UNIVERSITÀ DEGLI STUDI DEL PIEMONTE ORIENTALE

“AMEDEO AVOGADRO”

DIPARTIMENTO DI SCIENZE DEL FARMACO

Corso di Laurea Magistrale in Biotecnologie Farmaceutiche

TESI DI LAUREA

Optimization of expression and purification conditions for recombinant NAD-binding protein PARP15, a potential pharmacological target in oncology

Relatore

Prof. Riccardo Miggiano

Candidato

Daniele Rappo

Correlatore

Dott. Matteo Leoncini

Anno Accademico 2023/2024

Sessione autunnale

Abbreviations

AI: artificial intelligence

ART: ADP-ribosyl transferase

CV: column volumes

FT: flow-through

IEX: ion-exchange chromatography

IPTG: isopropyl β-D-1-thiogalactopyranoside

MART: mono-(ADP-ribosyl) transferase

MW: molecular weight

MWCO: molecular weight cutoff

NAD: nicotinamide adenine dinucleotide

NBP: NAD-binding protein

Ni-NTA: nickel-nitrilotriacetic acid

OD: optical density

o.n. overnight

pI: isoelectric point

PTM: post-translational modification

R -: resin before elution

R +: resin after elution

SDS-PAGE: sodium dodecyl sulfate-polyacrylamide gel electrophoresis

SEC: size exclusion chromatography

TCEP: tris-(2-carboxyethyl)-phosphine

W: wash-unbound

Table of contents

1. Introduction 8
1.1 Cancer and immunotherapy 9
1.2 Cancer and targeted therapy 11
1.3 NAD and NAD-binding proteins (NBPs) 12
1.4 The PARP superfamily 15
1.5 PARP15 21
1.6 The importance of structural biology in drug development 24
2. Aim of the work 26
3. Results and discussion 29
3.1 Isolation of high-purity pET-28a (+) plasmids from *E. coli* DH5α cells for subsequent transformations 30
3.2 Identification of expression conditions for PARP15_78-678 and PARP15_FL 30
3.2.1 Test expression in 2X YT medium 30
3.2.2 Test expression in autoinduction medium 33
3.3 Purification trials of PARP15_78-678 36
3.3.1 First purification 36
3.3.2 Second purification 40
3.3.3 Third and fourth purification 45
3.3.4 Fifth purification 50
3.4 Crystallization of PARP15 Catalytic Domain 52
4. Conclusions and future perspectives 53
5. Materials and methods 55
5.1 Recombinant protein production in *E. coli* 56
5.2 Chromatographic techniques for the purification of recombinant proteins 57
5.2.1 Ni-NTA affinity chromatography 58
5.2.2 Size exclusion chromatography (SEC) 58
5.2.3 Ion-exchange chromatography (IEX) 60
5.2.4 Heparin chromatography 61
5.3 Amplification and isolation of high-purity vectors encoding PARP15_78-678 and PARP15_FL proteins 62
5.3.1 Expression clones 62

5.3.2 Isolation of high-purity pET-28a (+) plasmids from *E. coli* DH5α cells for subsequent transformations 63
5.4 Transformation of *E. coli* cells 64
5.4.1 *E. coli* DH5α competent cells 64
5.4.2 *E. coli* BL21 (DE3) competent cells 64
5.4.3 Transformation of *E. coli* cells via heat-shock protocol 65
5.5 Growth of transformed *E. coli* BL21 (DE3) bacterial cells 65
5.5.1 2X YT medium and induction of recombinant proteins expression 65
5.5.2 Autoinduction medium for recombinant proteins expression 66
5.6 Screening of PARP15_78-678 and PARP15_FL expression conditions 67
5.7 Test expression procedure 69
5.8 Purification trials of PARP15_78-678 71
5.8.1 Large-scale cultures for bacterial pellets production 71
5.8.2 Cell lysis and clarification 71
5.8.3 First purification trial 72
5.8.4 Second purification trial 74
5.8.5 Third purification trial 76
5.8.6 Fourth purification trial 77
5.8.7 Fifth purification trial 77
5.9 Western blot for protein detection 78
5.10 SDS-PAGE analysis 79
5.11 Bradford's method for protein quantification 80
5.12 Crystallization procedure for PARP15 Catalytic Domain 81
Bibliography 84

1. Introduction

1.1 Cancer and immunotherapy

In recent decades, cancer has progressively become the leading cause of death worldwide. According to a 2019 report from the World Health Organization (WHO), cancer is one of the primary causes of death before the age of 70 in as many as 100 countries around the globe (fig. 1). Furthermore, the social impact of this disease continues to grow relentlessly (Sung *et al.,* 2021).

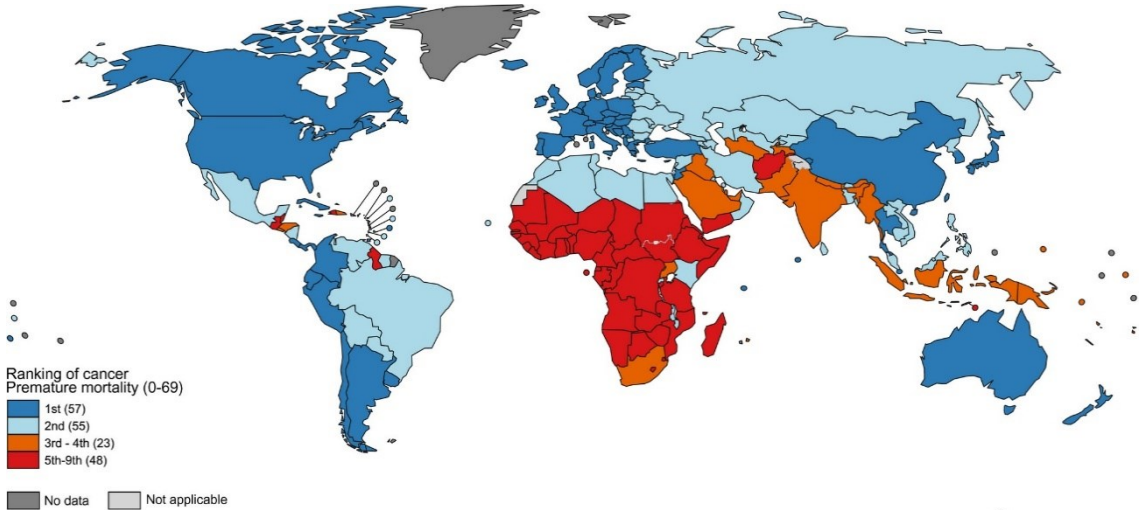

Fig. 1: national ranking of cancer as a cause of death at ages < 70 years in 2019. (Sung *et al.,* 2021)

Over the years, several therapeutic strategies have been developed against cancer, such as chemotherapy and radiotherapy. The culmination of these efforts is cancer immunotherapy, an innovative strategy that aims to enhance the antitumor immune response while reducing off-target effects compared to previous approaches (Rosenberg, 2014). Cancer immunotherapy uses agents to activate or boost the host's immune system, promoting the attack on cancer cells through physiological mechanisms often evaded during disease progression. As a result, it is recognized as a particularly promising therapeutic strategy for treating, and in some cases curing, various types of cancer (Riley *et al.,* 2019).

