employed for the renaturation and refolding of inclusion body proteins. However, these methods usually result in extremely low recoveries (Evans and Xu, 2011; Kaur et al., 2018; Kim et al., 2016). Besides this, when *E. coli* is used as an expression system, most proteins are expressed intracellularly, and suitable cell disruption methods are required. Although a cell disruption method is used that is thought to be suitable to produce a protein of interest, this method may not be the one that allows the greatest amount of this protein to be obtained, hence the importance of this project (Rosano and Ceccarelli, 2014; Martins et al., 2018; Smolskaya et al., 2020).

It is possible to do an optimization of numerous factors affecting expression of recombinant protein in soluble fraction. That optimization can be done at the concentration of the inducer, at the induction temperature, optical density (since the phase of growth of bacterial culture at the time of induction plays a significant role in soluble form of recombinant protein production), addition of glucose (reducing the toxic effect of protein before induction), and

addition of additives in culture medium (Kaur et al., 2018). This will be further discussed at the future perspectives chapter.

The current procedure for production of APRc using *E. coli* as heterologous system has as the main limitation the low, quantity of functional protease (yield) produced, which difficult the study and possible industrial applications of APRc. Therefore, in this project we aimed to increase this yield, focusing on designing the most adequate cell lysis procedure for *E. coli* disruption and the recovery of functional APRc.

1.4. Cell Lysis Methods

The increasing demand for intracellular bioproducts for applications in the food and pharmaceutical industries highlights the importance of cell lysis (cell disruption) operations. Cell disruption is a key unit operation to make valuable intracellular target products accessible for further downstream unit operations, such as protein purification for studying protein function and structure. Is a process in which the outer boundary or cell membrane is broken down or destroyed to release the intracellular materials such as DNA, RNA, proteins, organelles, among others. The choice of most appropriate cell lysis method is influenced by subsequent separation/purification steps, nature of the target molecules and desired quality and purity of the final products (Eggenreich et al., 2017; Islam et al., 2017).

The physical characteristics of the suspension that results from cell disruption (cell lysate) have a significant impact on the subsequent downstream processing operations. The particle size of the cell debris strongly influences the particle settling velocity in the centrifuges and the viscosity of the suspension is inversely related to the particle settling velocity (Balasundaram et al., 2009).

The cell lysis methods are divided essentially in mechanical and non-mechanical, as outlined in Figure 1.

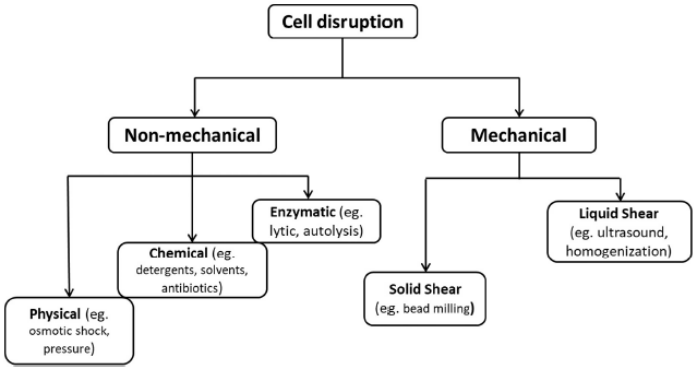


Figure 1 - Classification of conventional cell disruption methods. Adapted from Wang, 2020 (Wang et al., 2020).

It is important to take into consideration that much of the energy input needed by mechanical methods of cell disruption are dissipated as heat, which may degrade the heat-sensitive proteins. On the other hand, chemical cell disruption does not require energy-intensive instruments for the operation and therefore the costs of setup and operation are relatively lower. Mechanical methods can also be combined with non-mechanical methods to increase the overall efficiency of cell disruption (Balasundaram et al., 2009; Lo et al., 2016).

After this enlightenment, the focus of this project is to increase the recovery yields of active dimer of APRc through the application of different cell lysis methods. As mentioned before, this protease was previously produced in this laboratory throughout *E. coli* expression system and its extraction carried out using a conventional mechanical lysis, *i.e.*, using a HPH equipment. This extraction releases low amounts of active dimer, which is a limitation for the study of protease and the development of new target therapies. The current cell lysis method is thought to be aggressive for protein extraction because of the high pressure applied to the pellet, which involve subjecting the proteins to high temperatures, since temperature and pressure are directly proportional. In the next paragraphs, it will be presented an overview of the general cell lysis methods, and then the most commonly used cell lysis methods to release the intracellular content from *E. coli*, that is the expression system used in this project.

In the process for obtaining a bioproduct, cell disruption occurs after the clarification step, *i.e.*, after the separation between cells and culture medium, and the washing of these cells with a solvent. The liquid resulting from the cell disruption operation is called a homogenate or lysate, and consists of the target molecule, contaminant molecules, and cell membrane fragments. The undesirable compounds present in this lysate must be removed by suitable unit operations, which comprise the process of biomolecule purification (Pereira and Hasmann, 2020). The processes for obtaining a bioproduct consist of the steps represented in Figure 2.

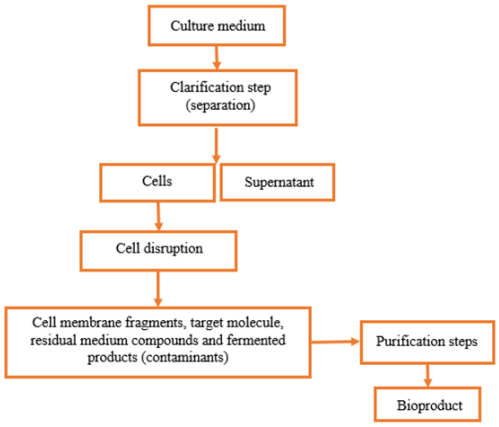


Figure 2 - General sequential process flow for obtaining a bioproduct. Adapted from Pereira, 2020 (Pereira and Hasmann, 2020).

Figure 2 presents a generic flowchart of a bioproduct production process. For obtaining intracellular products it is required a cell disruption step. The cell disruption step has the disadvantage of promoting an increase in the number and diversity of contaminant molecules in the medium containing the biomolecule to be purified, as well as an increase in viscosity caused by the release of components of the cytoplasm of the cells, such as DNA. It should be noted that with each purification step added to the process a percentage of the target molecule is lost (Pereira and Hasmann, 2020).

The criteria used in selecting the cell disruption technique must consider several factors, such as:

- The size and type of the cell, and its culture conditions
- Shear stress
- Temperature
- The operation time, energy, and investment cost
- Maximization of the process yield (high productivity)

In more detail, the shear stress is an important factor because when cell disruption is performed, for example, in homogenizers, the pressure and the shear stress applied on the cells besides promoting the cell disruption, can cause damage to the target molecule, with the possibility of losing its biological function, since that is a sudden drop in pressure during the passage of cell suspension through the exit orifice. Relatively to the temperature, for example in proteins, applying certain temperatures, they suffer vibrations inside them that damage their structure due to the disruption of non-covalent interactions, resulting in the loss of their biological functionality, thus denaturing by heat (Pereira and Hasmann, 2020).

A summary between diverse types of cell lysis techniques (mechanical and non-mechanical), that are mostly used, is presented in Table 1, including also the corresponding main advantages and disadvantages.

Table 1 - Overview of most common cell lysis methods. Adapted from Islam et al, 2017 (Islam et al., 2017).

Methods		Equipment or Technique used	Advantages	Disadvantages
Mechanical		- High pressure homogenizer - Bead mills	- High efficiency - Not cell dependent	- Heat generated could damage intracellular products - Expensive - Difficult to purify the lysate
Non-mechanical	**Physical**	Thermal lysis	-Independent of the cell type - Easy to implement	- Expensive - Damage to proteins and intracellular components
		Cavitation /Ultrasonication	- Independent of cell type - Large scale - Lower temperatures and energy level	- Expensive - Can cause damage to sensitive proteins - Difficult to purify sample from debris
		Osmotic shock	- Extract sensitive intracellular products	- Not suitable for all cell types
	Chemical	Alkaline lysis	- Extract sensitive intracellular components - Suitable for all kinds of cells	- Slow process (6–12 h)
		Detergent lysis	- Suitable for protein release	- Not suitable for isolating sensitive enzyme and proteins -Expensive - Removal of chemical reagent from sample after lysis is difficult - Complete lysis not possible
	Biological	Enzymatic lysis	- Specific for cell types - Extract proteins	- Complete lysis not possible - Expensive reagents - Must be used in combination of detergents for bacteria

Therefore, looking at Table 1, mainly in mechanical processes it is essential to reduce the operating temperature since they promote the increase of the temperature during cell disruption, which can denature the targeted protein (Pereira and Hasmann, 2020). However, non-mechanical methods have to be carefully studied, since they cannot be adequate for all cell types and so prevent the extraction of targeted protein. Also, non-mechanical methods can cause the lost of targeted protein because of surfactant/detergent compounds when used, or contamination of organelles with chemicals, and in case of sensitive proteins or application of heat in the cell lysis process, make damage to them. Finally, they can also make the purification more difficult because of separation between debris and protein of interest (Pereira and Hasmann, 2020).

As well, a crucial point to take into consideration is the preservation of the bioproduct during cell disruption operation. For that, an evaluation of the efficiency of cell disruption should consider the yield of the bioproduct in its active/functional form. The release of intracellular products after cell disruption, including proteases, might result in hydrolysis of the desired protein. Furthermore, during cell lysis, the high viscosity in the homogenate, because of the release of nucleic acids and structural proteins from the cell, can lead to diffusion problems, which in turn impose difficulties on the subsequent purification steps such as filtration or chromatography. Otherwise, the addition of nucleases or proteases, or an appropriate adjustment of the pH value are strategies that can improve the rheological characteristics of the medium. However, it is always important to verify that these do not destroy the biomolecule of interest (Pereira and Hasmann, 2020).

Next, recent studies of improvements in cell lysis methods applied to recover and purify intracellular proteins from *E. coli* are briefly presented. The mechanical lysis is the most used method for disruption of *E. coli* cells. Mechanical methods are based on the action of physical forces such as compression and shear. Taking this in consideration, the mechanical lysis methods, mainly due to the heat production during the disruption, which can cause the denaturation of the protein of interest. Therefore, it has become urgent to search for biomolecule-friendly alternatives that allow to obtain higher yields of the target-bioproduct biologically functional. In fact, the main challenge in bioprocesses using heterologous expression systems is to guarantee that the recombinant proteins are functional and active after all manufacturing stages (*e.g.*, extraction and purification) (Pereira and Hasmann, 2020).

Martins *et al*. proposed a new method to extract recombinant intracellular GFP from *E. coli* cells by using aqueous solutions of surface-active compounds combined with freezing-thawing cycles. In this work, the recovery of GFP from *E. coli* cells was achieved using

tensioactive compounds (like surface active ionic liquids), demonstrating some of ionic liquids solutions are more effective than the use of ultrasonic-assisted methods for the recovery of GFP. The results shown that, even at low concentrations, the surface-active compounds are able to extract the whole GFP content from the cells (Martins et al., 2018).

Dos Santos *et al.* also made studies in that field. They evaluate aqueous biphasic systems (ABS) composed of choline chloride and different polymers as effective platforms to separate and purify GFP from cell lysate produced in *E. coli*, after 4 cycles of freezing-thawing (physicochemical lysis). In this work, it was possible to obtain GFP (with >99% purity) into the polymeric or choline chloride rich phases, hence ABS showed potential for protein purification (Dos Santos et al., 2018).

Lo *et al.* have proposed a versatile cell disruption method based on alcohol permeabilization to extract an enhanced green fluorescent protein (EGFP) from *E. coli*. Regardless that the active EGFP obtained from the optimized cell disruption with an alcohol was comparable to that from ultrasonication treatment, the alcohol-based cell disruption was the advantage to be less energy intensive (in comparison with ultrasonication and glass bead vortex methods). This approach is a promising operation to reduce the operational costs even if performed as a large-scale cell disruption process (Lo et al., 2016).

This project aims to evaluate different cell lysis methods to enhance the liberation of organelles, and purify them, with the intention of increasing the recovery of functional APRc. Therefore, based on previous findings of both Supervisor's research groups, new strategies are designed to evaluate if an adequate choice and implementation of non-mechanical methods, such as physicochemical lysis (applying freezing-thawing cycles, surfactant solvents and temperature), can improve the recovery yields of active form of APRc. Physicochemical lysis encompasses two types of lysis, *i.e.*, freezing-thawing cycles and the addition of surfactant compounds as cell disruption additives. The latter strategy aims to increase membrane permeability and help to release the intracellular components of interest. However, to confirm if these are effectively the most efficient, the physicochemical lysis tests were compared with the method already established for the recovery of APRc, namely the HPH. For this mechanical cell disruption method, it was also performed a screening to find conditions that maximize the extraction of intracellular APRc, namely by making changes in the operating conditions like pressure and number of passages of the pellet throughout the equipment, among others. It is important to note, that the current cell lysis method using HPH can be optimized since there is still a need to increase the yield of active form of APRc. Anyway, the current recovery yields of APRc obtained with HPH are good starting points, which make to believe that these can be

increased if adequate tests and changes are included. In the following sections, the cell lysis methods used in this work are explained in more detail.

1.4.1. *Mechanical lysis with High-Pressure Homogenizer*

High-pressure homogenization (HPH) is the most frequently used cell disruption operation in the bioprocessing industries. With this method, cells in media are forced through an orifice valve using high pressures. The equipment consists of pistons designed to apply high pressures and force the cell suspension through a narrow orifice. Throughout this orifice, the suspension passes at high speed, undergoing a sudden depressurization and collision against a rigid and immobile surface inside a chamber at atmospheric pressure. The instantaneous reduction in pressure associated with the impact causes the cell disruption (Islam et al., 2017; Pereira and Hasmann, 2020).

The key operating variables in a HPH are the operating pressure, feed rate, temperature, piston oscillation frequency and its distance from the base, and the number of passes through the valve. The cell disruption occurs under high pressure conditions and, therefore, there is an increase in the suspension temperature during processing. Therefore, it is essential that the equipment has an efficient cooling system, especially in cases where the target molecule is thermosensitive (e.g., proteins). For instance, in bioprocesses that use HPH, the temperature of the suspension can increase 15 ºC after one cell disruption cycle. In processes for the recovery of intracellular proteins, it is recommended to reduce the operating temperature immediately before and after the process, in order to avoid denaturation of the proteins (S. T.L. Harrison, 2011; Islam et al., 2017; Pereira and Hasmann, 2020; Wang et al., 2020).

Sauer *et al*. proposed a model to relate the amount of protein released and the pressure applied in a HPH for the lysis *E. coli* cells. The model, represented by Equation (1), represents the variation in efficiency of this breaking process as a function of the number of passes of the suspension through the homogenizer valve (N) and the pressure difference (ΔP in N/m^2), under fixed temperature and with the same cell type (Islam et al., 2017; Pereira and Hasmann, 2020).

$$\ln\left(\frac{R_m}{R_m - R}\right) = k\, N\, \Delta P^{\alpha} \qquad \text{(Equation 1)}$$

where R is the concentration of protein released (kg/m^3), R_m is the maximum concentration of protein available for release (kg/m^3), P is pressure in MPa, N is the number of passes, k is the velocity constant (h^{-1}), and α is a constant of resistance of the cells to rupture.

The constant k depends on temperature, concentration and cell type, and α depends on cell type and growth conditions and can range from 0.86 to 2.9 (for *E. coli* suspension has a value of 2.2) (Islam et al., 2017; Pereira and Hasmann, 2020).

Equation 1 suggests that the amount of the released biomolecule is dependent on the pressure applied to the process, indicating that high operating pressure would be beneficial for intensifying the efficiency of the cell disruption process. In theory, the complete disruption of a cell suspension can be achieved with a simple homogenization step at high pressure; however, the efficiency of the disruption process can be improved, by recirculating the cell suspension, which must be confronted with the rise of the process cost, the increase of the degraded fraction of the target molecule and the size of the cell fragments, as these will have to be removed later by centrifugation (Pereira and Hasmann, 2020; Wang et al., 2020).

As mentioned before, the high shear forces caused by mechanical disruptions can destroy cell organelles and denature enzyme complexes or enzymes located close to the membrane. In addition, when cells are integrally disrupted, all intracellular contents, including nucleic acids, organelles, and cell fragments, are released along with the target molecule. Consequently, the cell homogenate will present high viscosity and proportion of contaminants, making the further purification steps of the target molecule difficult. So, other methods should be tested, to help and simplify the purification step, as well as, to increase the recovery yields.

1.4.2. *Physicochemical lysis*

The most common and simple physical lysis method is freezing-thawing. In this method, cells are subjected to repeated freezing-thawing cycles, which will guarantee lesions, or complete disruption, due to perforation of the wall and membrane by intracellular ice crystals creating pores that will allow the permeabilization of the bioproduct. Several factors control the efficiency of this type of disruption, among which the most important are cell type and age, temperature, velocity of the freezing-thawing cycles. Despite of these advantages, it is a time-consuming, high-cost process due to the energy required for freezing and unsuitable for freeze-sensitive biomolecules. In general, freezing-thawing is applied in combination with other disruption methodologies, like chemical lysis, to obtain an acceptable efficiency (Islam et al., 2017; Pereira and Hasmann, 2020).

Chemical lysis methods use addition of chemical agents for enhancing the disruption of the cell membrane. Chemicals/solvents used in cell disruption methods will break the cell membrane through different mechanisms, for instance: by changing the pH of lysis buffer (osmotic shock lysis); use of detergents to solubilize the membrane proteins and cell membrane

permeabilization. Detergents, also called surfactants, have the action of dissociating proteins and lipoproteins from the cell wall, causing the formation of pores and release of the target molecule. There is a wide variety of surfactant compounds already employed in cell disruption approaches, such as: bulky quaternary ammonium-based salts, anionic surfactants such as sodium lauryl sulphate, SDS (sodium dodecyl sulphate) or sodium sulfonate, cationic surfactants, non-ionic detergents, among others. Regardless of high solubilization ability, the use of surfactants as additives in cell disruption have the disadvantage of generating foam and the possibility of denaturing and/or precipitating proteins, implying, therefore, the need to remove the surfactant during the subsequent purification steps (Islam et al., 2017; Pereira and Hasmann, 2020; Wang et al., 2020).