According to the Cancer Research UK organization and by several studies, such as the one by Jiang *et al.,* 2024, there are different types of cancer immunotherapy approaches:

- Monoclonal antibodies (mAbs) stimulate the patient's immune system to recognize and destroy cancer cells as foreign. They may also inhibit the division of cancer cells.
- Checkpoint inhibitors block "checkpoint proteins" (such as CTLA-4, PD-1 and PD-L1), which signal the immune system to shut down to prevent autoimmune responses or damage to healthy tissues. In cancer, these checkpoints can prevent the immune system from attacking tumors.
- Cytokines, such as interferons and interleukins, naturally boost the immune response. Modified versions of these proteins have been developed specifically for cancer treatment.
- Cancer vaccines are designed for individuals who already have cancer. They identify specific proteins, known as "tumor antigens," unique to cancer cells. Once these proteins are recognized, the immune system is activated.
- CAR-T cell therapy involves *in vitro* genetically engineered T lymphocytes that recognize a tumor antigen as foreign and initiate an immune response against it. They express a chimeric T-cell receptor (TCR) and do not require a co-stimulatory signal to activate the cell-mediated response.
- Tumor-infiltrating lymphocytes (TILs) are immune cells capable of migrating from the bloodstream to the tumor, where they recognize and kill cancer cells.

Moreover, the recent study by Jiang *et al.,* 2024 emphasizes an intriguing finding: the effectiveness of cancer immunotherapy can be improved by modulating the gut microbiome with specific probiotic strains.

1.2 Cancer and targeted therapy

Targeted therapy works by focusing on specific proteins that regulate the growth, division, and spread of cancer cells (National Cancer Institute: targeted therapy to treat cancer). Fig. 2 highlights the concept of targeted therapy, using rare lung cancer as an example.

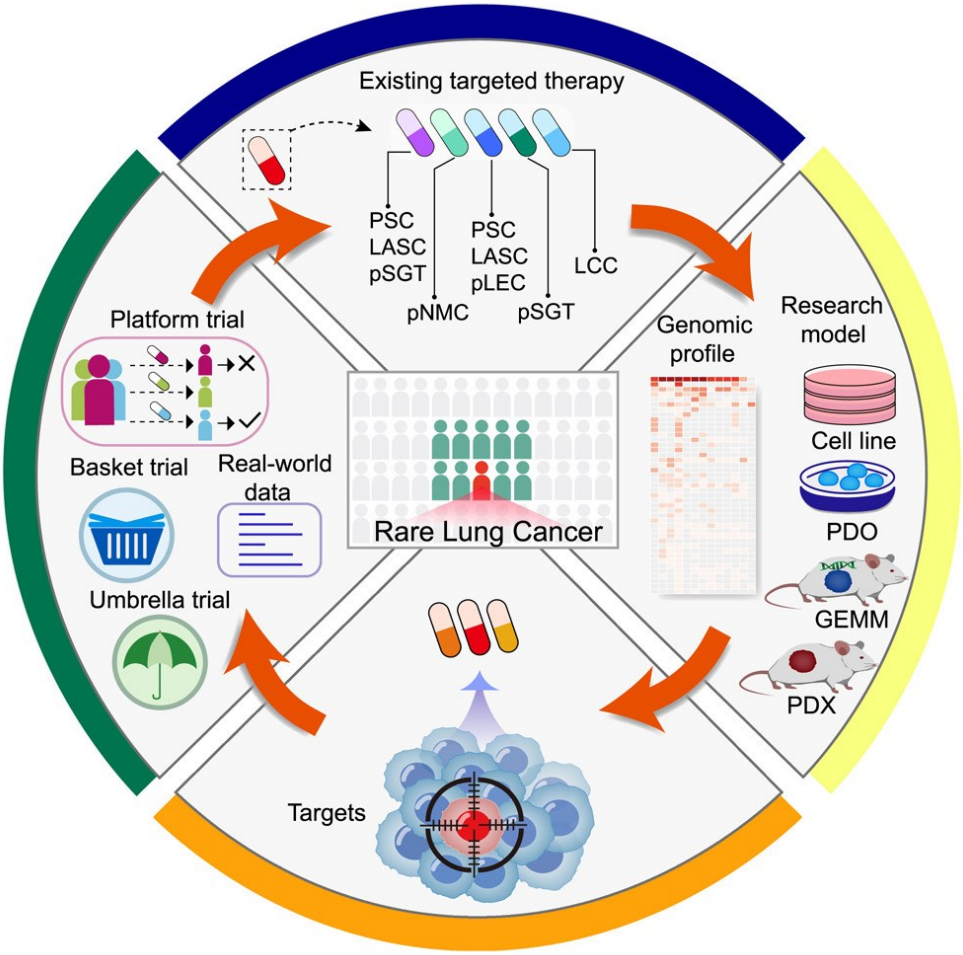

Fig. 2: targeted therapy. Patients can be grouped based on tumor specific characteristics and patients from different groups are treated with the most appropriate drug. (Wang *et al.,* 2023)

Targeted therapy may involve the use of the following agents (Min *et al.,* 2022):

- Small molecule inhibitors or antibodies that block signaling pathways responsible for cell growth, proliferation, and survival
- Hormonal therapies

- Immune checkpoint inhibitors

These drugs can be given to patients with not only altered targets, but also targets critical to specific pathways that drive tumor cell growth.

While this type of therapy has faced challenges, such as toxicity from unintended interactions with normal cells or drug resistance, it has led to significant survival improvements in certain cancers and is considered a step forward toward precision or personalized medicine (Min *et al.,* 2022).

In this context, it is crucial to identify which potential pharmacological targets can be addressed. Therefore, understanding the molecular mechanisms underlying the tumor environment is essential for developing targeted therapies. Over the years, the scientific community has identified numerous potential pharmacological targets, leading to the subsequent production of drugs against them.

1.3 NAD and NAD-binding proteins (NBPs)

NAD (nicotinamide adenine dinucleotide) is a coenzyme found in all living cells. It plays a crucial role in cellular metabolism and is involved in several biological processes, including energy production, ageing regulation, and DNA repair (Cantò *et al.,* 2012; Bogan and Brenner, 2008; Imai and Guarente, 2014):

- Energy metabolism: NAD is essential for cellular energy production. During the metabolism of glucose and fatty acids, NAD is involved in electron transport in redox reactions. It exists in two main forms (fig. 3): NAD^+ (oxidized) accepts electrons from other molecules, converting into NADH; NADH (reduced) donates electrons to the mitochondrial electron transport chain, helping in the synthesis of ATP.
- DNA repair: NAD is vital for the functioning of PARP proteins (poly-(ADP-ribose) polymerase), which are involved in repairing damaged DNA. Efficient DNA repair is essential to maintain genomic integrity and prevent mutations that could lead to diseases like cancer.

- Ageing regulation: NAD is involved in controlling the ageing process through the regulation of sirtuins, a group of enzymes that influence longevity and cellular stress resistance. Higher levels of NAD^+ are linked to better mitochondrial function and greater resistance to age-related diseases.
- Cellular communication: NAD acts as a signaling molecule, influencing both intracellular and intercellular communication. This role is particularly important in immune and inflammatory responses.

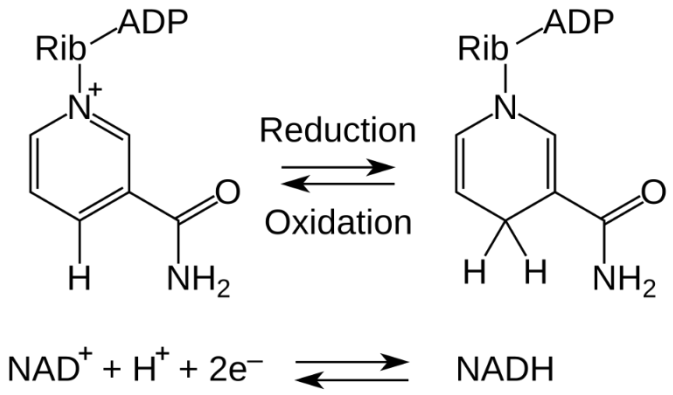

Fig. 3: oxidized and reduced forms structure formulas and redox reactions of NAD.

Physiologically, there are proteins that use NAD as a substrate; they are primarily involved in cellular survival mechanisms, DNA damage repair, calcium signaling, and transcriptional regulation. These proteins are known as "NAD-dependent enzymes", a subgroup of the larger class of NAD-binding proteins (NBPs), which includes all proteins that interact with NAD, whether it is used as a coenzyme or as a substrate (Duarte-Pereira *et al.,* 2023).

In 1996, the work of Bellamacina described the NAD-binding domain of all NBPs; it contains a βαβαβ unit and an additional β-strand, forming the minimal structure necessary for binding NAD. The first 30-35 amino acids of this motif form the so-called fingerprint region, which has a fold suitable for binding the dinucleotide. The fingerprint region exhibits several distinctive features: a phosphate-binding consensus sequence, GXGXXG; six positions typically occupied by small hydrophobic amino acids; a conserved negatively charged amino

acid residue (Glu or Asp) at the end of the second β-strand; a conserved positively charged amino acid residue (Arg or Lys) at the beginning of the first β-strand.

In the same year, Domenighini and Rappuoli published a study in which they specifically described three conserved consensus sequences that identify the NAD-binding site of certain NAD-dependent enzymes, particularly ADP-ribosylating enzymes. According to these researchers, in addition to the conserved negatively charged (Glu or Asp) and positively charged (Arg or Lys) amino acid residues present in the NAD-binding domain of all NBPs, for ADP-ribosylating enzymes: the Arg or Lys residue is preceded by a hydrophobic amino acid; within the β-strands, there is a consensus sequence arom-ph-Ser-Thr-Ser-ph (where "arom" refers to an aromatic amino acid and "ph" to a hydrophobic amino acid), while the Glu (or Asp) residue is part of the Glu/Gln-X-Glu motif; in some cases, the NAD-binding motif is Tyr-X_{10}-Tyr.

In more recent times, the characteristics of the NAD-binding site in other NBPs and NAD-dependent enzymes, including certain dehydrogenases, have also been elucidated (Metherell *et al.,* 2016). Fig. 4 shows the structure of a reference NAD-binding site.

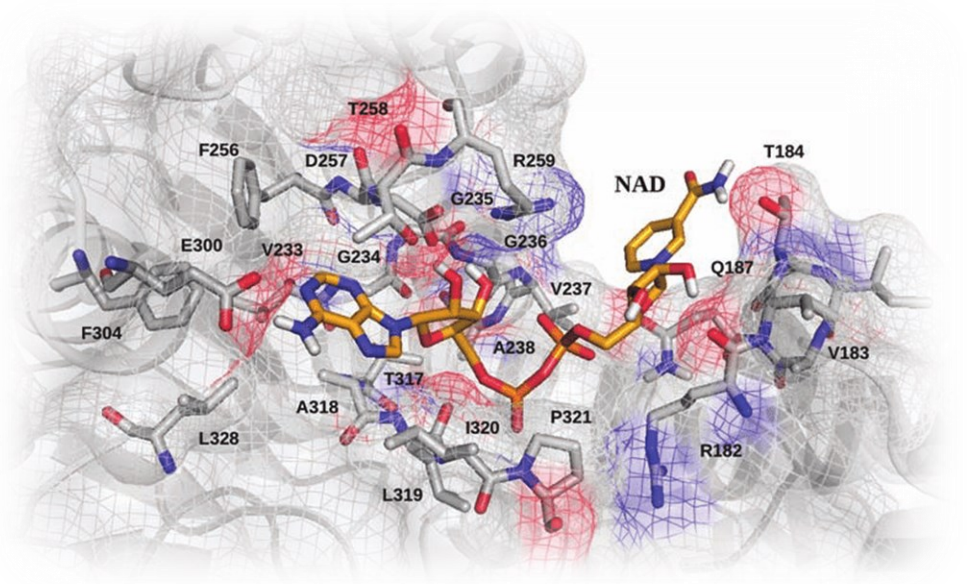

Fig. 4: 3D model structure of the NAD-binding site of NBPs. NAD is represented in yellow. (Metherell *et al.,* 2016)

The NAD-dependent enzymes group includes the previously mentioned sirtuins, as well as cyclic ADP-ribosyl cyclases, and poly- and mono-(ADP-ribose) polymerases (Duarte-Pereira *et al.,* 2023). Sirtuins are NAD-dependent histone deacetylases that regulate several critical processes in both prokaryotes and eukaryotes. Various studies have shown that these proteins are involved in processes such as inflammation, metabolism, oxidative stress, apoptosis, ageing and more. For this reason, they are also considered potential therapeutic targets in various pathological contexts (Wu *et al.,* 2022). ADP-ribosyl cyclases (ADPRCs) are enzymes that catalyse the synthesis of cyclic ADP-ribose, a calcium-active second messenger, and ADP-ribose from NAD^+, and the synthesis of nicotinic acid adenine dinucleotide phosphate (NAADP+) from $NADP^+$. Within this family, the most studied and well-known enzyme in mammals is CD38, a single-pass transmembrane protein (Astigiano *et al.,* 2022).

1.4 The PARP superfamily

Poly-(ADP-ribose) polymerases (PARPs) are a group of NBPs that catalyse the transfer of one or more ADP-ribose residues onto specific target proteins as a post-translational modification (PTM), using NAD as a substrate to synthesize the reaction product (Amé *et al.,* 2004). The addition of one or more ADP-ribose residues can serve as a PTM that activates proteins involved in DNA damage repair, chromatin remodelling, genetic and epigenetic regulation of transcription, programmed necrosis, and much more (fig. 5) (Morales *et al.,* 2016).