1.5. Project Aims

*Rickettsia'*s are causative agents of some of the most historically significant and severe bacterial diseases of humans. APRc is the aspartic protease from *Rickettsia conorii*, and has been identified as an immune evasion factor, contributing for *Rickettsia's* ability to resist serum bactericidal activity, clearly emphasizing its relevance as a virulence factor.

As detailed in previous sections, the current challenges to address the functional significance of APRc are related with limitations in purification yields of the canonical dimeric active form, hampering the structural characterization and design of specific inhibitors targeting the protease.

The release of intracellular proteins requires a step of cell lysis before the purification steps. This step plays a critical role in bioprocessing because it directly affects the quantity, biological activity, and subsequent purification of the target protein. The actual process to produce APRc in its functional form needs to be improved because of the low yields obtained. Therefore, the aim of this project is to evaluate the impact of different cell lysis methods for the improvement of APRc dimer recovery from *E. coli* cell lysates.

Therefore, the challenge of this project is to obtain higher yields of pure and homogenous recombinant APRc in is functional structure. Improving the purification yield of recombinant candidate protease targets, will help the development of novel antimicrobial protease-targeted therapies. So, with this work, the goal is to optimize the actual process of production of APRc, to obtain a greater amount of the purified and functional protease. Accordingly, several steps were carried out, namely:

i. Reproduce the process stablished in the laboratory - HPH.
ii. Analyse the standard procedure and possible steps to be changed.

iii. Apply different cell lysis methods and quantify the amount of protease that is obtained compared to the standard procedure.

iv. After all the trials done at the cell lysis step, analyse which one provides an optimization of the process and a purified and functional APRc.

2. Materials and Methods

2.1. Materials

The chemicals used in all the experiments were acquired from Sigma-Aldrich (Missouri, U.S.A), Merck (Germany), Fisher Scientific (Geel, Belgium), Discovery - Fine Chemicals (Wimborne, U.K.), EMD Millipore (U.S.A.), USB Corporation (Cleveland, U.S.A.), VWR (U.S.A), or from NZYTech (Portugal). All primers were acquired from Integrated DNA Technologies (IDT), BVBA (Leuven, Belgium). Table 2 includes detailed information about all chemicals used in this study.

Table 2 - Information about the chemicals used in the experiments.

Reagent	Molecular weight (g/mol)	Purity (%)	CAS number	Goal/Finality
Imidazole	68.08	99	288-32-4	Buffer
HEPES	238.3	99	7365-45-9	Buffer
Sodium chloride	58.44	99	7647-14-5	Buffer
LB Broth	-	-	-	Culture medium
Acrylamide	-	40	79-06-1	Density of mesh network
TEMED	116.21	99	110-18-9	Electrophoresis (initiator)
Glycerol bidistilled	92.09	99.5	56-81-5	Colouring solution
2-Propanol	60.1	99	67-63-0	Denaturing solution
Methanol	32.04	99	67-56-1	Denaturing solution
Trifluoroacetic acid (TFA)	114.02	99	76-05-1	Buffer for HPLC
Acetonitrile	41.05	99	75-05-8	Buffer for HPLC
Acetic acid	60.05	100	64-19-7	Denaturing solution
Hydrochloric acid fuming	36.46	37	7647-01-0	pH adjustment
IPTG (Dioxane free)	238.31	-	367-93-1	Protein expression
BugBuster	-	-	71456-3	
Ampicillin Sodium Salt	-	-	69-52-3	Antibiotic
Kanamycin	-	-	25389-94-0	Antibiotic
SDS	288.38	-	151-21-3	Electrophoresis gel
Bromophenol Blue Sodium Salt	691.95	-	62625-28-9	Colouring solution

2.2. Methods

2.2.1. *Isolation of plasmid DNA*

For plasmid DNA isolation, the NucleoSpin Plasmid QuickPure (Macherey-Nagel) was used, following the manufacturer's instructions. This isolation method is based on the alkaline lysis of *E. coli* cells followed by binding of plasmid DNA to the silica membrane column. The plasmid DNA was eluted in autoclaved Milli-Q water and quantified using NanoDrop® ND-1000 UV-Vis Spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA).

2.2.2. *Preparation of Escherichia coli BL21 star (DE3) competent cells*

The preparation of competent cells was made in two steps. Firstly, the pre-inoculum obtained from the BL21 stock was placed to incubate at 37ºC with agitation, overnight. In the second step, the pre-inoculum was diluted in 100 mL of medium and then grown at 37ºC with agitation. As soon as the OD_{600nm} reaches a value between 0.3 and 0.4, the culture was divided into two tubes and incubated in ice during 10 min. The cells were harvested by centrifugation at 3000 rpm, for 10 min at 4ºC, in an Avanti J-26S XPI Centrifuge with a JA-25.50 rotor (Beckman Coulter Life Sciences, Indiana, U.S.A). After removal of the supernatant, the pellets were resuspended in 5 mL of a 50 mM $CaCl_2$ solution. After another 10 min on ice, the cells were harvested by centrifugation under the conditions mentioned above. The supernatant was removed, and the cells were resuspended in 1 mL of the 50 mM $CaCl_2$ solution. Finally, 740 µL of an 87% glycerol solution was added to the 2 mL of bacteria resuspension from the previous step. Aliquots of 150 µL were frozen at -80°C.

2.2.3. *Transformation of the construct encoding APRc$_{110-231}$-HisShort into Escherichia coli BL21 star (DE3)*

For transformation of *E. coli* BL21 star (DE3) competent cells, 1 µL of DNA of the pET-23-based construct encoding APRc$_{110-231}$-HisShort sequence (appendix A) was inserted into the cells. Then, they were incubated on ice for 30 min, followed by 45 sec in a 42°C water bath, and finally 3 min on ice. 1 mL of LB medium was added, and the cells were grown for 2 h at 37ºC with shaking. Cells were harvested by centrifugation at 14800 rpm for 30 sec in a Sigma 1-14 centrifuge. Next, about 900 µL of the supernatant was removed, and with the remainder, the pellet was resuspended. The resuspended pellet was placed in the centre of a LB plate with agar and ampicillin, and after being spread over glass beads, it was left to grow overnight in the incubator.

2.2.4. *Production and purification of recombinant APRc*

Production of recombinant protein

After transformation of the DNA construct into competent *E. coli* BL21 Star (DE3) cells, these 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-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, for 3 h at 30ºC with agitation. The cells were harvested by centrifugation at 3000 rpm 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 30 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.

Cell lysis

For cell lysis, the frozen cells were thawed at room temperature and different lysis methods were tested. First, a "standard" protocol using High-Pressure Homogenizer (HPH) was performed, not only to have a starting point but also to verify if any of the methods studied can be further optimized. Afterwards, new, and different cell lysis methods were tested, namely, mechanical using the HPH and physicochemical employing freezing-thawing approaches. In each trial of a new lysis method, a control assay was performed in parallel using the standard HPH approach. Table 3 include the description of all tests performed.

Table 3 - Summary of all the lysis tests performed.

Trial	Cell lysis method	Buffers used	Experimental Conditions
1	Mechanical lysis – "standard" assay	20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl	3 cycles P_{sample} = 1700 bar
2	Mechanical lysis	Repetition of trial 1, for protein quantification and evaluation of the solid/solid ratio	
3	Physicochemical lysis: with freezing-thawing cycles	**(1)** 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl; **(2)** 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole, 500 mM NaCl and 250 mM SDS; **(3)** 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole, 500 mM NaCl and 500 mM SDS	Only 0.5 g of pellet and 5 g of solvent were applied; 4 cycles of freezing-thawing with buffers 1 and 2, 1 cycle with buffers 2 and 3; thawing at 20°C for 44 min and at 37°C for 29 min
4	Physicochemical lysis: with freezing-thawing cycles	20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl	4 cycles of freezing-thawing; one trial with thawing at 20°C for 44 min and at 37°C for 29 min and another only at 20°C for 44 min
5	Physicochemical lysis: with freezing-thawing cycles; Mechanical lysis	**(1)** 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl; **(2)** 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole, 500 mM NaCl and 100 mM SDS	1 cycle of freezing-thawing, with buffers 1 and 2 (at 20ºC for 44 min) 3 cycles and P_{sample} = 1700 bar (buffer 1) 3 cycles and P_{sample} = 0 bar (buffer 1) 6 cycles and P_{sample} = 1000-1250 bar (buffer 1)
6	Mechanical lysis	20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl	3 cycles and P_{sample} = 1700 bar 6 cycles and P_{sample} = 1000-1250 bar

As shown in Table 3, two types of lysis were tested: mechanical (HPH) and physicochemical (freezing and thawing cycles, with or without the addition of detergents in different buffered solutions). The mechanical lysis trials were carried out using the Emulsiflex^{R}-C3 equipment (Avestin, Ontario, Canada). To apply the cells, the equipment was pressurized

to 4 bar, washed with distilled water, and the pellet was placed under pressure. Depending on the specific test, the conditions to which the pellet was subjected were modified. It should be noted that the equipment has certain limitations, including the requirement for a minimum volume of 30 mL and the necessity of passing the pellet through it at least 3 times (3 cycles).

In the "standard" assay, in addition to these conditions, a pressure of 1700 bar was applied to the sample (pellet). In other assays, the number of cycles and the pressure applied to the sample were adjusted accordingly. For example, one trial was conducted with no pressure applied to the sample, and the pellet was passed through the narrowing of the equipment 3 times (3 cycles). In subsequent tests, the pellet was initially passed through the narrowing, and then the pressure was gradually increased until it reached a minimum pressure, at which point it was subjected to 6 cycles at that pressure. Finally, the equipment was washed with distilled water, soaked in a NaOH bath for 30 minutes, and then rinsed with a 70% ethanol solution.

For the physicochemical lysis, different freezing/thawing cycles (1 or 4) and thawing temperatures (20°C and 37°C) were tested in a water bath, along with different buffered solutions, both with and without surfactants (detergents). Initially, four triplicate trials were conducted using only 0.5 g of cellular pellet and 5 g of solvent. These trials involved 4 freezing-thawing cycles with two different buffers and 1 freezing-thawing cycle with the same two buffers. Freezing-thawing was carried out at 20°C for 44 min and at 37°C for 29 min.

To study the effect of increasing temperature during thawing, a separate assay was performed using only one buffer solution, and 4 freezing-thawing cycles were conducted. In this experiment, one pellet was thawed at 20°C for 44 min, and another pellet was thawed beyond that time at 20°C before being further thawed for 29 min at 37°C. In a subsequent trial, a freezing-thawing cycle at 20°C for 44 min was also tested with two different buffers.

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 cell lysis was performed using various methods, as described in the previous section. Subsequently, the resulting cell lysates were centrifuged at 24000 x*g*, at 4ºC for 20 min in an Avanti J-26S XPI Centrifuge with a JA 25.50 rotor. This was followed by ultracentrifugation at 41000 rpm at 4ºC for 20 min in an Optima l-100 XP ultracentrifuge with a 90Ti rotor (Beckman Coulter Life Sciences). The resulting supernatants were carefully recovered and applied to a Histrap HP 5 mL column (Cytiva, Massachusetts, U.S.A), which had been

previously equilibrated with a 20 mM sodium phosphate pH 7.4 buffer containing 10 mM imidazole and 500 mM NaCl.

Protein elution was carried out using a BioLogic DuoFlow™ chromatography system (Bio-Rad, California, U.S.A) with a four-step gradient of imidazole (50 mM, 150 mM, 200 mM, and 500 mM) in a 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 during the 150 mM/200 mM imidazole steps. These fractions were individually collected and then dialyzed using Spectra/Por™ 3 (3.5 kDa) dialysis membranes (Spectrum Labs, California, U.S.A) against a 20 mM HEPES pH 7.4 buffer solution, overnight at 4ºC with agitation.

The dialyzed proteins were subsequently ultracentrifuged at 41000 rpm at 4ºC for 20 min in an Optima l-100 XP ultracentrifuge with a 90Ti rotor. The resulting supernatants were further purified by cation exchange chromatography in a Mono-S 5/50 GL column (Cytiva), which had been pre-equilibrated with 20 mM HEPES pH 7.4 buffer. Protein elution was carried out using a BioLogic DuoFlow™ chromatography system with a linear gradient of NaCl (0-1 M) in a 20 mM HEPES pH 7.4 buffer, at a flow rate of 0.75 mL/min.

Finally, the eluted protein samples were analysed by SDS-PAGE, followed by Coomassie blue staining, and quantified using two different methods, as described in section 2.2.5.

Analytical-size Exclusion Chromatography

Analytical size exclusion chromatography was employed to assess the oligomerization state of APRc. Protein samples were injected into a Superdex 200 5/150 GL column (Cytiva), previously equilibrated with a 20 mM HEPES buffer pH 7.4 containing 100 mM NaCl. This analytical protocol was performed using a Prominence Shimadzu HPLC system (Shimadzu, Tokyo, Japan). The molecular weight of proteins was estimated by calibrating the column with standard proteins from the Gel Filtration LMW calibration kit (Cytiva): 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.

Reverse-phase HPLC

The proteolytic activity of the APRc was evaluated by Reverse-Phase HPLC, using oxidized insulin β-chain as the substrate. 3 μg of APRc were incubated overnight at 37ºC with 50 μg of oxidized insulin β-chain in 0.1 M sodium acetate buffer at pH 6. The samples were

then centrifuged for 8 min at 14800 rpm in a Sigma 1-14 centrifuge at room temperature. The supernatant was applied to a KROMASIL 100 C18 250 4.6 mm column (Teknokroma, Spain) previously equilibrated in 0.1% trifluoroacetic acid (TFA). Peptides were separated by Reverse-Phase HPLC using a Prominence Shimadzu HPLC system. The elution of peptides was performed with a linear gradient of acetonitrile (0-80%) in 0.1% TFA, at a flow rate of 1 mL/min, with absorbance monitored at 220 nm.

2.2.5. *Protein quantification*

Pierce™ 660 nm Protein Assay

In this colorimetric method, protein concentrations were estimated by referencing absorbances obtained for a series of standard BSA protein dilutions, which were assayed alongside the unknown samples. Initially, 10 µL of each replicate of standard, unknown samples, and the appropriate blank were added to the 96 microplate wells (Thermo Scientific™ Pierce™ 96-Well Plates). If the samples contained any detergent (such as SDS), 1 g of ionic detergent compatibility reagent was added to 20 mL of the Pierce 660 nm protein assay reagent before conducting the assay. Next, 150 µL of the protein assay reagent was added to each well. The plate was covered and mixed on a plate shaker at medium speed for 1 min, followed by incubation at room temperature for 5 min. In the end, the absorbance of the standards and unknown samples was measured at 660 nm. To determine the protein concentration of each unknown sample, a standard curve was prepared by plotting the average Blank-corrected 660 nm measurement for each BSA standard vs. its concentration in µg/mL.

NanoDrop (ND-1000 Spectrophotometer)

The NanoDrop® ND-1000 UV-Vis Spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA) was used to obtain a more feasible quantification of the protein. Firstly, the pedestal was cleaned with distilled water, and 2 µL of Milli-Q water was applied. The blank was prepared with 20 mM HEPES pH 7.4 buffer solution containing 10 mM NaCl (2 µL), and finally, 2 µL of each sample to be quantified was applied, with the pedestal being cleaned between applications. This equipment reads at an absorbance of 280 nm and provides the protein concentration in mg/mL.

2.2.6. *SDS-PAGE analysis (gel electrophoresis)*

Protein samples were also subjected to an 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), for 9 min at 92 ºC. Subsequently, these samples were applied to

12.5% polyacrylamide gels. 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 140 V and room temperature. The gels were stained with Coomassie Brilliant Blue (50% methanol, 10% acetic acid and 0.2% Brilliant Blue R) followed by distaining with a solution of 25% methanol and 5% acetic acid. Gel images were acquired using a VWR imager (VWR international, Pensylvania, EUA). The apparent molecular weight of proteins was determined by comparison with Precision Plus Protein^{TM} Unstained Protein Standards (Bio-Rad).

3. Results

This chapter presents and provides insights into the results obtained using different cell lysis methods. The goal of this study was to investigate how different cell lysis methods affect the recovery of the dimeric/active form of APRc. To achieve this, the recombinant soluble domain was expressed in *E. coli* and the cell extracts were subjected to various treatments, depending on the specific test conducted (as described in section 2.2.4. and are summarized in Table 3). The primary objective was to enhance the yield of APRc in the extracted soluble fraction.

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

In this section, it is provided a description of the methods applied and corresponding results, following chronological order in which they were performed. This approach will facilitate a better understanding of the influences of the changes made in the process.

3.1.1. *Mechanical lysis (High-Pressure Homogenizer) – "standard" assay*

To perform this assay, the goal was to establish a standard method and use it as a control for comparison with the modifications introduced in the following assays, in order to optimize the cell lysis method. Therefore, it was started by using the lysis conditions previously established in the laboratory, which is called the "standard" assay. These conditions consisted of cell disruption with Emulsiflex^{R}-C3 plugged into a pressure of 4 bar of nitrogen. As can be seen in Table 4, the pressure placed on the sample/pellet was 1700 bar, and the equipment had limitations, such as the use of a minimum volume of 30 mL and the passage of the pellet at least 3 times (3 cycles of compression). The sample was diluted in the last passage in Emulsiflex^{R}-C3 in 20 mM sodium phosphate buffer (pH 7.4) containing 10 mM imidazole and 500 mM NaCl, leaving a final 50 mL volume of total lysate.

Table 4 - Emulsiflex^{R}-C3 "standard" conditions.

Pressure placed on the system (bar)	Pressure placed on the sample (bar)	Minimum volume (mL)	Number of cycles
4	1700	30	3

After centrifugation, the total soluble fractions were applied onto a Histrap HP 5 mL column, and protein elution was carried out by implementing a four-step gradient of imidazole, as mentioned in section 2.2.4. One of the applications of imidazole is the purification of His-tagged proteins (such as APRc) using immobilized metal affinity chromatography (IMAC)

(Sigma-Aldrich, 2021). The buffer with imidazole was passed through the column, competing with Ni^{2+}, to prevent non-specific protein binding and freeing the His-tagged proteins from the column. Chromatograms of IMAC-Ni^{2+} purification of $APRc_{110-231}$-HisShort obtained after the mechanical lysis under "standard" conditions are presented in Figure 3 (A).