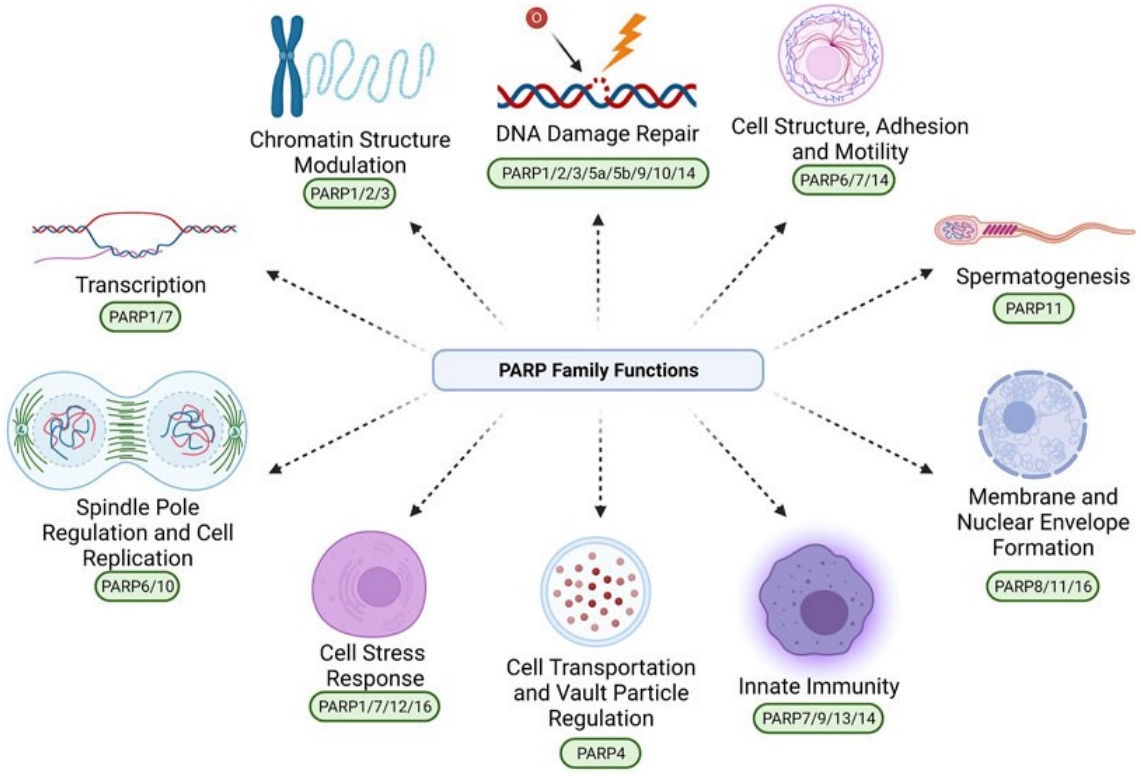

Fig. 5: confirmed and proposed roles of the PARP proteins. (Richard *et al.,* 2022)

In 2023, a study conducted by Suskiewicz *et al.* was published, proposing an updated classification and domains architecture definition of proteins belonging to the PARP family. This classification was based on structural models of the respective biomolecules generated by the AI algorithm AlphaFold2. First, the mentioned study specifies that, to date, there are 17 canonical members belonging to the PARP protein family. As also reported in other studies, such as the one by Richard *et al.,* 2022, these can be broadly divided into two main groups: the proteins that transfer single ADP-ribose residues, also known as MARTs (mono-(ADP-ribosyl) transferases); and the true PARPs (poly-(ADP-ribose) polymerases), which are enzymes that catalyse the transfer of chains of varying lengths of ADP-ribose residues (fig. 6).

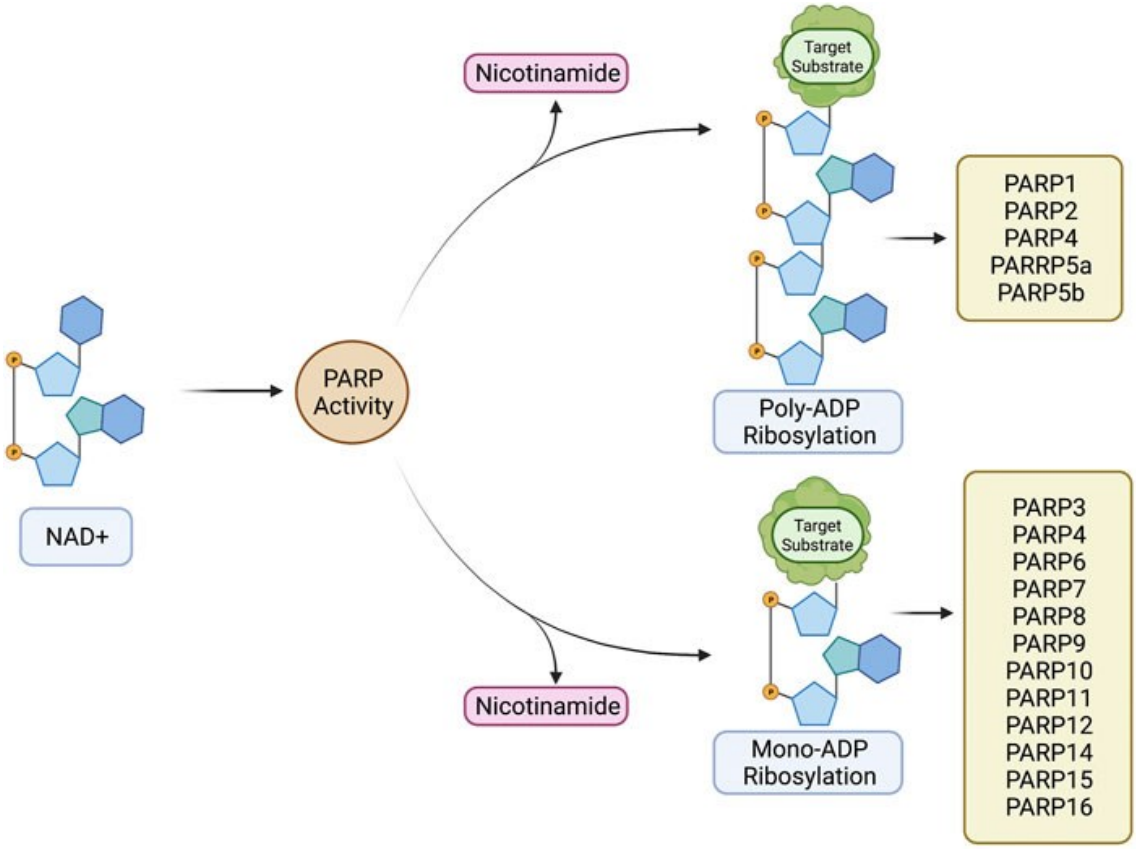

Fig. 6: PARP family-dependent poly- and mono-ADP ribosylation. (Richard *et al.,* 2022)

The multifunctionality of PARP proteins can be explained by the fact that the 17 members have quite different architectures, as highlighted in fig. 7. For some PARP proteins, in addition to the linear architecture of their constituent domains, the 3D structure is also well-known and defined, as is the case for PARP1 and PARP2. Consequently, their functions are well understood, making them potential pharmacological targets in certain pathological contexts, such as cancer (Morales *et al.,* 2016).