Each chromatogram shows the monitoring of protein elution at an absorbance of 280 nm and the respective imidazole gradient. As described by Cruz *et al*., from previous purifications of $APRc_{110-231}$-HisShort, it was observed that the fractions eluted with 200 mM imidazole corresponded to the dimeric forms of APRc and the proteins eluted at 150 mM imidazole were monomers (Cruz et al., 2014).

The fractions were individually fractionated and dialyzed. After dialysis, the two fractions were injected to a cation-exchange chromatography, as described previously in section 2.2.4. The APRc is positively charged at pH 7.4, binding to the negatively charged column due to ion exchange mechanism, allowing the elution of the proteins (Bio-Rad, 2021). The charged salt ions compete with bound proteins for the charged resin functional groups, with increasing NaCl concentration, causing protein elution (Bio-Rad, 2021). Representative chromatograms of the cation-exchange purifications of $APRc_{110-231}$-HisShort (150 mM and 200mM fractions) for mechanical lysis, under "standard" conditions, are presented in Figure 3 - (B) and (C) graphics. Each chromatogram shows protein elution at an absorbance of 280 nm and the respective NaCl concentration given by the buffer conductivity. Samples from the different purification steps were also analysed by SDS-PAGE followed by Coomassie blue staining, being the respective gel images presented in the Figure 3 (D).

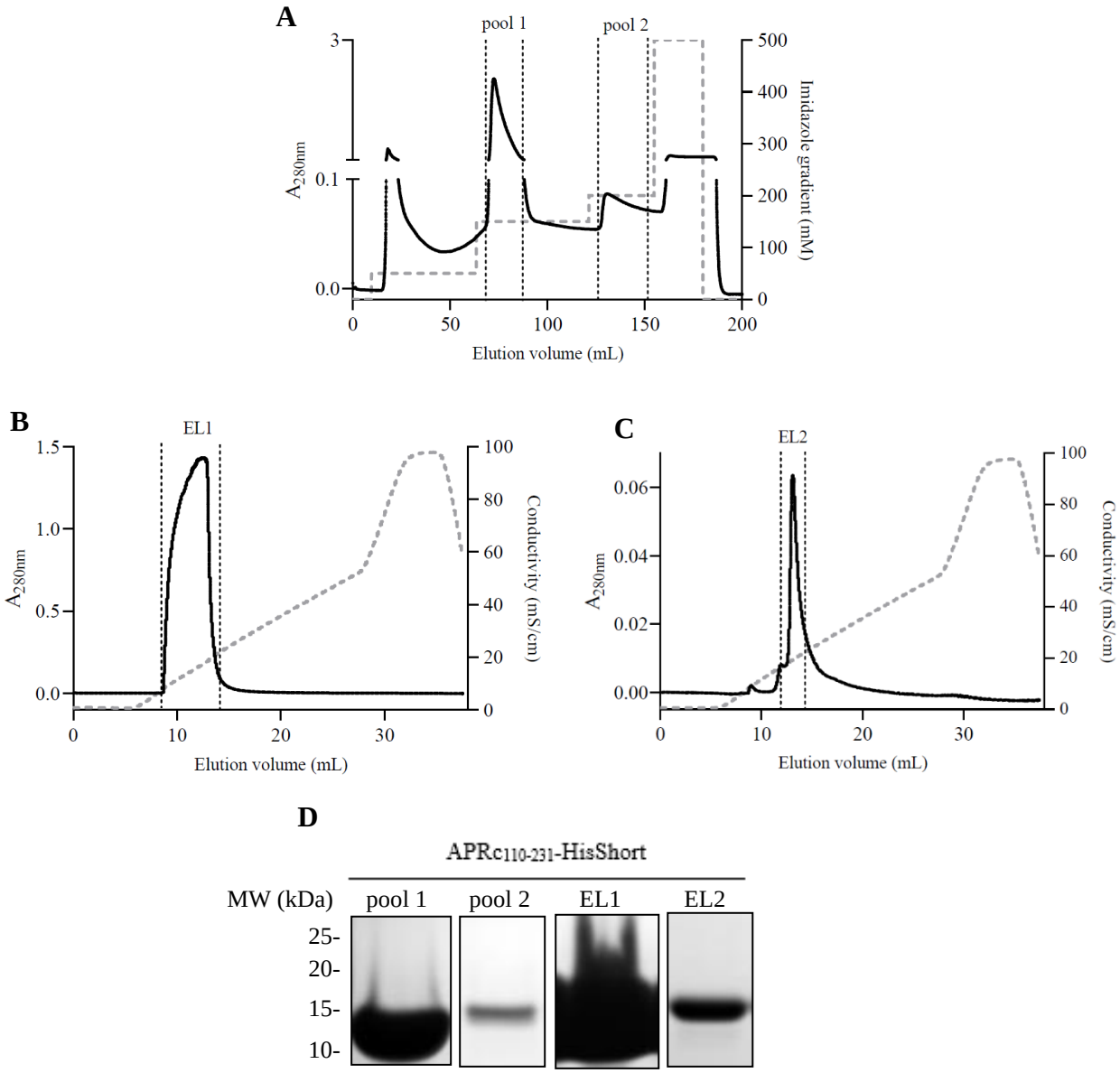
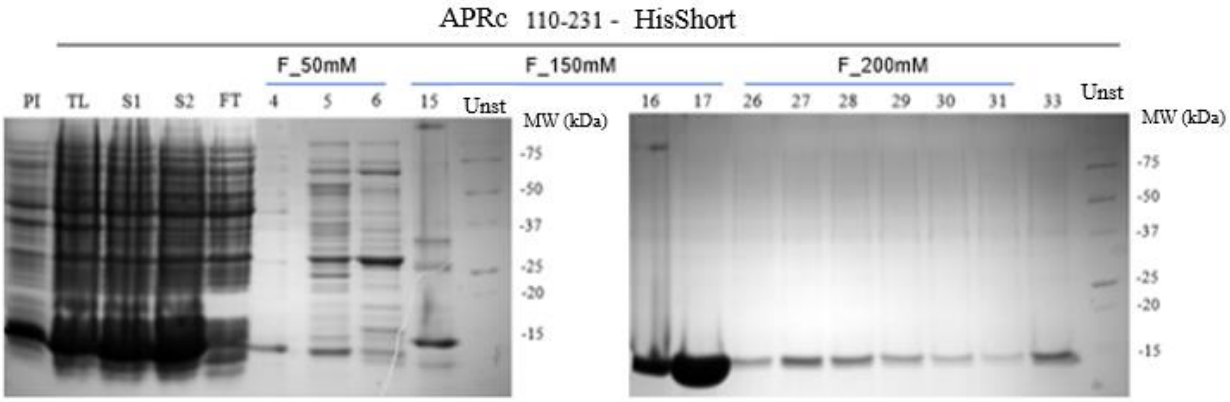


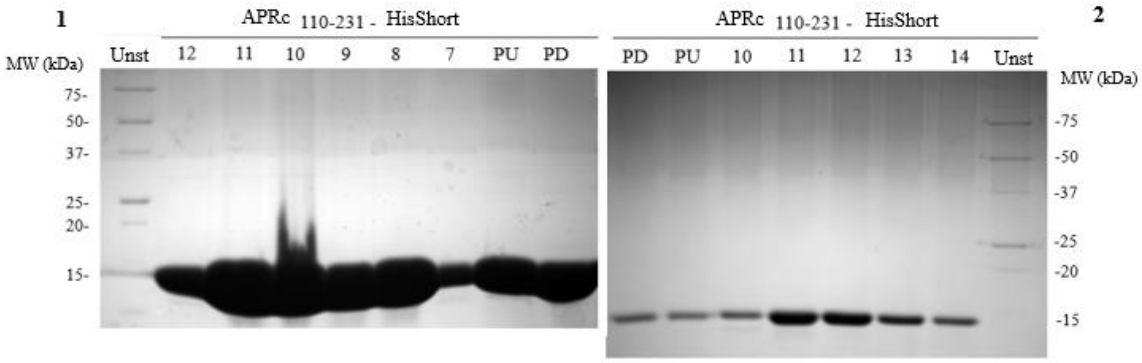
Figure 3 - Purification of APRc$_{110-231}$-HisShort by IMAC Ni^{2+} and cation-exchange chromatography, with APRc produced by mechanical lysis under "standard" conditions. **(A)** The recombinant protein APRc$_{110-231}$-HisShort was purified by IMAC Ni^{2+} on a Histrap HP 5 mL column, previously equilibrated 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 150 mM and 200 mM imidazole gradient steps were independently pooled (pool 1 and pool 2 – A) and dialyzed overnight against 20 mM HEPES pH 7.4 buffer. **(B) (C)** After dialysis, pools 1 and 2 were purified individually 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 fractions (EL1 and EL2) are outlined by dashed lines. **(D)** SDS-PAGE analysis followed by Coomassie Blue staining of the pooled eluted fractions (pool 1 and 2) resultant from the purifications by IMAC Ni^{2+}, and the selected eluted fractions (EL1 and EL2) 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.

Next, a total screening of all the different samples was performed to obtain a better visualization of the eluted proteins in the different steps of the process for a qualitative assessment of the protein. As previously reported by Cruz *et al*., APRc has a molecular weight of 14 kDa, which allows its identification by finding a band in that MW, as well as if the APRc that is being recovered is more or less pure, depending how "clean" are the samples. Initially, 20 µL of each of the fractions collected during Histrap chromatography were applied to SDS-PAGE to obtain a better insight into the separation of proteins from the column as the imidazole concentration increased. The quantity of protein was normalized to volume. This was done because with 20 µL of sample, it is not known how much protein is being used for analysis, not being a viable test to achieve the intended goal of this project. Thus, 6 µg of protein from each sample was first quantified using either the Pierce™ 660 nm protein assay or NanoDrop method, as detailed in section 2.2.5., and then applied to each well of an SDS-PAGE. All the SDS-PAGE performed for this assay are presented in Figure 4 and 5 (20 µL of each sample was applied) and Figure 6 (6 µg of total protein was applied per well).



Legend: PI – Total *E.coli* extract post-induction; TL – Total *E.coli* lysate; S1 – Soluble fraction; S2 – Clarified soluble fraction; FT – Flow through; 4, 5, 6 – Fractions eluted at 50 mM imidazole; 15, 16, 17 – Fractions eluted at 150 mM imidazole; 26 a 31 – Fractions eluted at 200 mM imidazole; 33 – Fraction eluted at 500 mM imidazole; Unst - Molecular weight marker.

Figure 4 - SDS-PAGE from the trial with mechanical lysis under "standard" conditions – Histrap Chromatography. Initially, 20 µL of each of the fractions collected along the Histrap were applied to SDS-PAGE to get a better insight into the separation of proteins from the column as the imidazole concentration increases. The protein of interest, APRc, is known to have a molecular weight of approximately 14 kDa, and it is found that along Histrap there is greater removal of this protein at a concentration of 150 mM imidazole.



Legend 1: Unst - Molecular weight marker; PD – Post-dialysis of the fraction eluted at 150 mM Imidazole; PU – Post-ultracentrifugation of the fraction eluted at 150 mM Imidazole; 7 a 12 – Fractions eluted from Mono S column.
Legend 2: PD – Post-dialysis of the fraction eluted at 200 mM Imidazole; PU – Post-ultracentrifugation of the fraction eluted at 200 mM Imidazole; 10 a 14 – Fractions eluted from Mono S column; Unst - Molecular weight marker.

Figure 5 - SDS-PAGE from the trial with mechanical lysis under "standard" conditions – Mono-S Chromatography. The images 1 and 2 refer to the pools previously removed from 150 mM imidazole fraction and 200 mM imidazole fraction, respectively. 20 µL of each of the fractions collected along the Mono-S were applied to SDS-PAGE to get a better insight into the separation of proteins from the column as the NaCl concentration increases. The protein of interest, APRc, is known to have a molecular weight of approximately 14 kDa, so only bands at this molecular weight are observed in the SDS-PAGE because in the Mono-S chromatography there is already a greater purification of APRc, and cleaner bands can be observed.

Initially, 20 µL of each of the fractions collected along the Histrap were applied to SDS-PAGE to get a better insight into the separation of proteins from the column as the imidazole concentration increases. Similarly, 20 µL of each of the fractions collected along the Mono-S were applied to SDS-PAGE to get a better insight into the separation of proteins from the column as the NaCl concentration increases. The protein of interest, APRc, is known to have a molecular weight of approximately 14 kDa, and it is found that along Histrap there is greater removal of this protein at a concentration of 150 mM imidazole. In case of Mono S, only bands at this molecular weight are observed in the SDS-PAGE because in the Mono-S chromatography there is already a greater purification of APRc, and cleaner bands can be observed. So, to make changes to the cell lysis method, it was categorical to evaluate the yields of protein and humid cell mass obtained in the different trials. For this, it was normalized the amount of total protein applied per well to better observe which proteins were present in each sample applied. Total protein was quantified using the PierceTM 660 nm protein assay, and 6 µg of total protein were applied to each well. Therefore, it was possible to see in which samples there was a higher amount of the protein of interest (APRc), or where this protein was present, as shown in Figure 6. Relative to the yields of protein and humid cell mass obtained in each assay, the results are presented in a further section.

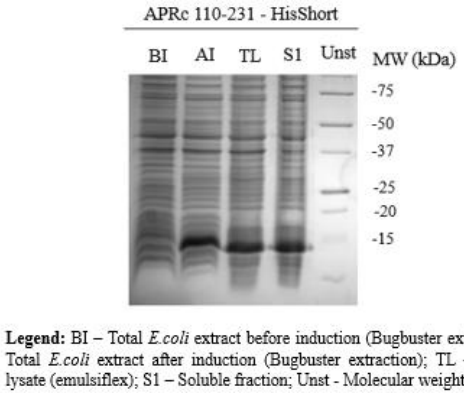


Figure 6 - SDS-PAGE from trial with mechanical lysis under "standard" conditions – first protein quantification. The samples applied were first quantified by Pierce™ 660 nm protein assay and then 6 µg of total protein were applied by well. This allows a better visualization of the presence of APRc.

As shown in Figure 6, and as expected, prior to induction with IPTG there is no production of APRc. After induction, the APRc appeared in the soluble phase. In the samples from before and after induction, the cell lysis of the sample, in this and all other tests performed, was performed using 100 µL of BugBuster - Protein Extraction Reagent, because only a sample from the trial was used for protein quantification and identification (on SDS-PAGE). Additionally, it is notable that the total *E. coli* lysate and the soluble fraction, obtained after the first purification, did not seem to present differences in protease extraction, presenting the same amount of cell debris and total protein.

3.1.2. *Physicochemical lysis – initial trial*

After the first evaluation using the standard cell lysis method, there was no difference in the amount of total protein extracted from the *E. coli* cell lysate and the total protein present in the soluble fraction, as it can be seen in Figure 6. Taking this into consideration and based on the results obtained by Professor Jorge Pereira and colleagues in its projects, it was decided to proceed with tests that did not use HPH to try to increase the protein present in the soluble fraction (Barro, A. M., 2020; Martins et al., 2018). In this way, it was decided to conduct several new trials testing physicochemical lysis because it has already shown promising results and because it is a less aggressive method to the cells than subjecting them to high pressures (as seen in section 1.4.).

First, it was established that this assay would not include the purification step, as the goal would be to test as many different physicochemical lysis conditions as possible, in order to try to experiment more conditions in less time, excluding those that were less promising than the assay already established in the laboratory - the "standard" assay - by comparing the protein

removed in each assay, analysing the quantification of total protein and the SDS-PAGE of each assay.

Therefore, after APRc expression in *E. coli*, four different triplicate tests were performed, dividing the pellet obtained from the expression. The cellular pellet (0.5 g) was placed in each of 12 falcon tubes, and then each pellet was resuspended in 5 g of solvent, testing different solvents, and different lysis conditions. Since the pellets could not be thawed at room temperature (~20°C) for the first time, the thawing step was increase to 37°C for total thawing of the sample. Thawing was performed at 20ºC for 44 min and at 37ºC for 29 min. A detailed description of the conditions of each triplicate assay is presented in Table 5.

Table 5 - Different conditions under which the physicochemical lysis tested in these triplicate assays.

Triplicate assay	Buffer	Freezing/thawing cycles
1	20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl	4
2	20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole, 500 mM NaCl and 250 mM SDS	4
3	20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole, 500 mM NaCl and 250 mM SDS	1
4	20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole, 500 mM NaCl and 500 mM SDS	1

Note that besides the application of freezing-thawing cycles (physical lysis), a surfactant, SDS (chemical lysis), was also added. The addition of SDS appears to be a solution, because previous studies have shown significant improvements in the combination of physical and chemical lysis. For instance, Martins *et al* have used freezing-thawing treatment combined with tensioactive compounds to successfully improve the purification of Green Fluorescent Protein (Martins et al., 2018). A surfactant is added to the standard buffer that is normally used for pellet resuspension (which is constituted by 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl) to help in the rupture of the cell membrane and in the removal of cellular debris because the SDS-PAGE of the standard assay shows that there are many other proteins released. Thus, it is intended to "clean" the gel using the surfactant

SDS, making the membrane more selectively permeable; however, it should be noted that this compound can denature the protein. It is important to mention that in this test, there was great difficulty in handling the pellets because the SDS made it very difficult to resuspend the pellet, mostly in the buffer with 500 mM SDS. The buffer with 500 mM SDS became quite thick, as there were quite a few difficulties in solubilizing this compound, which required heat to solubilize completely. This difficulty in solubilization contributed to the difficulty of thawing at room temperature, being only possible to thaw completely at 37ºC. This need for heating to thaw shows that the presence of SDS forms aggregates with the pellet; hence, it does not thaw at room temperature, revealing that the crystalline structure is stable. In addition, it should be kept in mind that thawing at 37ºC can cause changes in the entropy of the system because the higher the temperature, the higher is the entropy.

After the lysis step, samples were taken from the total lysate and the soluble fraction (obtained after centrifugation of the total lysate) and subsequently analysed in an SDS-PAGE. As mentioned in the previous section, in each SDS-PAGE, only 6 μg of total protein was applied per well. Also, using the program *ImageJ*, it was possible to analyse the ratios between $APRc_{110-231}$-HisShort and all lane, enabling a better visualization of the "cleaning" of SDS-PAGE, and they were calculated to evaluate if were indeed extracting more total soluble APRc. It is only possible to observe in terms of total soluble APRc because in this case it cannot be said that it is the dimer, because under reducing conditions of the SDS-PAGE, the molecular weight of the protein is the same, under monomeric or dimeric forms. This analysis allows a better comparison of different trials performed because it can make see the SDS-PAGE through numbers, and not just by looking at it. These results are presented in Figure 7, which compares the total and soluble fractions of each trial, and Figure 8, in which it is possible to evaluate the soluble fractions with different concentrations of SDS, enabling a better visualization of the influence of SDS on APRc.

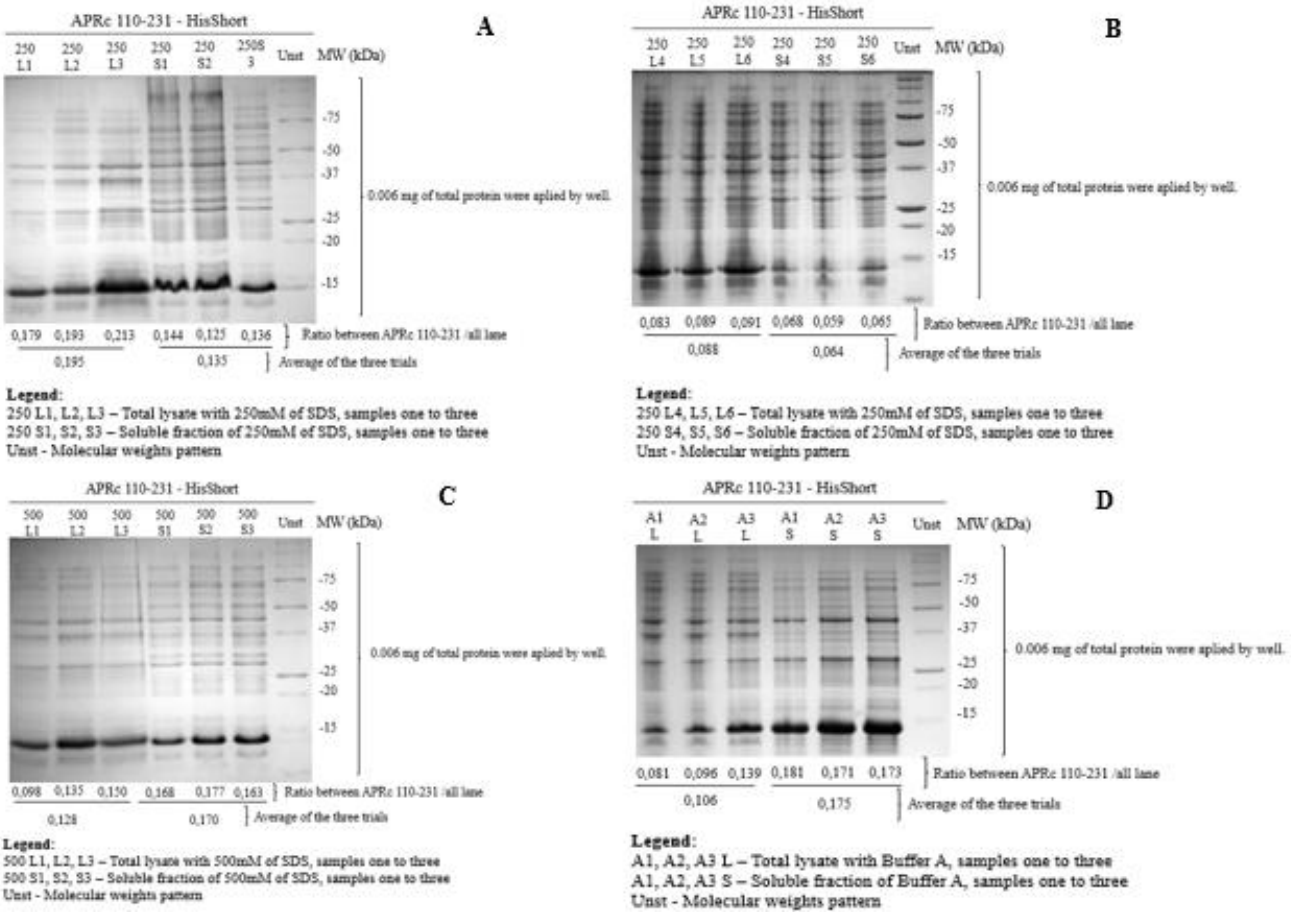


Figure 7 - SDS-PAGE from first trial with physicochemical lysis. The samples applied were first quantified by Pierce™ 660 nm protein assay and then 6 µg of total protein were applied by well. **(A)** Lysis with the standard buffer containing 250 mM SDS, and one cycle of freezing-thawing. Comparison between total and soluble fractions. **(B)** Lysis with the standard buffer containing 250 mM SDS, four cycles of freezing-thawing. Comparison between total and soluble fractions. **(C)** Lysis with the buffer containing 500 mM SDS, one cycle of freezing-thawing. Comparison between total and soluble fractions. **(D)** Lysis with standard buffer, four cycles of freezing-thawing. Comparison between total and soluble fractions.

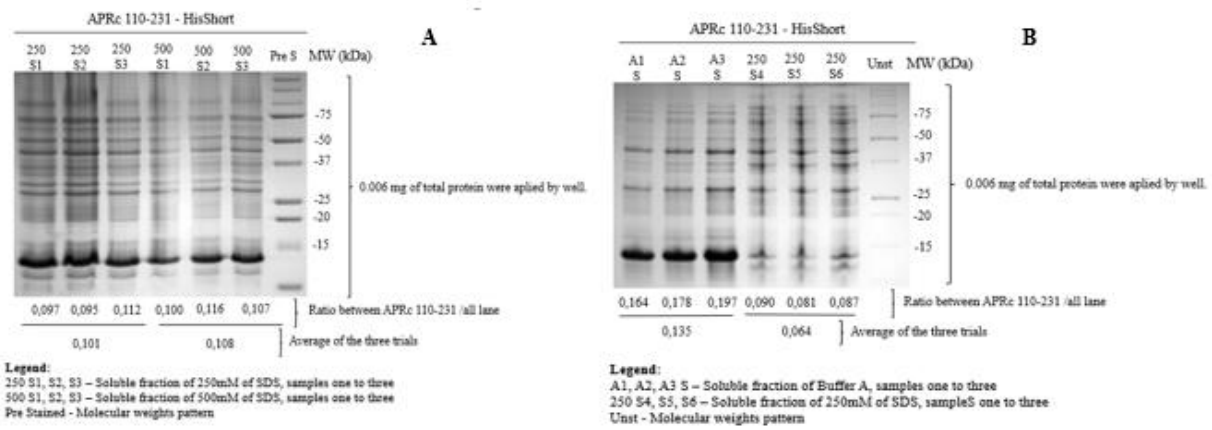


Figure 8 - Comparisons from SDS-PAGE from first trial with physicochemical lysis. The samples applied were first quantified by PierceTM 660 nm protein assay and then 6 μg of total protein were applied by well. **(A)** Soluble fractions from assays with standard buffer containing 250 mM SDS and 500 mM SDS, and one cycle of freezing-thawing. Comparison between soluble fractions of different concentrations of SDS. **(B)** Soluble fractions from assays with standard buffer and buffer containing 250 mM SDS, and four cycles of freezing-thawing. Comparison between soluble fractions of standard buffer and 250mM of SDS.

From Figure 7 it can be observed that by performing only one freezing-thawing cycle, more protein appears in the soluble phase. In addition, in the assay of standard buffer with 250 mM SDS and four freezing-thawing cycles, all APRc was lost compared to the assay in which only one cycle was performed. Moreover, in this assay, it is possible to conclude that the protein probably precipitated because it appeared in the total lysate sample, but not in the soluble fraction. The SDS-PAGE with better resolution, Figure 7 (D), corresponds to the assay with no SDS in the buffer; the lanes are clearer, and there is no protein drag along the lane, as in the assay with 500 mM SDS, Figure 7 (C).

By comparing the soluble fractions from different assays, it was possible to verify that between different concentrations of SDS, namely 250 mM and 500 mM, 500 mM presented clear lanes, but 250 mM appeared to have extracted more protein, since the bands were more loaded, mainly at the APRc molecular weight, as it can be seen in Figure 8 (A). Comparing the contrast between standard buffer and the presence of 250 mM SDS, it is clear that performing four cycles with 250 mM SDS is not a viable option for APRc extraction, as it is showed by Figure 8 (B).

Interestingly, some samples were not sufficiently solubilized to be sampled and applied to the gel. Therefore, it was necessary to heat the samples to enable solubilization and pipetting. It is unclear how this heating to solubilize SDS affected the release and stability of the protease and, in the case of SDS-PAGE, its running.

After performing this assay, and in comparison, with the previous one, it is possible to exclude the hypothesis of optimizing the process in the presence of SDS, because it seems that APRc is not being extracted or precipitated. However, with freezing-thawing cell disruption, there is a noticeable improvement in the cleanliness of the lanes, as this process makes membrane rupture less aggressive, and due to a cell permeabilization, releasing less protein contaminants and cellular debris. For the next steps, only trials based in physicochemical lysis with standard buffer were carried out as a hypothesis to improve the process. However, physicochemical lysis was used to study the influence of temperature on APRc, performing the next test of this method in duplicate, thawing one of the pellets only at 20 ºC and the at 37 ºC.

3.1.3. *Physicochemical lysis – study of the influence of temperature*

Based on the results obtained from the first trial of physicochemical lysis, this trial was performed using the standard buffer (20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl), and four freezing-thawing cycles. As mentioned before, since the pellets could not be thawed at room temperature (~20°C) for the first time, and the thawing step was increased to 37 °C, an assay was performed to study the influence of heating on obtaining APRc and its proteolytic activity. The thaw was performed at 20 ºC during 44 min and at 37 ºC with a first step equal at 20 ºC and another step at 37 ºC for 29 min. This study is interesting because there are many references to the denaturation of proteins caused by an increase in the temperature of the medium in which the target protein located.

So, after the cell lysis step was completed, the pellet was subjected to an increase up to 37 ºC, and a more fluid appearance was observed. However, after centrifugation of the total lysate, the soluble fraction became viscous, indicating that there was still DNA attached to the cells that had not been solubilized. Next, the total soluble fractions were applied to a Histrap column, and the above-mentioned fractions were independently pooled and dialyzed overnight, the two pools obtained were applied onto a cation-exchange column, as previously detailed in section 2.2.4. Chromatograms of IMAC-Ni^{2+} purifications of $APRc_{110-231}$-HisShort produced by physicochemical lysis are shown in Figures 9 and 10 (A). Representative chromatograms of the cation-exchange purifications of $APRc_{110-231}$-HisShort (150 mM and 200mM fractions) for physicochemical lysis, at 20 ºC and 37 ºC, are shown in Figure 9 and 10 (B and C graphics), respectively. Each chromatogram shows the protein elution at an absorbance of 280 nm and the respective NaCl concentration given by the buffer conductivity. Samples from the different purification steps were analysed by SDS-PAGE followed by Coomassie blue staining, as shown in Figures 9 and 10 (D). The complete and detailed screenings are presented below.

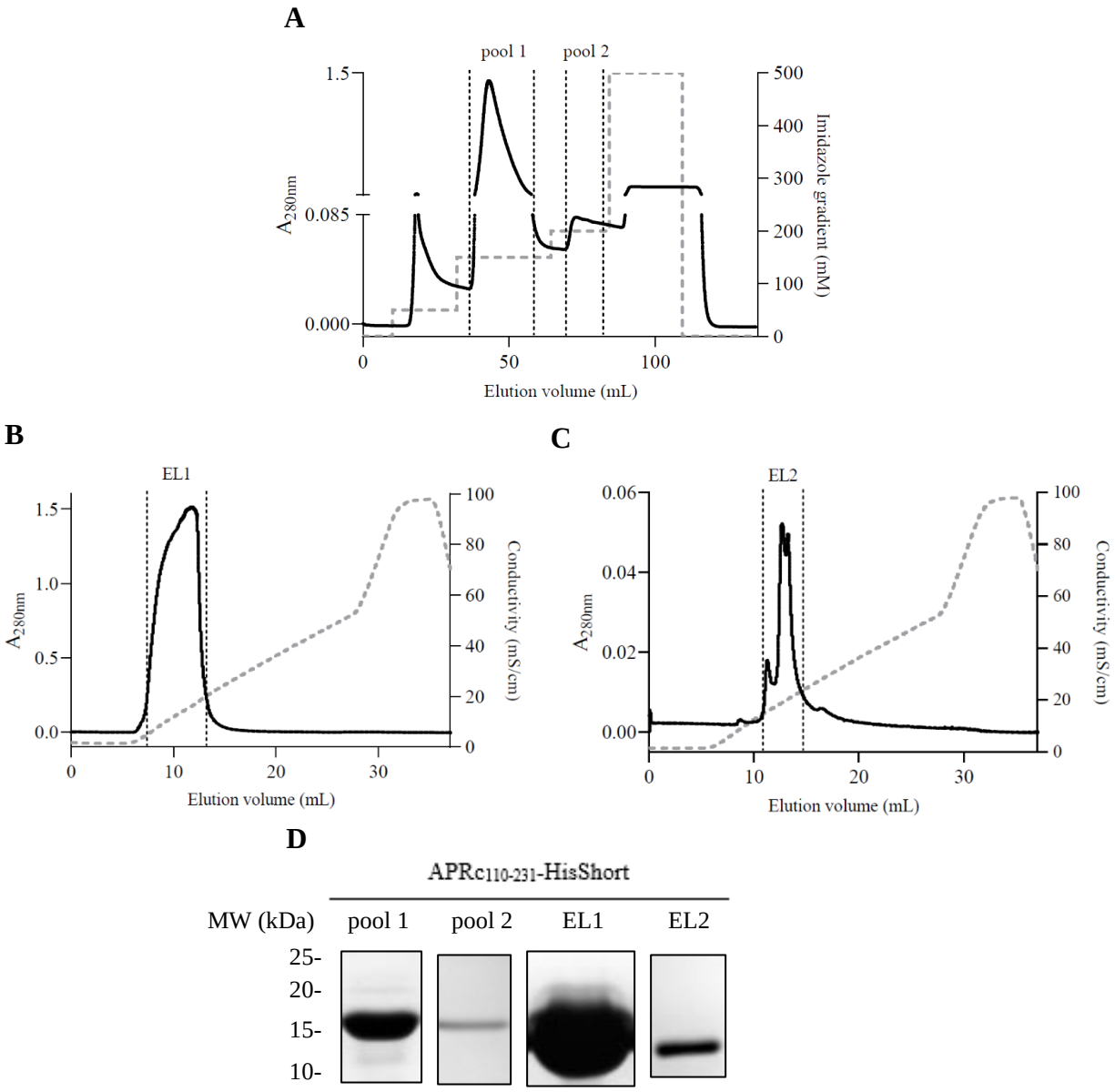


Figure 9 - Purification of APRc$_{110-231}$-HisShort by IMAC Ni^{2+} and cation-exchange chromatography, with APRc produced by physicochemical lysis, at 20ºC. (A) The recombinant protein APRc$_{110-231}$-HisShort was purified by IMAC Ni^{2+} on a Histrap HP 5 mL column, previously equilibrated in 20 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 150 mM and 200 mM imidazole gradient steps were independently pooled (pool 1 and pool 2 – A) and dialyzed overnight against 20 mM HEPES pH 7.4 buffer. **(B) (C)** After dialysis, pools 1 and 2 were purified individually 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 fractions (EL1 and EL2) are outlined by dashed lines. **(D)** SDS-PAGE analysis followed by Coomassie Blue staining of the pooled eluted fractions (pool 1 and 2) resultant from the purifications by IMAC Ni^{2+}, and the selected eluted fractions (EL1 and EL2) from the purification by cation-exchange chromatography. 6 µL of target protein of each sample were denatured with 6x loading buffer with DTT and then applied to a 12.5% polyacrylamide gel.

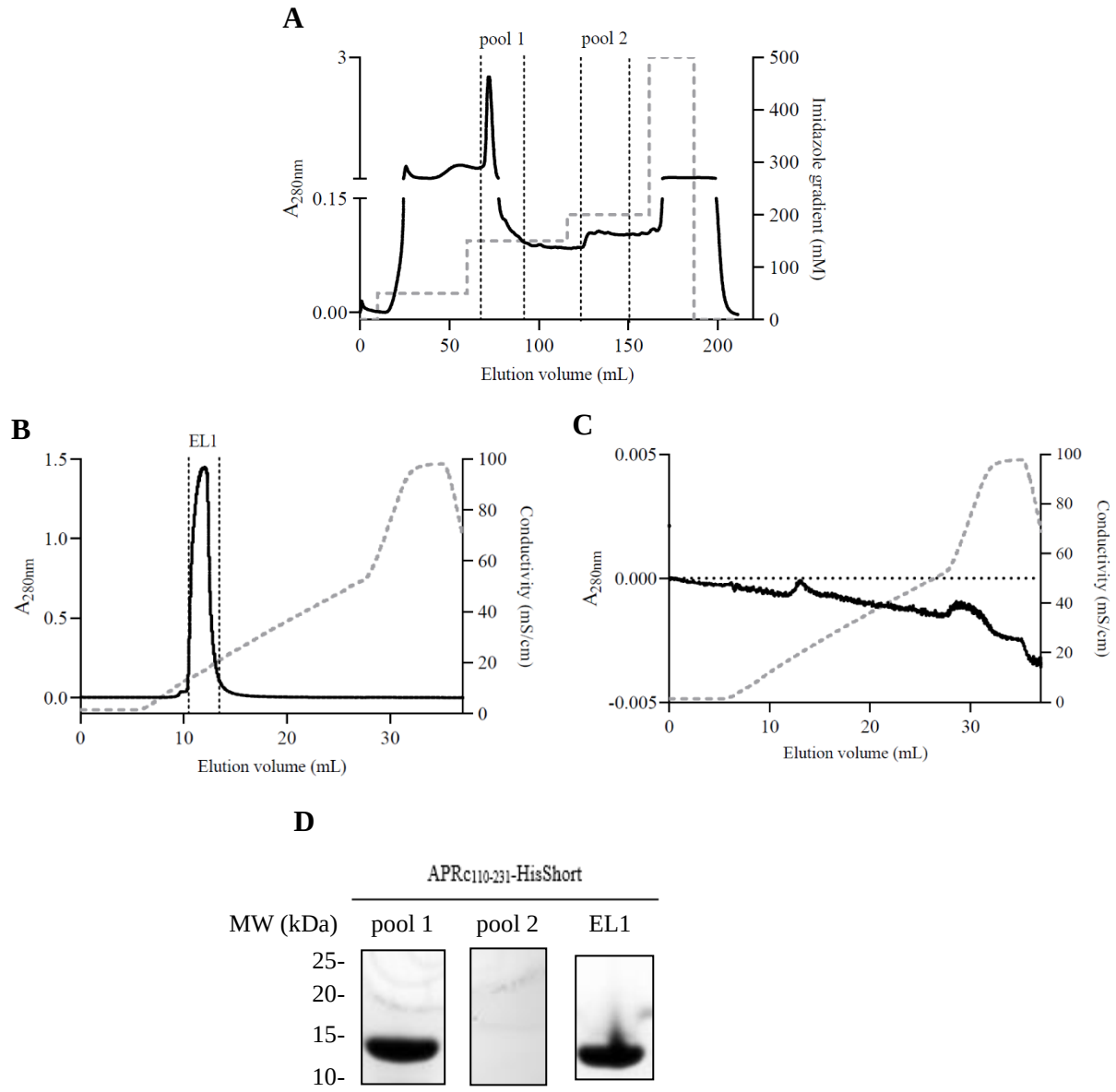


Figure 10 - Purification of APRc$_{110-231}$-HisShort by IMAC Ni^{2+} and cation-exchange chromatography, with APRc produced by physicochemical lysis, at 37ºC. (A) The recombinant protein APRc$_{110-231}$-HisShort was purified by IMAC Ni^{2+} on a Histrap HP 5 mL column, previously equilibrated in 20 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 150 mM and 200 mM imidazole gradient steps were independently pooled (pool 1 and pool 2 – A) and dialyzed overnight against 20 mM HEPES pH 7.4 buffer. **(B) (C)** After dialysis, pools 1 and 2 were purified individually 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 (EL1) is outlined by dashed lines. **(D)** SDS-PAGE analysis followed by Coomassie Blue staining of the pooled eluted fractions (pool 1 and 2) resultant from the purifications by IMAC Ni^{2+}, and the selected eluted fraction (EL1) from the purification by cation-exchange chromatography. 6 µL of target protein of each sample were denatured with 6x loading buffer with DTT and then applied to a 12.5% polyacrylamide gel.