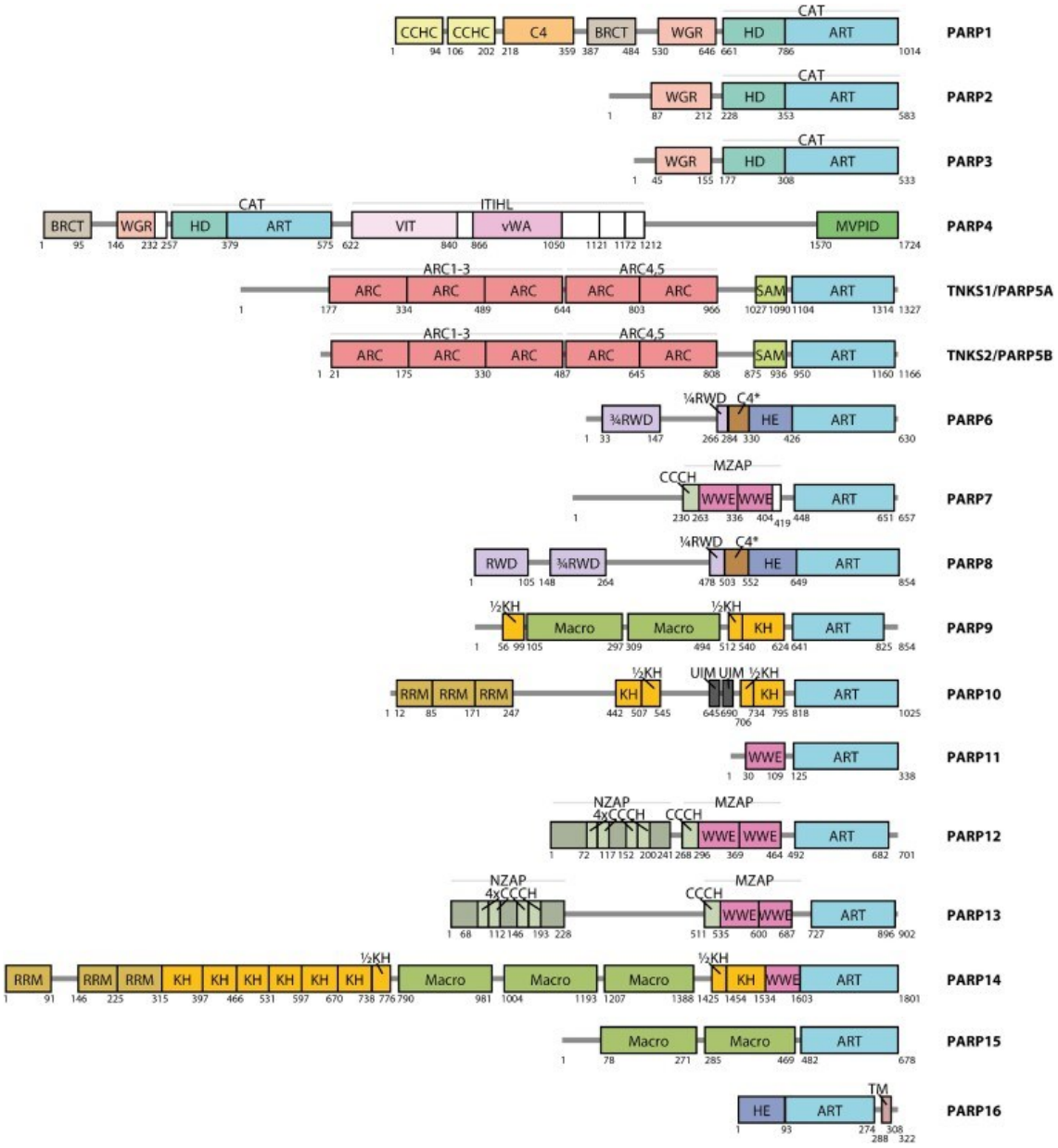

Fig. 7: domain architecture of the human members of the PARP protein family. Domains and other longer structured elements are represented as boxes and labelled with their respective domain or motif names. (Suskiewicz *et al.,* 2023)

Within the PARP protein family, PARP1 and PARP2 are undoubtedly the most studied and well-known. PARP1 is a poly-(ADP-ribose) polymerase involved in the cellular response to genotoxic stress by PARylating (i.e., attaching poly-(ADP-ribose) chains to) the DNA near

the damaged site. This site is recognized by the protein through interaction with the nucleic acid via its N-terminal domain, thereby recruiting a variety of proteins and factors involved in DNA damage repair (Dawicki-McKenna *et al.,* 2015). Specifically, through this action, PARP1 is involved in several DNA repair mechanisms: base excision repair (BER), homologous recombination (HR), non-homologous end joining (NHEJ), and nucleotide excision repair (NER) (De Vos *et al.,* 2012; Robu *et al.,* 2013). In addition to this role, PARP1 appears to be involved in other cellular processes, such as gene expression, chromatin remodelling, and programmed cell death (PCD) mechanisms (Gibson and Kraus, 2012). PARP2 is another PARP protein with functions that are analogous to those described for PARP1. Like PARP1, it operates within the nucleus, playing a key role in the previously mentioned DNA damage repair mechanisms. PARP3 also has a similar functionality to the others (Suskiewicz *et al.,* 2023).

Since these proteins are involved in various DNA damage repair mechanisms and given that cancer cells have an exceptional ability to efficiently repair these breaks using PARPs, the idea arose to develop PARP inhibitors (PARPi) as anti-tumor agents (fig. 8). These drugs work by deactivating these proteins, preventing the repair of DNA damage and inducing the death of cancer cells. The use of PARP inhibitors is particularly indicated for patients who are unable to produce BRCA1 or BRCA2 proteins, key players in homologous recombination for repairing double-strand breaks (DSBs) in DNA, due to mutations in the corresponding genes (a condition known as *BRCAness*). In these cases, it is possible to induce “synthetic lethality”: the patient is treated with a chemotherapeutic agent (e.g., cisplatin) capable of inducing single-strand breaks (SSBs) in DNA. The patient is then treated with a PARP inhibitor, which blocks PARP proteins from resolving the SSBs, leading to the formation of DSBs. However, due to the patient's *BRCAness* condition, these DSBs cannot be repaired either. This results in unresolved damage in the cancer cell, leading to induced cell death. This strategy is promising because cells deficient in BRCA1 or BRCA2 are much more sensitive to PARP inhibition (Farmer *et al.,* 2005; Lord *et al.,* 2017).