By means of Figures 9 and 10, it is observed that in the 20 ºC test there seems to be a greater amount of dimer removed, in comparison with the standard test represented in section 3.1.1, Figure 3. However, it is verified that heating up to 37 ºC causes the loss of all dimeric form of the protein, which is not therefore a viable method to produce active APRc. Therefore, in the next assays in which physicochemical lysis were tested, the freezing-thawing stage must not exceed room temperature.

Next, a total screening of all different samples was performed for the same reasons mentioned in Section 3.1.1. Thus, 6 μg of protein from each sample was first quantified using either the Pierce™ 660 nm protein assay or NanoDrop, as detailed in Section 2.2.5., and then applied to each well of an SDS-PAGE. All SDS-PAGE images obtained for this assay are presented in Figures 11 to 13.

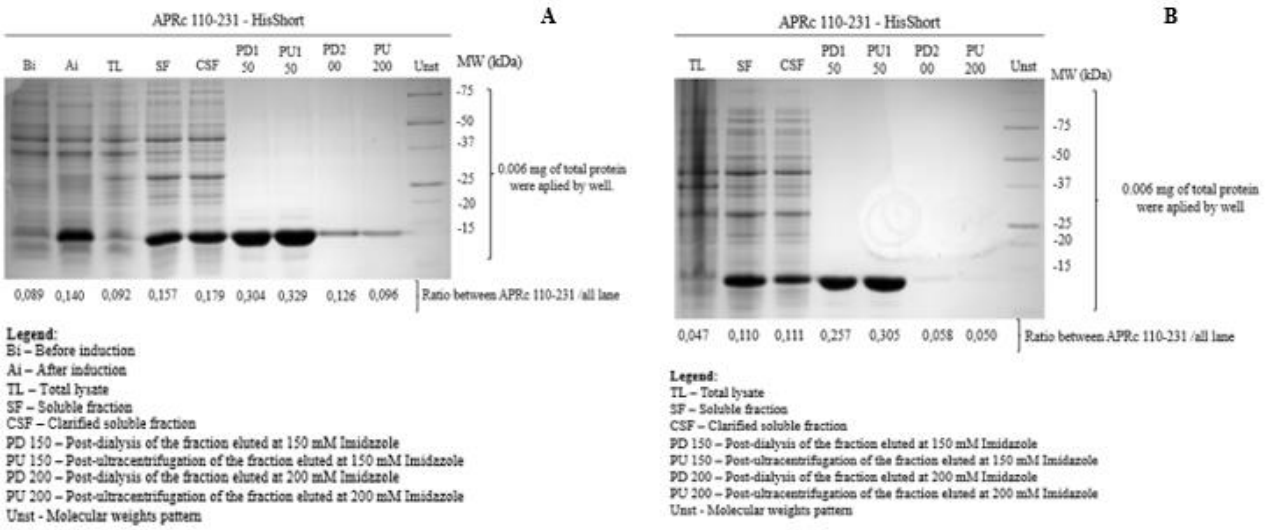


Figure 11 - SDS-PAGE from trial done with physicochemical lysis at 20 ºC (A) and 37 ºC (B) – study of the influence of temperature. The samples applied were first quantified by Pierce™ 660 nm protein assay and then 6 μg of total protein were applied by well. In the assay of 37 ºC **(B)** there is not before or after induction samples because they were impossible to manage due to high viscosity.

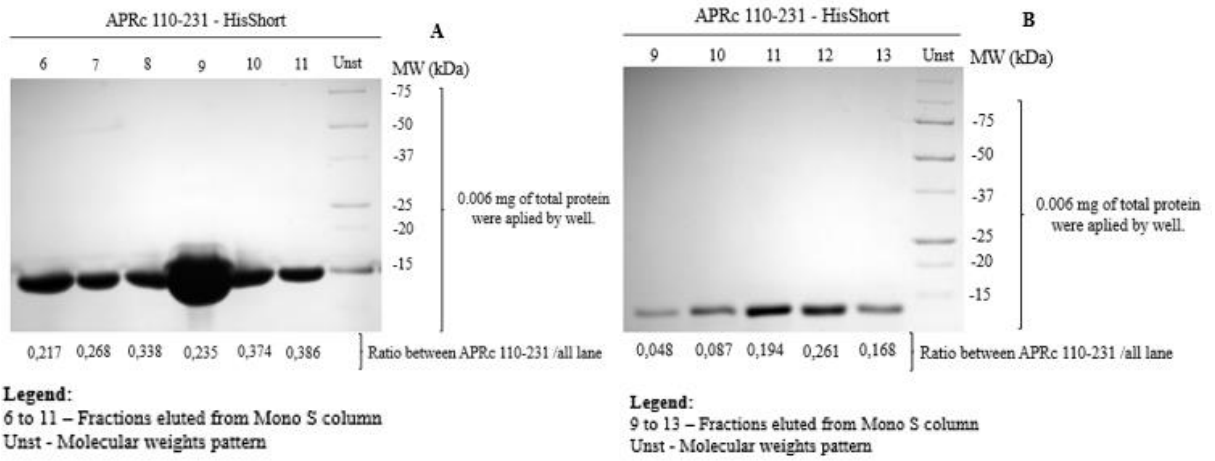


Figure 12 - SDS-PAGE from the trial done with physicochemical lysis at 20 ºC – Mono-S Chromatography. The images A and B refer to the pools previously removed from 150 mM imidazole fraction and 200 mM imidazole fraction, respectively. The samples applied were first quantified by Pierce™ 660 nm protein assay and then 6 µg of total protein were applied by well.

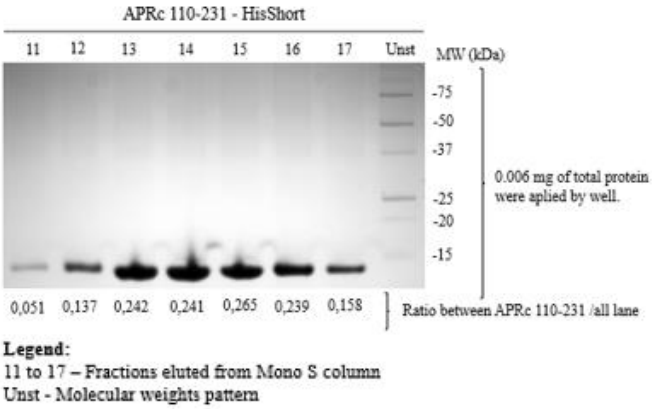


Figure 13 - SDS-PAGE from the trial done with physicochemical lysis at 37 ºC – Mono-S Chromatography. The image refers to the pools previously removed from 150 mM imidazole fraction. The samples applied were first quantified by Pierce™ 660 nm protein assay and then 6 µg of total protein were applied by well.

Figure 11 confirms that there are larger amounts of protein than normal based on the intensity of the bands. Recall that, as shown in Figure 12 (of Mono S), the peaks had greater intensity with this method, and therefore, a greater amount of protein. In Figure 11 (B), SDS-PAGE confirmed the loss of functional APRc at 37 ºC, and the lower quantity obtained when compared to 20 ºC – *cf.* Figures 11, 12 and 13. Regarding these observations, all following physicochemical lysis assays were then performed at 20 ºC.

3.1.4. *Physicochemical lysis – final trials*

Two more physicochemical lysis tests were performed to verify that there is no optimization of the APRc extraction by these methods, regardless of the changing conditions. So, 1 cycle of freezing-thawing was applied, using a buffer with a smaller concentration of SDS (20 mM sodium phosphate pH 7.4 buffer containing 10 mM imidazole, 500 mM NaCl and 100 mM SDS) and the standard buffer (20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl).

After the cell lysis step, it was observed that the pellet with 100 mM SDS buffer was almost impossible to handle, possibly due to the release of too much intracellular cell content, suggesting that they are mainly contaminants. However, the total soluble fractions were applied to a Histrap column, as previously detailed in Section 3.1.1. Even though purification was carried out with some difficulties, the samples were characterized by SDS-PAGE. Therefore, a total screening of all samples was performed. Hence, 6 μg of protein from each sample was first quantified using either the Pierce™ 660 nm protein assay or NanoDrop, and then applied to each well of an SDS-PAGE. The SDS-PAGE performed for this assay is presented in Figure 14.

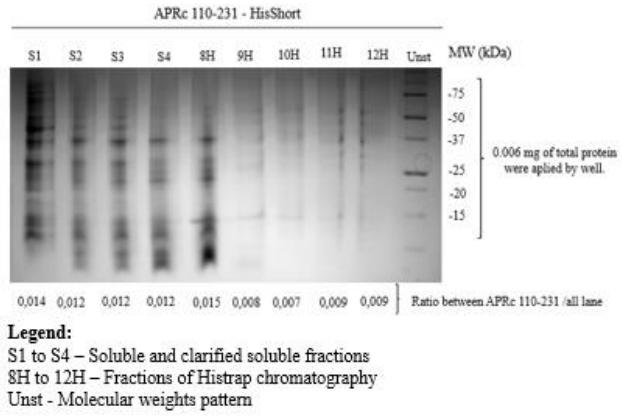


Figure 14 - SDS-PAGE from trial done with physicochemical lysis at 20 ºC with one cycle of freezing-thawing and 100mM SDS buffer. The samples applied were first quantified by PierceTM 660 nm protein assay and then 6 μg of total protein were applied by well.

As shown in Figure 14, the assay with a buffer containing 100 mM SDS, with only one freezing-thawing cycle, in the soluble fractions, there appeared to be no APRc, but the first purification step (with Histrap chromatography) was performed and no protein was eluted for further steps; therefore, this hypothesis was excluded.

As mentioned at beginning of this section, an assay applying one cycle of freezing-thawing with the standard buffer (20 mM sodium phosphate pH 7.4 buffer, containing 10 mM

imidazole and 500 mM NaCl) was performed as well, since it was only tested with four cycles of freezing-thawing, and it is intended to exclude all the hypotheses.

After the cell lysis step, the total soluble fractions were purified as explained previously in section 3.1.1. Chromatograms of IMAC-Ni^{2+} purifications of $APRc_{110-231}$-HisShort produced by physicochemical lysis are shown in Figure 15 (A). Representative chromatograms of the cation-exchange purifications of $APRc_{110-231}$-HisShort (150 mM and 200mM fractions) for this physicochemical lysis assay are shown in (B) and (C) graphics of Figure 15. Each chromatogram shows the protein elution at an absorbance of 280 nm and the respective NaCl concentration given by the buffer conductivity. Samples from the different purification steps were analysed by SDS-PAGE followed by Coomassie blue staining, as shown in the (D) graphic of Figure 15. The complete and detailed screenings are presented below.

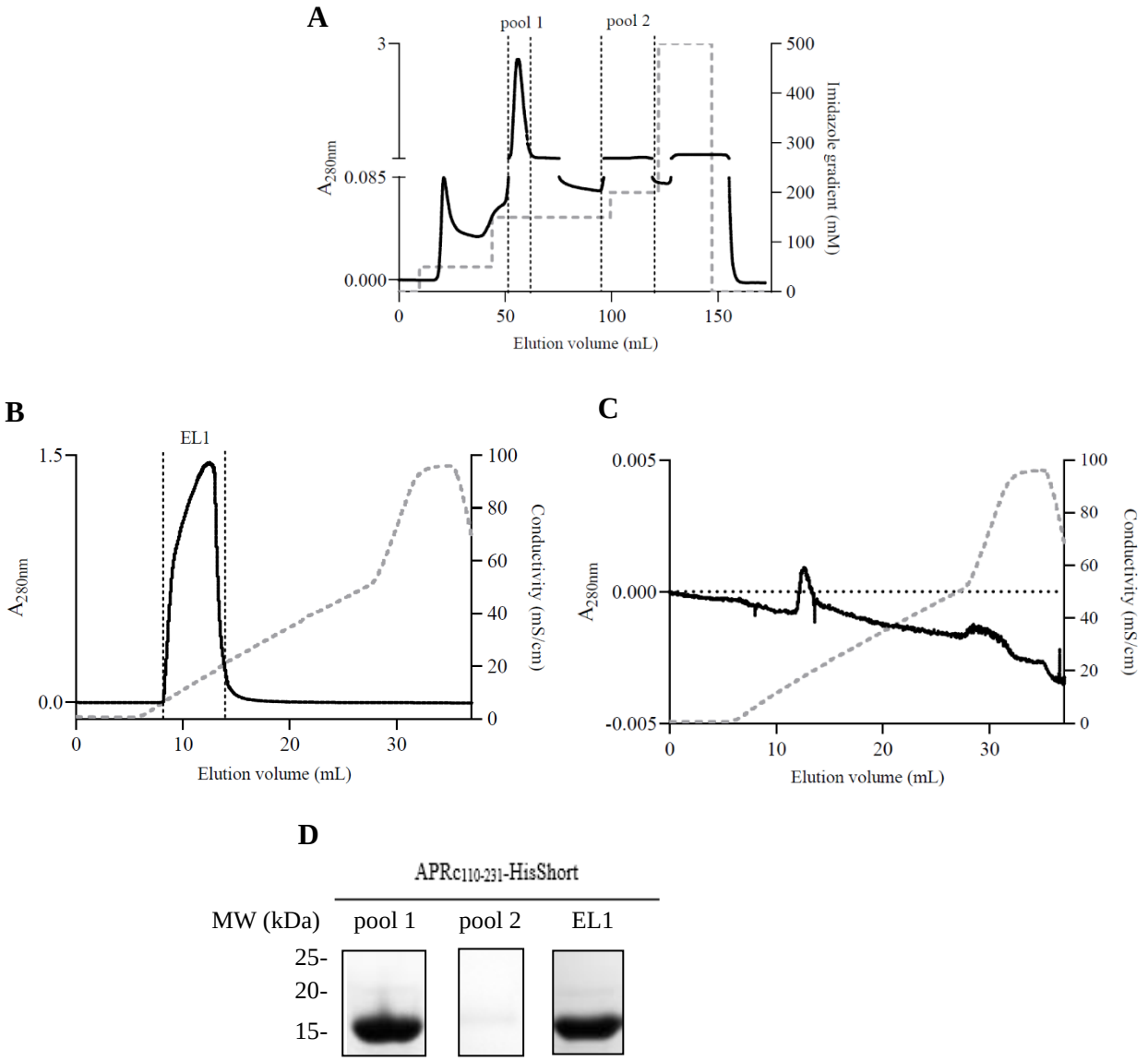


Figure 15 - Purification of APRc$_{110-231}$-HisShort by IMAC Ni^{2+} and cation-exchange chromatography, with APRc produced by physicochemical lysis, with just one cycle of freezing-thawing. (A) The recombinant protein APRc$_{110-231}$-HisShort was purified by IMAC Ni^{2+} on a Histrap HP 5 mL column, 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 150 mM and 200 mM imidazole gradient steps were independently pooled (pool 1 and pool 2 – A) and dialyzed overnight against 20 mM HEPES pH 7.4 buffer. **(B) (C)** After dialysis, pools 1 and 2 were purified individually 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 (EL1) is outlined by dashed lines. **(D)** SDS-PAGE analysis followed by Coomassie Blue staining of the pooled eluted fractions (pool 1 and 2) resultant from the purifications by IMAC Ni^{2+}, and the selected eluted fraction (EL1) from the purification by cation-exchange chromatography. 6 µL of target protein of each sample were denatured with 6x loading buffer with DTT and then applied to a 12.5% polyacrylamide gel.

Next, a total screening of all the different samples was performed, for the same reasons mentioned at section 3.1.1. Thus, 6 µg of protein from each sample was first quantified using either the Pierce™ 660 nm protein assay or the NanoDrop, as detailed in section 2.2.5., and then applied to each well of the SDS-PAGE. All the SDS-PAGE performed for this assay are presented in Figure 16.