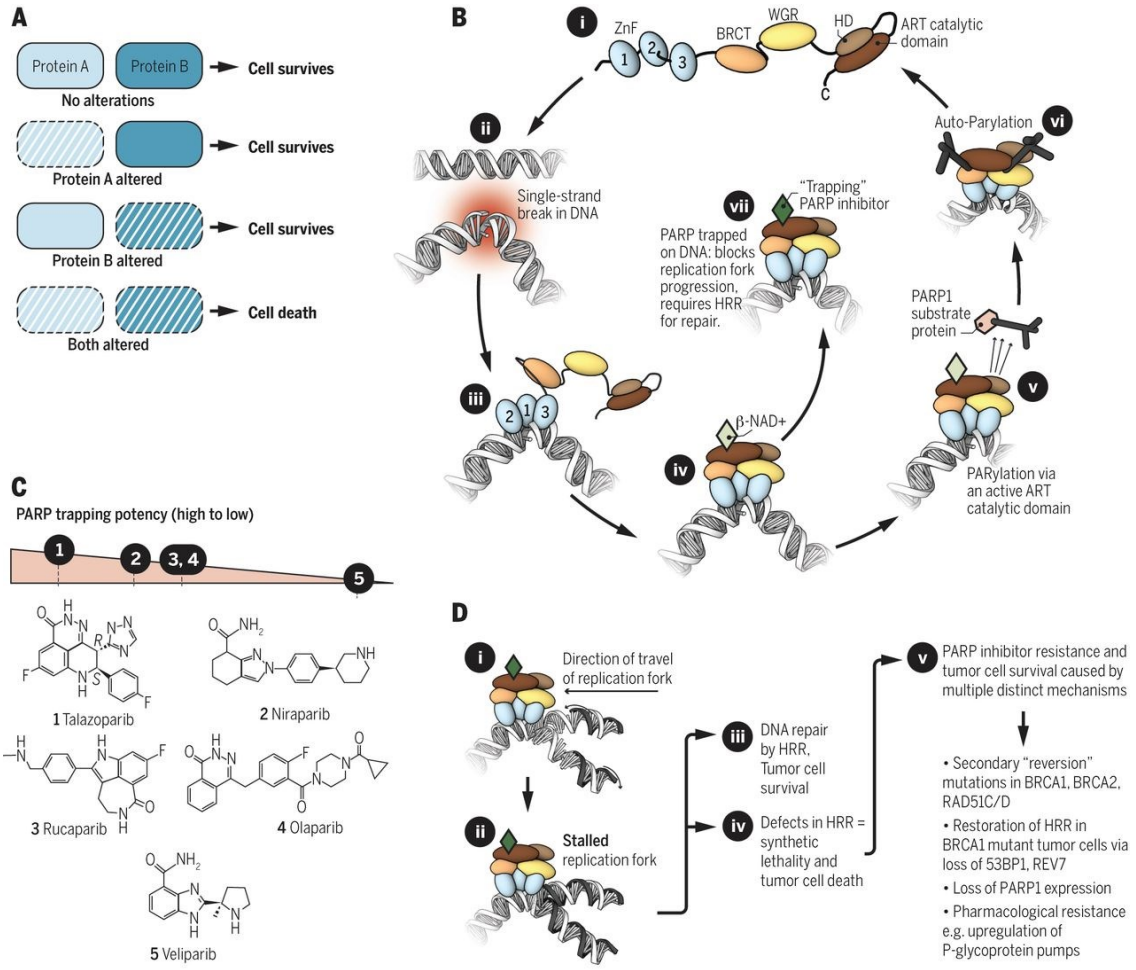

Fig. 8: mechanism of action of PARPi. (A) Logic of synthetic lethality. Only the simultaneous mutation of two genes or proteins (referred to as A and B) induces cell death. (B) Description of the PARP1 activity cycle. (C) Clinical PARP inhibitors. (D) A model of PARP inhibitor synthetic lethality. (Lord *et al.,* 2017)

In addition to these, there are other PARPs that, while certainly less well-known than the previous ones, have still been partially studied in terms of structure and function.

PARP4 has been identified as an associated component of vaults, ribonucleoprotein structures in eukaryotic cells that remain largely unknown (Kickhoefer *et al.,* 1999). This association with vaults occurs through its C-terminal end (MVP-interacting domain), while the N-terminal is homologous to that of PARP1, suggesting that this PARP might also be capable of recognizing DNA breaks (Suskiewicz *et al.,* 2023). Other studies suggest a possible involvement of PARP4 in the modulation of innate immunity (Garantziotis *et al.,* 2007).

Tankyrases PARP5A and PARP5B are PARylating enzymes involved in cellular signaling mechanisms. Within the PARP protein family, they are the only ones that contain ankyrin repeat cluster (ARC) domains, which serve as substrate recruitment sites (Guettler *et al.,* 2011).

PARP6, PARP8, and PARP16 are characterized by a helical appendage in the ART domain, which may have a function in protein-protein interactions (Karlberg *et al.,* 2012).

PARP7, PARP11, PARP12, and PARP13 form a subset of PARPs with a WWE domain, named after the conserved amino acid residues, that characterize them, and whose structure is very similar to the β-grasp fold of ubiquitin (Aravind, 2001).

Finally, PARP9, PARP10, PARP14, and PARP15 form a subset characterized by unannotated KH domains, sequence-specific RNA- or, more rarely, single-stranded DNA-binding modules (Valverde *et al.,* 2008). In PARP9, PARP10, and PARP14, one KH domain is split by a long insertion, which contains macro-domains in PARP9 and PARP14 that seem to function as ADP-ribose binding domains (Ekblad *et al.,* 2018; Schuller *et al.,* 2017). In PARP10, this insertion contains different domains, known as UIMs, which may recognize polyubiquitin chains (Nicolae *et al.,* 2014). In all cases, an additional KH domain with an extended helix is located between the split KH domain and the ART domain, connected to the ART domain by a loop or, in the case of PARP14, by a WWE domain (He *et al.,* 2012).

1.5 PARP15

PARP15, also known as BAL3, belongs to the same subset of PARP9, PARP10 and PARP14 and has been found only in humans and related species (Perina *et al.,* 2014). The structure, as well as the function, of PARP15 is largely unknown to date. Regarding structural information, based on predictions from the AI algorithm AlphaFold2, it appears to be composed of two macro-domains followed by the ART domain, which remain highly flexible, and it does not possess RNA-binding domains (fig. 9).

In 2024, David Baker, Demis Hassabis, and John Jumper, the founding fathers of the AlphaFold algorithm, were awarded the prestigious Nobel Prize in Chemistry (Nobel Prize

in Chemistry 2024). This recognition highlights the tremendous impact and invaluable insights this AI tool provides in predicting previously unknown protein structures.

Fig. 9: 3D structure of PARP15 predicted by AlphaFold2.

Currently, the information available about this protein is rather limited. It is known to be a MART protein, which carries out mono-(ADP-ribosylation), and is associated with mechanisms of telomere maintenance, transcriptional regulation, stress response, and immune signaling (Gupte *et al.,* 2017). However, its exact role in these processes has not yet been fully clarified.

Recent studies, such as the one by Zhou *et al.,* 2022, have elucidated the structure of the catalytic domain of PARP15 and how it can interact with specific inhibitors designed to target this pocket (fig. 10).

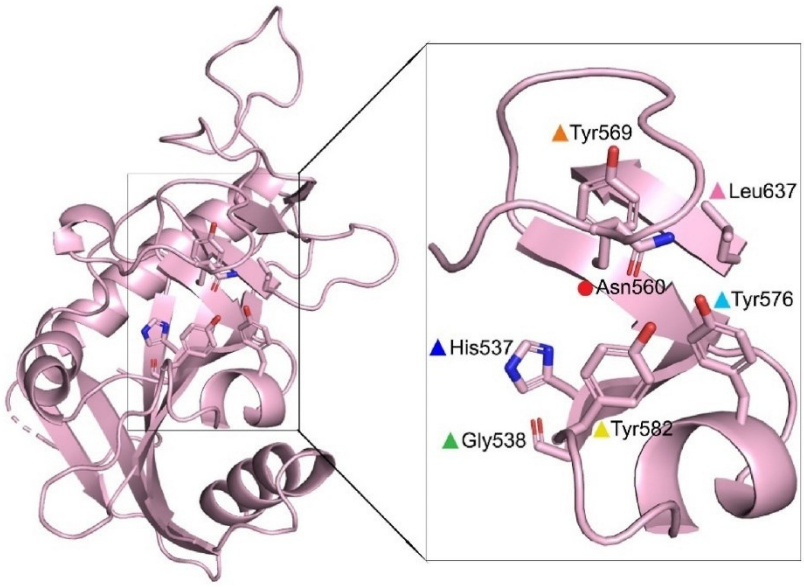

Fig. 10: structure of the active catalytic domain of PARP15. (Zhou *et al.,* 2022)

The crystallographic structure of the full-length protein has not yet been deposited in the PDB, so there is still no confirmation that it matches the prediction by AlphaFold2. In contrast, the catalytic domain has been reported as crystallized in dimeric form.

The same study by Zhou *et al.,* 2022 also elucidated the structure of the NAD-binding pocket, the active site for certain PARP inhibitors, in PARP15. As shown in fig. 11, it consists of a donor site (NAD^+ site) and an acceptor site. The donor site contains a glycine residue at position 558, replacing a flexible tyrosine residue found in tankyrases, where it is responsible for the opening and closing mechanism of the binding site for the nicotinamide portion of NAD, known as the "NI-subsite". In PARP15, this subsite is recessed into a deep pocket within the protein, without connecting to the acceptor site exposed on the opposite side. Furthermore, the binding site for the adenosine portion of NAD, known as the "AD-subsite", is surrounded by negatively charged amino acid residues.

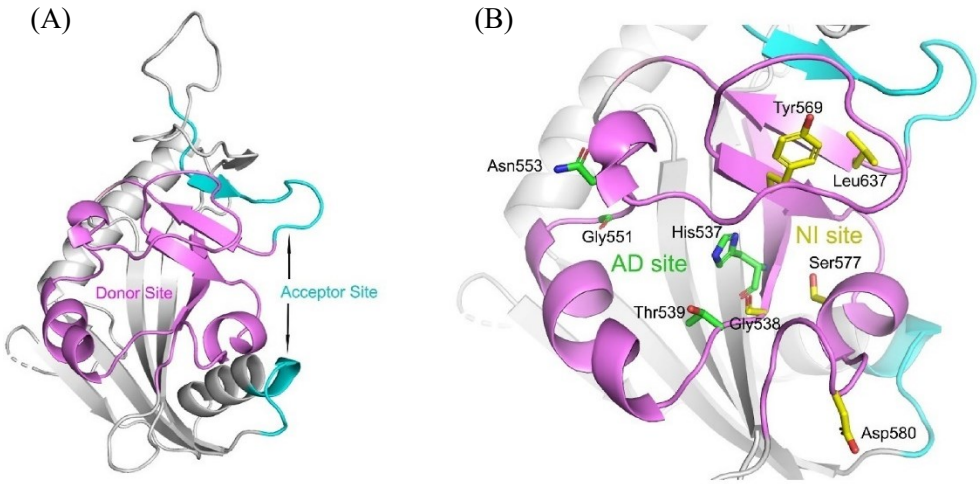

Fig. 11: (A) The donor site (purple) and acceptor site (blue) in the PARP15 catalytic domain. (B) The AD-subsite and NI-subsite in the donor site of PARP15 with certain conserved residues shown in sticks. Residues in the AD site are colored in green and those in the NI site are in yellow. (Zhou *et al.,* 2022)

1.6 The importance of structural biology in drug development

The basic research is the starting point from which the entire drug development process can take shape and develop, eventually leading to the actual drug available for the treatment of a particular disease. More specifically, determining the three-dimensional structure of the pharmacological target is essential to initiate structure-based drug design studies. This approach allows for the identification of a series of potential molecules, ranging in complexity, that could interact with the target at a specific domain. This interaction between the drug and the target biomolecule is the key event that leads to the desired therapeutic effect.

Once the 3D structure of the target biomolecule is determined, it is necessary to identify possible drug binding sites and characterize the interactions that occur between the two entities involved. During the drug design phase, this information enables the initiation of molecular modelling experiments, often computational in nature. Starting from a compound library, these molecules are scanned by docking them with the target moiety. In this way, compounds that could potentially have the desired effect are selected, allowing for subsequent chemical studies aimed at optimizing their structure to refine interactions at the

binding site ("lead optimization" phase). Generally, these efforts aim to maximize affinity and selectivity towards the target biomolecule. Furthermore, based on the target's structural information, molecular modelling also allows for "*de novo* design" experiments, which involve virtually "building" new molecular entities that might interact with the target (Lescrinier, 2011).

This context directly relates to the study presented here, which lays the groundwork for structural and activity analyses of PARP15, a new potential pharmacological target in cancer research that remains largely unknown and underexplored.

2. Aim of the work

In cancer research, the primary goal is to identify potential targets involved in the mechanisms of tumor growth and progression to subsequently develop drugs specifically designed to target them, following the concept of targeted therapy.

If the target is a protein, it becomes necessary to produce it in recombinant form so it can be characterized. Most of the anticancer drugs developed to date are inhibitors or activators of specific target proteins involved in various pro-tumor processes.

Having the protein available in recombinant form is essential for:

- Performing targeted biochemical and biophysical assays to define its activity and conducting screenings of libraries of potential inhibitors or activators.
- Conducting crystallographic studies aimed at defining its 3D structure and assembly, information that is crucial for the subsequent phase of structure-based drug screening.