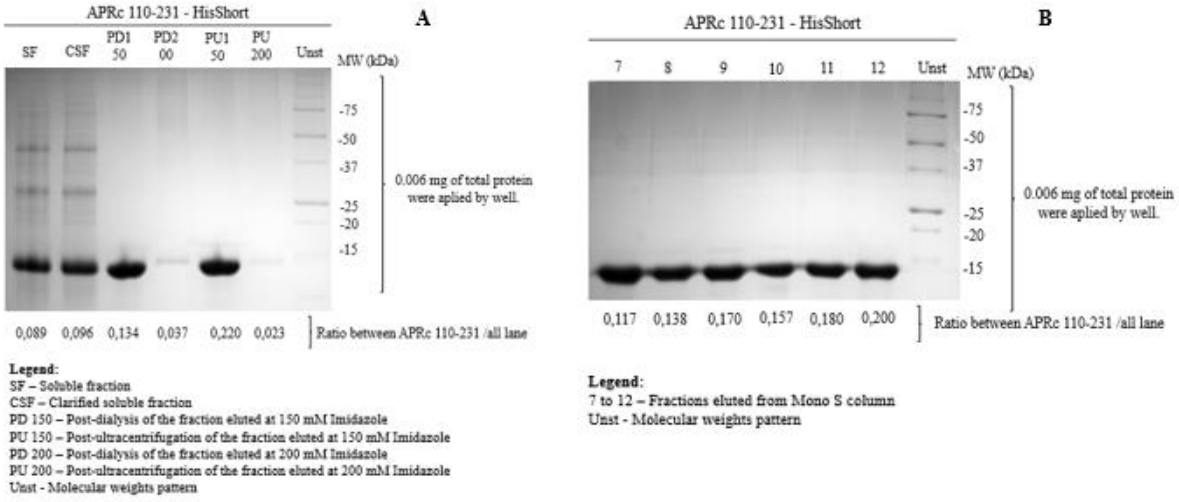


Figure 16 - SDS-PAGE from the trial done with at 20 ºC with one cycle of freezing-thawing and standard buffer. (A) Represents the general samples from all the steps before the first purification step **(B)** Represents the samples of soluble fractions eluted at 150 mM SDS, from the Mono-S chromatography. The samples applied were first quantified by Pierce™ 660 nm protein assay and then 6 µg of total protein were applied by well.

In both Figure 15 and 16, regarding purification and SDS-PAGE, respectively, it can be seen that this assay is not feasible, as no dimer is removed in the 200 mM imidazole fraction in Mono-S chromatography, and there are no APRc in the corresponding lane of SDS-PAGE.

3.1.5. *Mechanical lysis (High-Pressure Homogenizer) - modified assays*

After evaluating all physicochemical methods and understanding that there was no optimization of the process to obtain the protein through them, it was decided to focus on the process that was previously established in the laboratory. Furthermore, the next steps consist of modifying the operating conditions of the HPH, such as the pressure applied to the sample and the number of cycles of compression. Following the tests performed up to this stage, this seems to be the only method that favours the accumulation of the protein in its active form. These modifications are listed in Table 6.

Table 6 - Conditions used in HPH, seeking the optimization of this method.

Pressure placed on the system (bar)	Pressure placed on the sample (bar)	Number of cycles
4	0	3
4	1000 - 1250	6

Note that only the number of cycles and pressure applied to the sample were changed. Thus, a trial was performed in which no pressure was applied to the sample, and the pellet was passed only in the narrowing of the equipment, three times (3 cycles). In the following tests, the pellet was first passed in the narrowing and the pressure was gradually increased, until it was verified that the pellet was being exposed to a (minimum) pressure and subjected to six cycles at that pressure.

In these tests, pellet temperature was measured before and after each compression cycle. In the "standard" assay, the temperature varied about 12ºC in each cycle, and in the assay in which changes were made (6 cycles at lower pressure) the temperature variation was about 6ºC. Knowing this temperature variation was important because it was hypothesized that the temperature could influence the amount of dimer obtained, that is, by applying pressure to break the cells there was a temperature increase that could denature the protein of interest.

After cell lysis, the total soluble fractions were purified, as previously explained in section 3.1.1. Chromatograms of IMAC-Ni^{2+} purification of $APRc_{110-231}$-HisShort produced by physicochemical lysis are shown in Figure 17 (A). Representative chromatograms of the cation-exchange purifications of $APRc_{110-231}$-HisShort (150 mM and 200mM fractions) for mechanical lysis, with and without pressure applied in the sample, are shown in (B) and (C) graphics of Figures 17 and 18. Each chromatogram shows the protein elution at an absorbance of 280 nm and the respective NaCl concentration given by buffer conductivity. Samples from the different purification steps were analysed by SDS-PAGE followed by Coomassie blue staining, as shown in the (D) graphics of each Figure (17 and 18). The complete and detailed screenings are presented below.

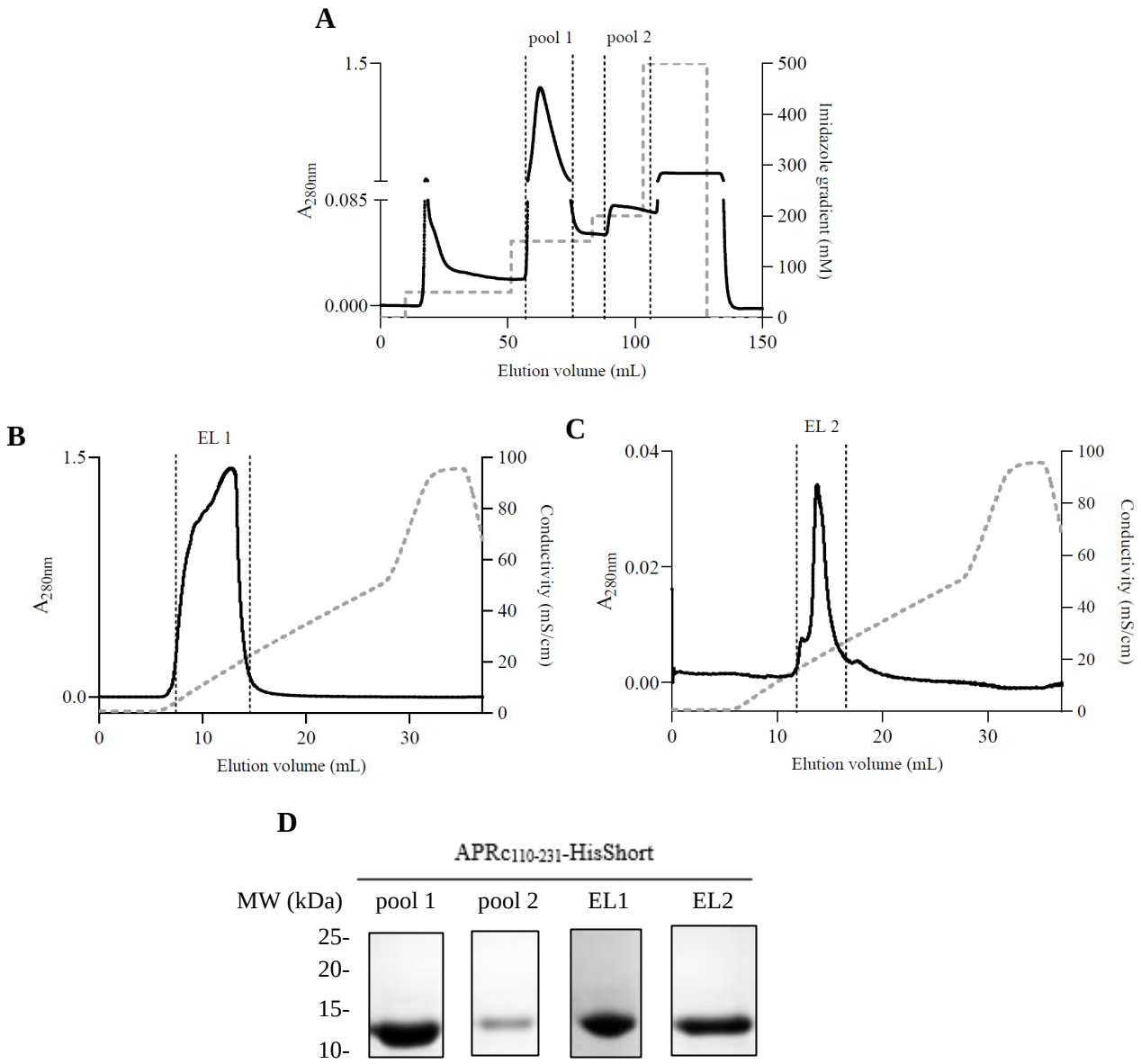


Figure 17 - Purification of APRc$_{110-231}$-HisShort by IMAC Ni^{2+} and cation-exchange chromatography, with APRc produced by mechanical lysis, under modified conditions (with no pressure applied in the sample). (A) The recombinant protein APRc$_{110-231}$-HisShort was purified by IMAC Ni^{2+} on a 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 150 mM and 200 mM imidazole gradient steps were independently pooled (pool 1 and pool 2 – A) and dialyzed overnight against 20 mM HEPES pH 7.4 buffer. **(B) (C)** After dialysis, pools 1 and 2 were purified individually 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 fractions (EL1 and EL2) are outlined by dashed lines. **(D)** SDS-PAGE analysis followed by Coomassie Blue staining of the pooled eluted fractions (pool 1 and 2) resultant from the purifications by IMAC Ni^{2+}, and the selected eluted fractions (EL1 and EL2) from the purification by cation-exchange chromatography. 6 µL of target protein of each sample were denatured with 6x loading buffer with DTT and then applied to a 12.5% polyacrylamide gel.

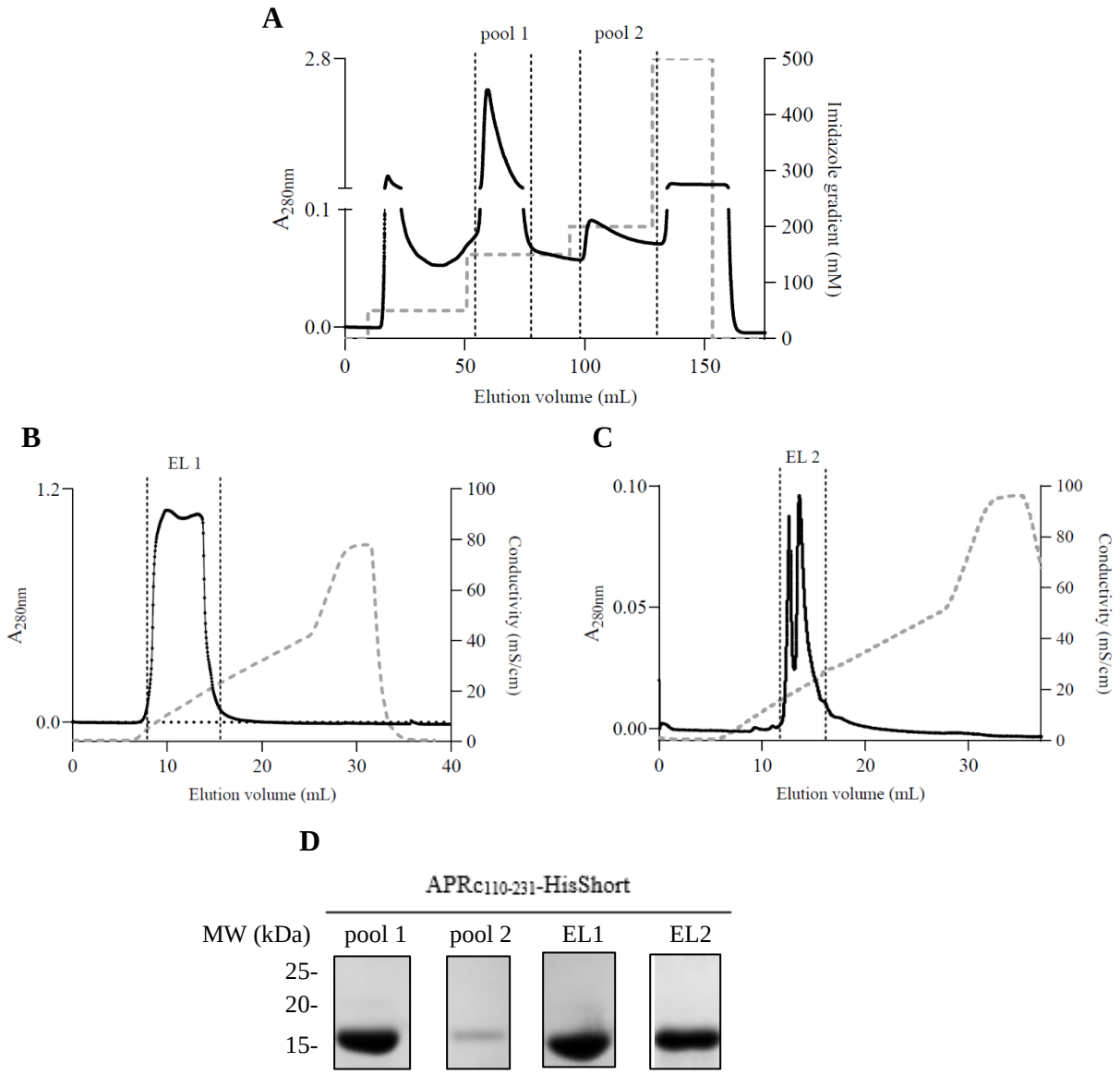


Figure 18 - Purification of APRc$_{110-231}$-HisShort by IMAC Ni^{2+} and cation-exchange chromatography, with APRc produced by mechanical lysis, under modified conditions (with pressure applied in the sample). (A) The recombinant protein APRc$_{110-231}$-HisShort was purified by IMAC Ni^{2+} on a Histrap 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 150 mM and 200 mM imidazole gradient steps were independently pooled (pool 1 and pool 2 – A) and dialyzed overnight against 20 mM HEPES pH 7.4 buffer. **(B) (C)** After dialysis, pools 1 and 2 were purified individually 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 fractions (EL1 and EL2) are outlined by dashed lines. **(D)** SDS-PAGE analysis followed by Coomassie Blue staining of the pooled eluted fractions (pool 1 and 2) resultant from the purifications by IMAC Ni^{2+}, and the selected eluted fractions (EL1 and EL2) from the purification by cation-exchange chromatography. 6 µL of target protein of each sample were denatured with 6x loading buffer with DTT and then applied to a 12.5% polyacrylamide gel.

Then, a total screening of all the different samples was performed, for the same reasons mentioned at section 3.1.1. Thus, 6 µg of protein from each sample was first quantified using either the NanoDrop, as detailed in section 2.2.5., and then applied to each well of the SDS-PAGE. All the SDS-PAGE performed for this assay are presented in Figures 19 to 22.

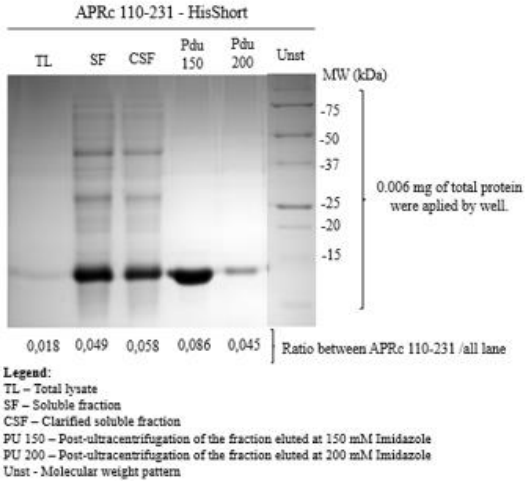


Figure 19 - SDS-PAGE from trial done with mechanical lysis at HPH, with no pressure applied in the sample and 3 cycles of compression. The samples applied were first quantified by NanoDrop (ND-1000 Spectrophotometer), and then 6 µg of total protein were applied by well.

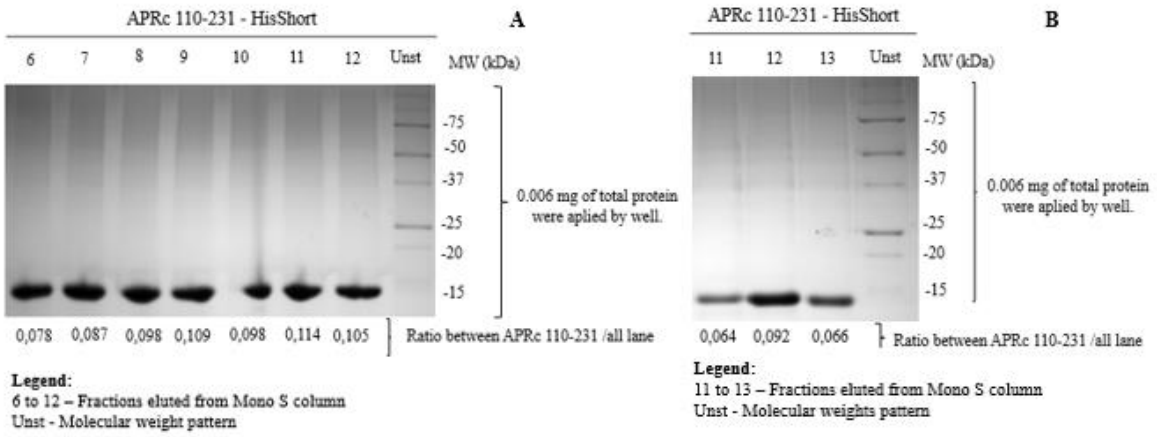


Figure 20 - SDS-PAGE from the trial done with mechanical lysis at HPH, with no pressure applied in the sample and 3 cycles of compression – Mono-S chromatography. The images **A** and **B** refer to the pools previously removed from 150 mM imidazole fraction and 200 mM imidazole fraction, respectively. The samples applied were first quantified by NanoDrop (ND-1000 Spectrophotometer), and then 6 µg of total protein were applied by well.

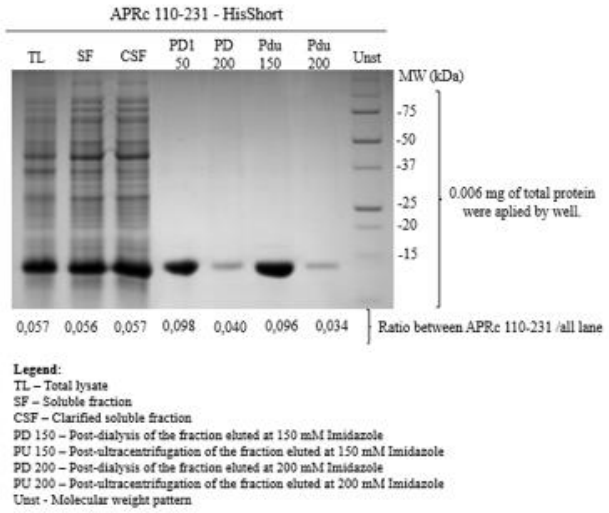


Figure 21 - SDS-PAGE from trial done with mechanical lysis at HPH, with minimum pressure applied in the sample and 6 cycles of compression. The samples applied were first quantified by NanoDrop (ND-1000 Spectrophotometer), and then 6 μg of total protein were applied by well.

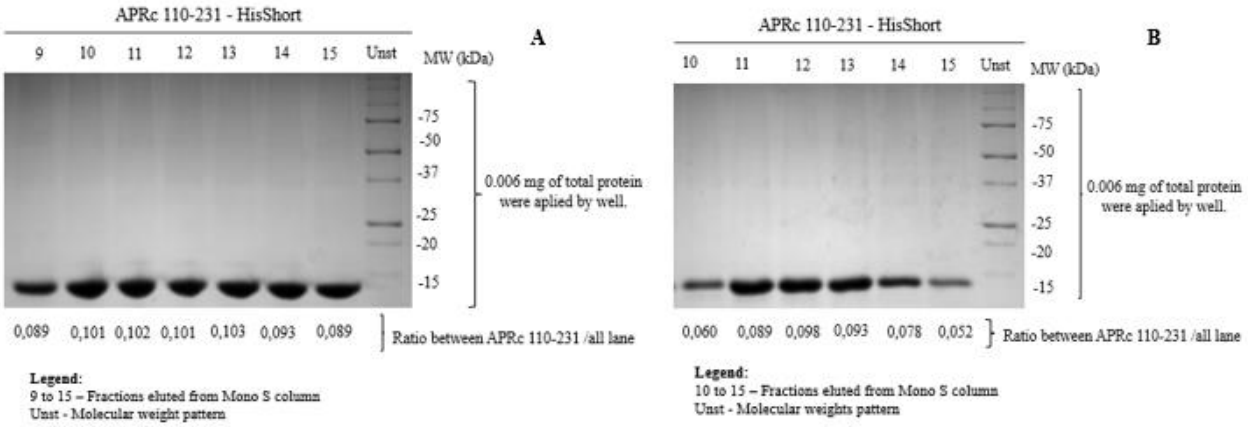


Figure 22 - SDS-PAGE from the trial done with mechanical lysis at HPH, with minimum pressure applied in the sample and 6 cycles of compression – Mono-S chromatography. The images **A** and **B** refer to the pools previously removed from 150 mM imidazole fraction and 200 mM imidazole fraction, respectively. The samples applied were first quantified by NanoDrop (ND-1000 Spectrophotometer), and then 6 μg of total protein were applied by well.

As shown in Figures 17 and 18, under modified conditions, there are several differences in the results obtained compared to the standard method, which was the starting point of this project. When no pressure was applied to the sample, the amount of APRc recovered from purification steps was lower, but when only the pressure necessary to cause intracellular release was applied, it was noted that the quantity of APRc recovered was higher comparing to standard assay (section 3.1.1.), indicating that the process can be optimized using these conditions. These assumptions were confirmed using SDS-PAGE, (*cf.* Figures 19 to 22).

3.1.6. *Evaluation of the oligomerization state of APRc*

Since APRc is described as an active enzyme with accumulation in the dimeric form that is essential for proteolytic activity (Cruz et al., 2014; M. Li et al., 2015), the oligomerization state and activity of these recombinant forms of APRc were also evaluated, in order to analyse whether the obtained APRc is a monomer or a dimer and if it is a dimer, if it is active, meaning that the APRc is in its functional form.

First, the oligomerization state of APRc was evaluated because it was confirmed that purified protein was in the correct oligomerization state, as we were interested in the dimeric/active form, as mentioned before.

The oligomerization state of the all-purified fractions from cation-exchange chromatography was evaluated, injecting the samples in a Superdex 200 column as previous explained in section 2.2.4. In retropepsins, dimerization of the enzyme is essential for proteolytic activity (M. Li et al., 2015), so it is essential to evaluate the oligomerization state. Size-exclusion chromatography was performed for this purpose. The macromolecules were separated by molecular size in a packed column with porous beads. The molecular weights of protein samples were estimated by calibrating the column with standard molecular weight proteins (ovalbumin and conalbumin) (Cytiva, 2021). A 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+}, for both methods of cell lysis) is shown in Figure 23. The chromatogram shows the monitoring of protein elution at an absorbance of 220 nm and the respective elution volumes. The elution volumes for the standard molecular weight proteins are shown as black points.

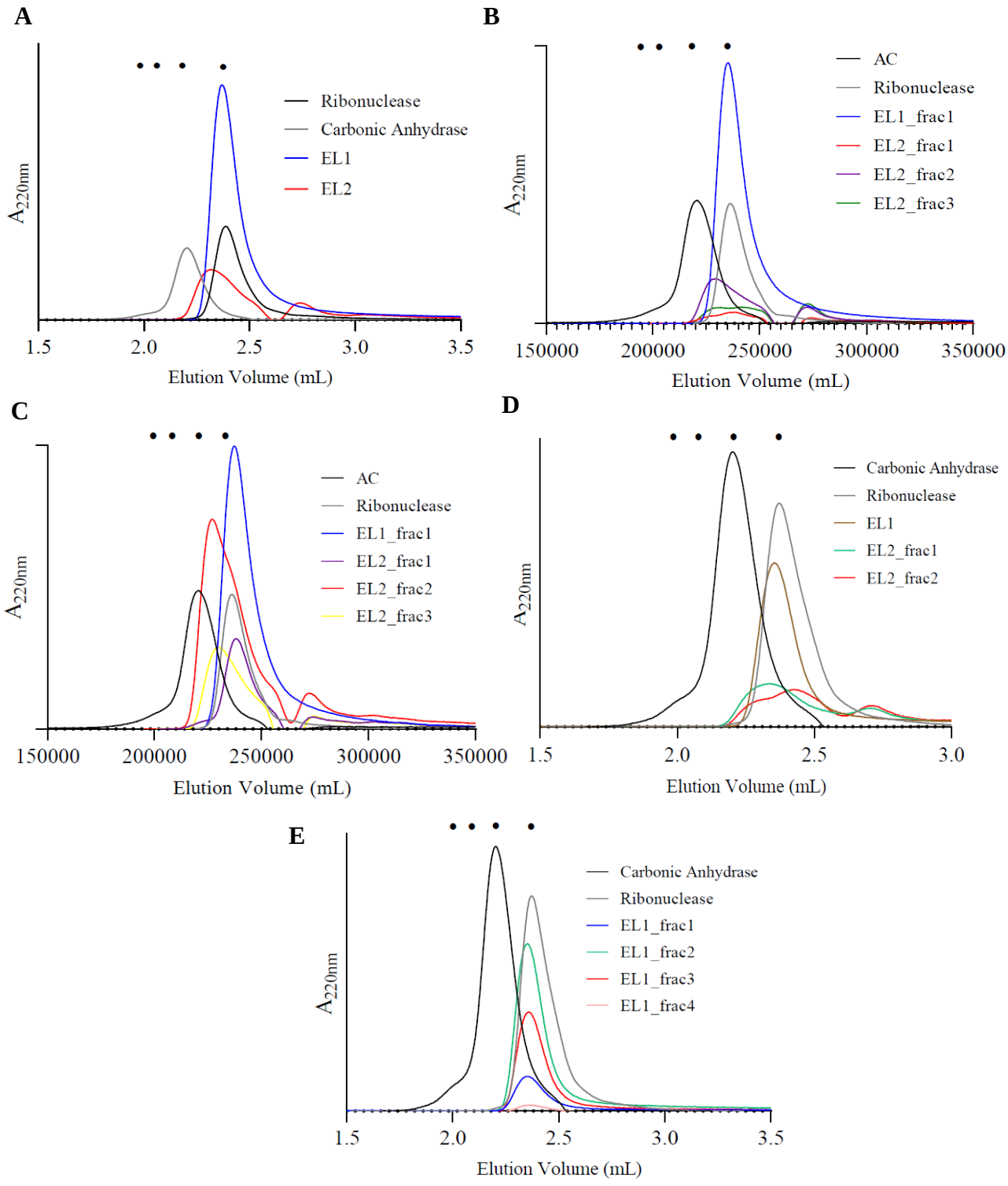

A
Ribonuclease
Carbonic Anhydrase
EL1
EL2
A220nm
1.5
2.0
2.5
3.0
3.5
Elution Volume (mL)
B
AC
Ribonuclease
EL1_frac1
EL2_frac1
EL2_frac2
EL2_frac3
A220nm
150000
200000
250000
300000
350000
Elution Volume (mL)
C
AC
Ribonuclease
EL1_frac1
EL2_frac1
EL2_frac2
EL2_frac3
A220nm
150000
200000
250000
300000
350000
Elution Volume (mL)
D
Carbonic Anhydrase
Ribonuclease
EL1
EL2_frac1
EL2_frac2
A220nm
1.5
2.0
2.5
3.0
Elution Volume (mL)
E
Carbonic Anhydrase
Ribonuclease
EL1_frac1
EL1_frac2
EL1_frac3
EL1_frac4
A220nm
1.5
2.0
2.5
3.0
3.5
Elution Volume (mL)

Figure 23 (previous page) - Evaluation of the oligomerization state of APRc. The purified samples of the recombinant protein $APRc_{110-231}$-HisShort (eluted fractions from 150 mM and 200 mM imidazole gradient steps of IMAC Ni2+), 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. **(A) (B)** and **(C)** are the SEC for mechanical lysis under the "standard" conditions, and modified conditions with no pressure applied on the sample and with lower pressure applied on the sample, respectively. **(D)** and **(E)** are the SEC for physicochemical lysis with 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl, with 4 cycles of freezing-thawing at 20ºC and 37ºC, respectively. 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) and ribonuclease A (13.7 kDa).

To calculate the molecular weight of each eluted fraction, a calibration curve was constructed using standard molecular weight proteins. The MW of each sample was calculated using the equation of the calibration line. That said, the MW for the monomeric forms of the protein is approximately 14 kDa, and the MW for the dimeric forms is approximately 28 kDa. The theoretical MW of APRc is 14 kDa.

From Figure 23, it can be observed that the fractions eluted at the same volume as ribonuclease are monomers, and the fractions eluted at the carbonic anhydrase volume are dimers because of their molecular weight. Therefore, it is possible to conclude that the assumption that the 150 mM imidazole pool was a monomer, and the 200 mM imidazole pool was a dimer was correct.

3.1.7. *Evaluation of the activity of APRc*

Since the oxidized insulin β-chain is usually a substrate of aspartic proteases, researchers assessed the activity of APRc towards this protein (Teixeira, B. D., 2020). Thus, the enzymatic activity of all the purified fractions of APRc was evaluated by Reverse-Phase chromatography. Therefore, APRc samples were incubated with the oxidized insulin β-chain at pH 6, since APRc displays optimal activity in vitro at pH 6, like previous described in section 2.2.4. (M. Li et al., 2015). The representative chromatograms are shown in Figure 24.

In Figure 24, the activities of the dimeric forms of APRc110-231-HisShort towards the substrate are compared to the activity of its monomeric forms and the elution profile of the oxidized insulin β-chain. Figures 24 (A), (B) and (C) show the activities of APRc obtained by the mechanical lysis methods. Figures 24 (D) and (E) show the APRc activities obtained using the physicochemical lysis methods. Note that insulin was not inserted in the first runs, only in the last run, hence the deviation in elution volume of the peaks in Figures 24 – (A), (D), and (E).

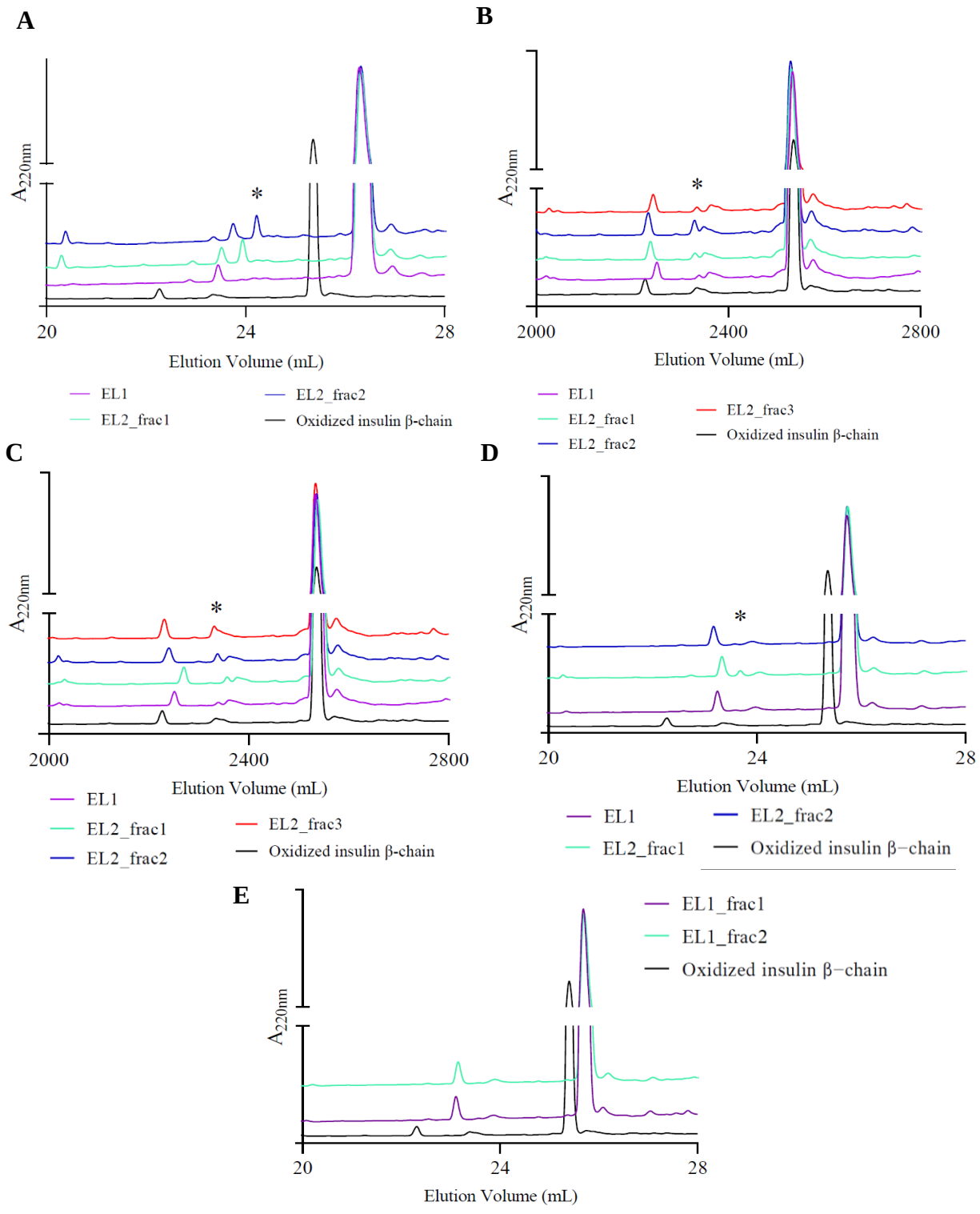

A
A220nm
*
20
24
28
Elution Volume (mL)
EL1
EL2_frac1
EL2_frac2
Oxidized insulin β-chain
B
A220nm
*
2000
2400
2800
Elution Volume (mL)
EL1
EL2_frac1
EL2_frac2
EL2_frac3
Oxidized insulin β-chain
C
A220nm
*
2000
2400
2800
Elution Volume (mL)
EL1
EL2_frac1
EL2_frac2
EL2_frac3
Oxidized insulin β-chain
D
A220nm
*
20
24
28
Elution Volume (mL)
EL1
EL2_frac1
EL2_frac2
Oxidized insulin β−chain
E
A220nm
20
24
28
Elution Volume (mL)
EL1_frac1
EL1_frac2
Oxidized insulin β−chain

Figure 24 (previous page) - Evaluation of the proteolytic activity of APRc. Eluted fractions from 150 mM and 200 mM imidazole gradient step of recombinant $APRc_{110-231}$-HisShort (monomer and dimer, respectively) 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 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. **(A)** Eluted fractions obtained from mechanical lysis, under "standard" conditions. **(B)** Eluted fractions obtained from mechanical lysis, under modified conditions: no pressure applied on the sample. **(C)** Eluted fractions obtained from mechanical lysis, under modified conditions: with lower pressure applied on the sample. **(D)** Eluted fractions obtained from physicochemical lysis with 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl, with 4 cycles of freezing-thawing at 20ºC. **(E)** Eluted fractions obtained from physicochemical lysis with 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl, with 4 cycles of freezing-thawing at 37ºC.

The reverse phase chromatograms (Figure 24) show that $APRc_{110-231}$-HisShort displayed activity in relation to the oxidized insulin β-chain substrate, as new peaks corresponding to peptides originating from the digestion of the substrate were detected. For the monomeric form of $APRc_{110-231}$-HisShort, no activity towards this substrate was observed, as expected. These results confirm that the dimeric forms of APRc are active, whereas the monomeric form is not. In addition, APRc displayed no functional form after the cell lysis assay performed at 37 ºC.

3.2. Solid/Solid Ratio and Yield – a comparison between all the cell lysis methods

From the previous analysis, it was shown that some of the methods did not accumulate protein in the dimeric form, mainly from physicochemical lysis. For those that allowed the recovery of protein in the active form, the yield was calculated as described below.

To evaluate the efficiency of the implemented cell lysis method, the yield was calculated to determine the amount of monomer and dimer. The solid/solid ratio was obtained as the ratio of wet cells to solvent mass, and the monomer *vs* dimer yield was calculated by quantifying the protein (described in Section 2.2.5) and the relation OD_{280nm} = 2.85 mg/mL. With this relationship between the OD measured at 280 nm (by NanoDrop®) and the concentration, the concentration of each of the eluted protein fractions can be determined, and consequently the yield. The results are presented in Table 7.

Table 7 – Evaluation of the efficiency of the implemented cell lysis method.

Cell lysis method	**Solid/Solid ratio (%)**	**Yield (monomer *vs* dimer) (mg)**
"Standard" mechanical lysis	29.4	61.6 *vs* 0.45
Mechanical lysis with no pressure applied in the sample	26.8	55.85 *vs* 0.36
Mechanical lysis with minimum pressure applied in the sample		43.11 *vs* 1.74
Physicochemical lysis	31.3	44.52 *vs* 0.54

As can be seen in Table 7, the amount of dimer obtained was far less than that of the monomer, as predicted and established in the laboratory. The best yield of the dimeric fraction was obtained for the assay in which a mechanical lysis was applied, with minimum pressure applied to the sample. These assays showed promising results, with the increasing from 0.45 to 1.74 mg (74%) in the yield of the dimer, and optimization was achieved with success.

4. Discussion

4.1. Rickettsial Infections and APRc

In recent years, the identification of new rickettsioses has continued to rise, underscoring the global emergence of rickettsial infections. The understanding of *Rickettsia*'s pathogenicity has significantly expanded, with an increasing number of studies in this area. Despite this progress, the development of a reliable protective vaccine or specific treatment for rickettsioses remains elusive. This reinforces the critical need to identify new protein factors as potential targets for the development of novel therapeutics.

One such protein factor is APRc, a novel retropepsin-like enzyme, discovered and characterized by Cruz *et al.* (Cruz et al., 2014). Retropepsins, including APRc, require homodimerization of two monomeric units in order to form a functional enzyme. Retroviral proteases, for example, are active as symmetric dimers, with two identical monomers and a single active site formed by residues one of the monomers (M. Li et al., 2015). Recent studies have revealed that hat APRc serves as a novel moonlighting protein in *Rickettsia*, acting as an evasin and contributing to *Rickettsia*'s immune evasion mechanisms (Curto et al., 2021). Despite its structural similarity to other retropepsins at the monomeric level, the dimer observed in APRc crystals differs entirely from the canonical dimer seen in retroviral enzymes. This APRc dimer cannot form a functional active site, and it is believed to be an artifact of the expression and crystallization conditions. In other words, Even though the topology of the dimer is the same in all three crystal forms, the angle between the two molecules is considerably different, suggesting that the observed dimer must be an artefact of the expression and crystallization conditions (M. Li et al., 2015).

4.2. Challenges in the production and recovery of active APRc

APRc was previously produced and purified from *E. coli* in the laboratory of Dr. Isaura Simões. However, they encountered challenges related to obtaining high yields of active dimer and lower yields of monomer, which limited their ability to achieve the canonical dimer of APRc and conduct studies on structure-function relationship of APRc for the design of specific inhibitors. These limitations were primarily associated with the established method for APRc production, specifically in the release and extraction of the intracellular protein from *E. coli* cells, widely known as cell lysis.

Mechanical lysis, particularly using the HPH method, is commonly employed for disrupting the cellular wall of *E. coli*. However, this method has its disadvantages, including the generation of heat during disruption, which can denature the protein of interest. As a result, it became imperative to explore new, cost-effective alternatives that would yield higher amounts of the desired bioproduct.

The most significant challenge in this process is extracting functional and active proteins from heterologous expression systems like *E. coli*. Recent studies in this field have suggested that physicochemical lysis methods may offer better chances of recovering the target protein compared to mechanical lysis (Dos Santos et al., 2018; Lo et al., 2016; Martins et al., 2018). For instance, Martins et al. developed a novel method for extracting GFP from *E. coli* cells using tensioactive compounds, such as ionic liquids. Their results showed that certain ionic liquids were more effective than ultrasonic-assisted extraction, capable of extracting the entire GFP content from the cells (Martins et al., 2018).

In this project, initially, the goal was to reproduce the standard method and use it for comparison with the modifications introduced in subsequent assays to optimize the cell lysis method. However, after evaluating the standard cell lysis method, it appeared that was no significant difference between the total protein extracted in the *E. coli* lysate and the total protein present in the soluble fraction. As a result, it was decided to conduct a comprehensive sequence of new trials testing physicochemical lysis.

During these physicochemical method trials, it became evident that optimizing the process in the presence of SDS was not feasible, as APRc either was not being extracted or precipitated. Subsequently, we conducted experiments to study the influence of thawing temperature on APRc recovery. This included thawing at 20 ºC for 44 min and at 37 ºC, with a step initially at 20 ºC followed by another at 37 ºC for 29 min. It was observed that the 20 ºC test resulted in higher release of APRc dimer compared to the standard test. However, heating up to 37 ºC led to the loss of all the protein in its dimeric form, making it impractical method for APRc production. Even in assays applying only physicochemical lysis, *i.e.*, with a 20 mM sodium phosphate pH 7.4 buffer, containing 10 mM imidazole and 500 mM NaCl and no addition of surfactants, no improvements were observed in APRc extraction.

After evaluating all physicochemical methods and realizing that optimization was not feasible for obtaining the protein through these methods, it was decided to turn the focus back to the previously established laboratory HPH process. It was decided to modify the operating conditions, resulting in several differences in the results compared to the standard method,

which served as the starting point of this project. When no pressure was applied to the sample, the amount of APRc recovered from purification steps was lower. However, when only the necessary pressure to cause intracellular release was applied, it was noted that a higher quantity of APRc recovered.

With this work, it was demonstrated that cell lysis through physicochemical methods does not appear to be a suitable option for the improving the recovery of functional APRc from *E. coli* cells, despite previous studies suggesting their potential, at least for permeabilization of these bacterial cells and release of intracellular proteins and other biomolecules. Instead, modifying the established mechanical lysis method with HPH proved to be an effective strategy for optimizing the process. Specifically, applying a minimum pressure in the sample, slightly lower than that used in the standard method, resulted in the extraction of APRc dimer from 0.45 to 1.74. Although this optimization is promising, it is important to explore potential synergies with other techniques, such as aqueous biphasic systems, to further enhance the process.

5. Conclusions and Future Perspectives

This project aimed to address the challenges associated with the recovery of recombinant APRc from *E. coli* cells, which included low functional protease concentration, the need for higher extraction and purification yields of functional protein for making possible to identifying specific APRc inhibitors. Since APRc is expressed in *E. coli*, the primary and main goal of this work was to select the most effective cell disruption method for obtaining an APRc-rich lysate supernatant.

Therefore, this project was started by replicating the cell lysis method currently used for obtaining APRc (entitled as standard method), to check the limitations and what could be improved or tested. The challenges described above were verified. So, several mechanical (using an HPH) and physicochemical (surfactant application, temperature, freezing-thawing cycles) lysis methods were tested during this project.

Modifying the standard procedure allowed for process optimization, resulting in an increased yield of the dimeric form (functional structure) of APRc from 0.45 to 1.74 (an improvement of 74%). However, the amount of the monomeric form decreased from 61.6 to 43.11. The desired goal of achieving a higher dimer-to-monomer ratio than the "standard" lysis method was thus achieved, as the dimeric form was proven to be functional in high-performance liquid chromatography.

It worth noting that in all the physicochemical lysis assays, handling the cellular pellet was challenging due to the presence of surfactant (i.e., SDS), which hindered proper resuspension and subsequent thawing. It was concluded that the use of detergents/surfactants, such as SDS, did not lead to process optimization for APRc production. In many tests, the protein become denatured or precipitated, and when obtained, the protein content did not exceed the yield of the "standard" method and lacked the same level of activity as the protein obtained through mechanical lysis. Additionally, it was observed that performing cell lysis at temperatures above room temperature resulted in the release of non-functional protein.

For future perspectives, it is suggested to explore the potential of using aqueous biphasic systems, incorporating ionic liquids, polymers, salts and/or other surfactants, for the recovery, concentration and purification APRc. Furthermore, developing an integrated platform for APRc purification by sequentially utilizing compounds that allow the cell permeabilization and APRc solubilization, and which can also be used as phase forming compounds of the subsequent aqueous biphasic systems. This could lead to the overall integration of APRc production and

purification process. Additionally, in terms of recombinant production, experiments can be conducted to optimize several factors affecting expression of recombinant protein in the soluble fraction, such as the inducer concentration, induction temperature, expression time, and optical density.

References

Abdad, M. Y., Abdallah, R. A., Fournier, P. E., Stenos, J., Vasoo, S. (2018). A concise review of the epidemiology and diagnostics of rickettsioses: *Rickettsia* and Orientia spp. Journal of Clinical Microbiology, 56(8).

Angelakis, E., Raoult, D. (2017). *Rickettsia* and *Rickettsia*-Like Organisms. In Infectious Diseases (Fourth Edi). Elsevier Ltd.

Azad, A. F., Beard, C. B. (1998). Rickettsial pathogens and their arthropod vectors. Emerging Infectious Diseases, 4(2), 179–186.

Balasundaram, B., Harrison, S., Bracewell, D. G. (2009). Advances in product release strategies and impact on bioprocess design. Trends in Biotechnology, 27(8), 477–485.

Barro, A. M. (2020). Characterization of APRc as an immune evasion factor of *Rickettsia*. FCT/UNL, UNL.

Bernaudat, F., Frelet-Barrand, A., Pochon, N., Dementin, S., Hivin, P., Boutigny, S., Rioux, J. B., Salvi, D., Seigneurin-Berny, D., Richaud, P., Joyard, J., Pignol, D., Sabaty, M., Desnos, T., Pebay-Peyroula, E., Darrouzet, E., Vernet, T., Rolland, N. (2011). Heterologous expression of membrane proteins: Choosing the appropriate host. PLoS ONE, 6(12).

Bio-Rad. Ion Exchange Chromatography. Available in: https://www.bio-rad.com/en-pt/applications-technologies/ion-exchange-chromatography?ID=MWHAY9ESH (Accessed September 1, 2021)

Blanton, L. S. (2019). The Rickettsioses: A practical Update. 33, 213–229.

Botelho-Nevers, E., Socolovschi, C., Raoult, D., Parola, P. (2012). Treatment of *Rickettsia* spp. infections: A review. Expert Review of Anti-Infective Therapy, 10(12), 1425–1437.

Chan, Y. G. Y., Riley, S. P., Chen, E., Martinez, J. J. (2011). Molecular basis of immunity to rickettsial infection conferred through outer membrane protein B. Infection and Immunity, 79(6), 2303–2313.

Cruz, R., Huesgen, P., Riley, S. P., Wlodawer, A., Faro, C., Overall, C. M., Martinez, J. J., Simões, I. (2014). RC1339/APRc from *Rickettsia conorii* Is a Novel Aspartic Protease with Properties of Retropepsin-Like Enzymes. PLoS Pathogens, 10(8).

Curto, P., Barro, A., Almeida, C., Vieira-Pires, R. S., Simões, I. (2021). The Retropepsin-Type Protease APRc as a Novel Ig-Binding Protein and Moonlighting Immune Evasion Factor of *Rickettsia*. MBio, 12(6).

Cytiva. Fundamentals of size exclusion chromatography. Available in: https://www.cytivalifesciences.com/en/us/solutions/protein-research/knowledge-center/protein-purification-methods/size-exclusion-chromatography (Accessed September 1, 2021)

De Sousa, R., Dória Nóbrega, S., Bacellar, F., Torgal, J. (2003). Mediterranean spotted fever in Portugal: Risk factors for fatal outcome in 105 hospitalized patients. Annals of the New York Academy of Sciences, 990, 285–294.

Dos Santos, N. V., Martins, M., Santos-Ebinuma, V. C., Ventura, S. P. M., Coutinho, J. A. P., Valentini, S. R., Pereira, J. F. B. (2018). Aqueous Biphasic Systems Composed of Cholinium Chloride and Polymers as Effective Platforms for the Purification of Recombinant Green Fluorescent Protein. ACS Sustainable Chemistry and Engineering, 6(7), 9383–9393.

Eggenreich, B., Rajamanickam, V., Wurm, D. J., Fricke, J., Herwig, C., Spadiut, O. (2017). A combination of HPLC and automated data analysis for monitoring the efficiency of high-pressure homogenization. Microbial Cell Factories, 16(1), 1–10.

Evans, T. C., Xu, M.-Q. (2011). Heterologous gene expression in E. coli: methods and protocols.

Harrison, S. T. L. (2011). Cell Disruption. In Comprehensive Biotechnology, Second Edition (Second Edi, Vol. 2). Elsevier B.V.

Islam, M. S., Aryasomayajula, A., Selvaganapathy, P. R. (2017). A review on macroscale and microscale cell lysis methods. Micromachines, 8(3).

Jeffery, C. J. (1999). Moonlighting proteins. Trends Biochem Sci, 8–11.

Kaur, J., Kumar, A., Kaur, J. (2018). Strategies for optimization of heterologous protein expression in E. coli: Roadblocks and reinforcements. International Journal of Biological Macromolecules, 106, 803–822.

Kazimírová, M., Štibrániová, I. (2013). Tick salivary compounds: Their role in modulation of host defences and pathogen transmission. Frontiers in Cellular and Infection Microbiology, 4(AUG), 1–19.

Kim, J., Kim, S. I., Hong, E., Ryu, Y. (2016). Strategies for increasing heterologous expression of a thermostable esterase from Archaeoglobus fulgidus in Escherichia coli. Protein Expression and Purification, 127, 98–104.

Konvalinka, J., Kräusslich, H. G., Müller, B. (2015). Retroviral proteases and their roles in virion maturation. Virology, 479–480, 403–417.

Li, H., Walker, D. H. (1992). Characterization of rickettsial attachment to host cells by flow cytometry. Infection and Immunity, 60(5), 2030–2035.

Li, M., Gustchina, A., Cruz, R., Simões, M., Curto, P., Martinez, J., Faro, C., Simões, I., Wlodawer, A. (2015). Structure of RC1339/APRc from *Rickettsia conorii*, a retropepsin-like aspartic protease. Acta Crystallographica Section D: Biological Crystallography, 71, 2109–2118.

Lo, S. C., Ramanan, R. N., Tey, B. T., Tan, W. S., Show, P. L., Ling, T. C., Ooi, C. W. (2016). A versatile and economical method for the release of recombinant proteins from Escherichia coli by 1-propanol cell disruption. RSC Advances, 6(67), 62291–62297.

Martins, M., Ooi, C. W., Neves, M. C., Pereira, J. F. B., Coutinho, J. A. P., Ventura, S. P. M. (2018). Extraction of recombinant proteins from Escherichia coli by cell disruption with aqueous solutions of surface-active compounds. Journal of Chemical Technology and Biotechnology, 93(7), 1864–1870.

Mcclure, E. E., Chávez, A. S. O., Shaw, D. K., Carlyon, J. A., Roman, R., Noh, S. M., Wood, D. O., Bavoil, P. M., Brayton, K. A., Juan, J., Mcbride, J. W., Valdivia, R. H., Munderloh, U. G., Joao, H. F. (2017). Challenges and Paradigms. 15(9), 544–558.

Osterloh, A. (2020). The neglected challenge: Vaccination against rickettsiae. In PLoS Neglected Tropical Diseases (Vol. 14, Issue 10).

Parola, P., Paddock, C. D., Raoult, D. (2005). Tick-borne rickettsioses around the world: Emerging diseases challenging old concepts. Clinical Microbiology Reviews, 18(4), 719–756.

Parola, P., Paddock, C. D., Socolovschi, C., Labruna, M. B., Mediannikov, O., Kernif, T., Abdad, M. Y., Stenos, J., Bitam, I., Fournier, P.-E., Raoult, D. (2013). Update on tick-borne rickettsioses around the world: a geographic approach. Clinical Microbiology Reviews, 26(4), 657–702.

Pereira, J. F. B., Hasmann, F. A. (2020). CAPÍTULO 3 Rompimento celular. In Purificação de produtos biotecnológicos (pp. 67–104).

Renesto, P., Ogata, H., Audic, S., Claverie, J. M., Raoult, D. (2005). Some lessons from *Rickettsia* genomics. FEMS Microbiology Reviews, 29(1), 99–117.

Rosano, G. L., Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: Advances and challenges. Frontiers in Microbiology, 5(APR), 1–17.

Sahni, A., Fang, R., Sahni, S. K., Walker, D. H. (2019). Pathogenesis of Rickettsial Diseases: Pathogenic and Immune Mechanisms of an Endotheliotropic Infection. Annual Review of Pathology: Mechanisms of Disease, 14(August 2018), 127–152.

Sahni, S. K., Narra, H. P., Sahni, A., Walker, D. H. (2013). Recent molecular insights into rickettsial pathogenesis and immunity. 1265–1288.

Sharma, A. K., Dhasmana, N., Dubey, N., Kumar, N., Gangwal, A., Gupta, M., Singh, Y. (2017). Bacterial Virulence Factors: Secreted for Survival. Indian Journal of Microbiology, 57(1).

Sigma-Aldrich. HisTrap™ High Performance. Available in: https://www.sigmaaldrich.com/PT/en/product/sigma/ge29051021 (Accessed September 1, 2021)

Smolskaya S., Logashina YA., Andreev YA. (2020). Escherichia coli Extract-Based Cell-Free Expression System as an Alternative for Difficult-to-Obtain Protein Biosynthesis. Int J Mol Sci, 21(3):928.

Stewart, A. G., Stewart, A. G. A. (2021). An Update on the Laboratory Diagnosis of *Rickettsia* spp. Infection. Pathogens, 10(10), 1319.

Teixeira, B. D. (2020). Elucidating autoprocessing activity and function of the rickettsial retropepsin APRc in mammalian cells. Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Walker, D. H., Valbuena, G. A., Olano, J. P. (2003). Pathogenic mechanisms of diseases caused by *Rickettsia*. Annals of the New York Academy of Sciences, 990, 1–11.

Wang, M., Chen, S., Zhou, W., Yuan, W., Wang, D. (2020). Algal cell lysis by bacteria: A review and comparison to conventional methods. Algal Research, 46(January), 101794.

Wlodawer, A., Gustchina, A. (2000). Structural and biochemical studies of retroviral proteases. Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology, 1477(1–2), 16–34.

Wood, D. O., Azad, A. F. (2000). Genetic manipulation of Rickettsiae: A preview. Infection and Immunity, 68(11), 6091–6093.

Appendices

Appendix A. Representation of the amino acid sequence of the construct of APRc used throughout the project

pET_APRc$_{110-231}$-HisShort

110-EVGEIIIARNRDGHFYINAFVNNVKIKFMVDTGASDIALTKEDAQKLGFDLTKLKYTRTYLTANGENKAAPITLNSVVIGKEFKNIKGHVGLGDLDISLLGMSLLERFKGFRIDKDLLILNY HHHHHH

Figure 25 - Representation of the amino acid sequence of the construct pET_APRc110-231-HisShort. The construct pET_APRc110-231-HisShort includes the coding sequence of APRc amino acids 110-231 fused with the C-terminal tag sequence HHHHHH. Adapted from Barro, A (Barro, A. M., 2020).