The host system used is selected based on the desired protein yield and the intended application. Moreover, it is often not even necessary to produce the full-length recombinant protein; in fact, it is frequently sufficient to have a protein variant capable of supporting the biochemical activity under study (Gräslund *et al.,* 2008).

This thesis project is part of a broader series of studies aimed at defining the structural and biochemical characteristics of proteins within the NBPs family. This work fits and focuses on the very first steps of PARP15 protein analysis. PARP15 remains largely unknown both in terms of its structure and physiological role, as well as its potential involvement in pathological contexts.

Specifically, the work carried out at the hosting biochemistry and structural biology lab of Professor Riccardo Miggiano aimed to carry out expression and purification of the PARP15 protein in two different forms:

- A truncated variant, called PARP15_78-678, which lacks the first 77 amino acids of the N-terminal end that, according to structural predictions from the AI program AlphaFold2, tends to maintain an unfolded structure that could potentially interfere with the recombinant expression in the host system.

- The full-length protein, called PARP15_FL.

For this project, the aim was to develop the expression and purification protocols for both variants of PARP15 to obtain them at high concentrations and purity level, making them available for subsequent structural and functional studies, as well as for testing potential protein inhibitors.

Regarding the determination of the structure of PARP15, it is true that the scientific community now has access to highly sophisticated AI algorithms capable of predicting the 3D structure of any protein from its primary sequence (e.g., AlphaFold2). However, these programs still have limitations, and their predictions are not 100% certain. For this reason, classical structural biology techniques (such as X-ray crystallography, 2D and 3D NMR, and Cryo-EM) are essential to confirm and enrich data.

The therapeutic implications of targeting PARP15 extend beyond cancer treatment. Recent findings suggest that PARP15 may play a role in acute myeloid leukemia (AML), positioning it as a candidate for therapeutic intervention in this malignancy (Aljabban, 2023). Furthermore, the development of selective PARP15 inhibitors could provide insights into the broader biological functions of ADP-ribosylation in various diseases, including inflammatory conditions and neurodegenerative disorders (Li *et al.,* 2022). In conclusion, the discovery of selective small molecule inhibitors for PARP15 represents a promising frontier in drug development. By leveraging advanced screening techniques and structural biology, researchers are making significant strides in identifying compounds that can effectively target this enzyme, potentially leading to novel therapeutic strategies for cancer and other diseases.

3. Results and discussion

3.1 Isolation of high-purity pET-28a (+) plasmids from *E. coli* DH5α cells for subsequent transformations

According to the protocol provided by Jena Bioscience (see *Materials and Methods* chapter), triplicates of the plasmids encoding PARP15_FL and PARP15_78-678 were quantified using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific), yielding the following results:

	PARP15_FL (ng/µL)	PARP15_78-678 (ng/µL)
A	574.1	554.3
B	698.5	715.5
C	741.0	629.6

Table 1: plasmid concentrations obtained after their amplification.

The plasmid samples were finally diluted to final concentrations of 10 ng/µL.

These data demonstrate the high amplification efficiency that can be achieved using the *E. coli* DH5α strain for these purposes, highlighting their genetic predisposition that facilitates achieving such results.

3.2 Identification of expression conditions for PARP15_78-678 and PARP15_FL

3.2.1 Test expression in 2X YT medium

To define best expression condition for PARP15_78-678 and PARP15_FL, test expression was performed making *E. coli* BL21 (DE3) bacterial cells growing in 2X YT medium and inducing protein expression (see *Material and Methods* chapter), following the conditions reported in table 2.

Condition	Temperature (°C)	Time of growth	IPTG concentration (mM)
1	37	4 hours	0.3
2	37	4 hours	0.5
3	37	4 hours	0.7
4	25, then 16	1 hour, then o.n.	0.3
5	25, then 16	1 hour, then o.n.	0.5
6	25, then 16	1 hour, then o.n.	0.7

Table 2: expression conditions tested with 2X YT medium.

Final SDS-PAGE analyses of the total, pellet, soluble, and elution fractions from the small-scale Ni-NTA affinity chromatography are reported below:

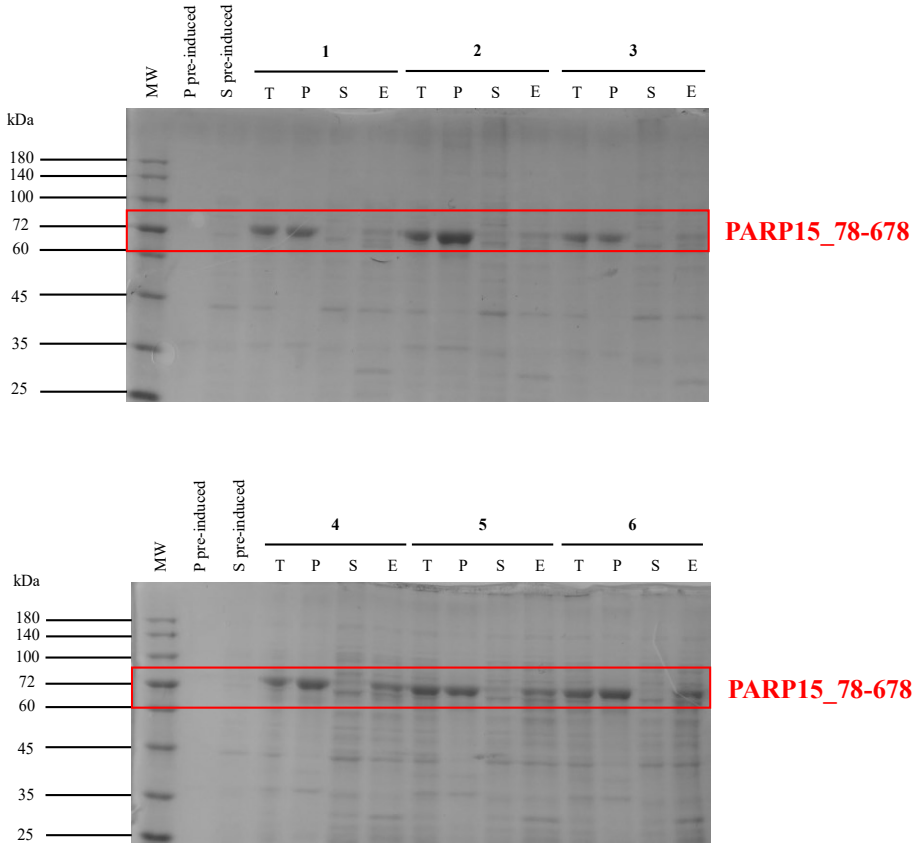

Fig. 12: SDS-PAGE (12% polyacrylamide gels) referring to T (total), P (pellet), S (soluble) and E (elution) samples of the pre and post induced cultures of PARP15_78-678, using 2X YT medium and induction with IPTG. Expression conditions from 1 to 6 are described in *Material and Methods* chapter.

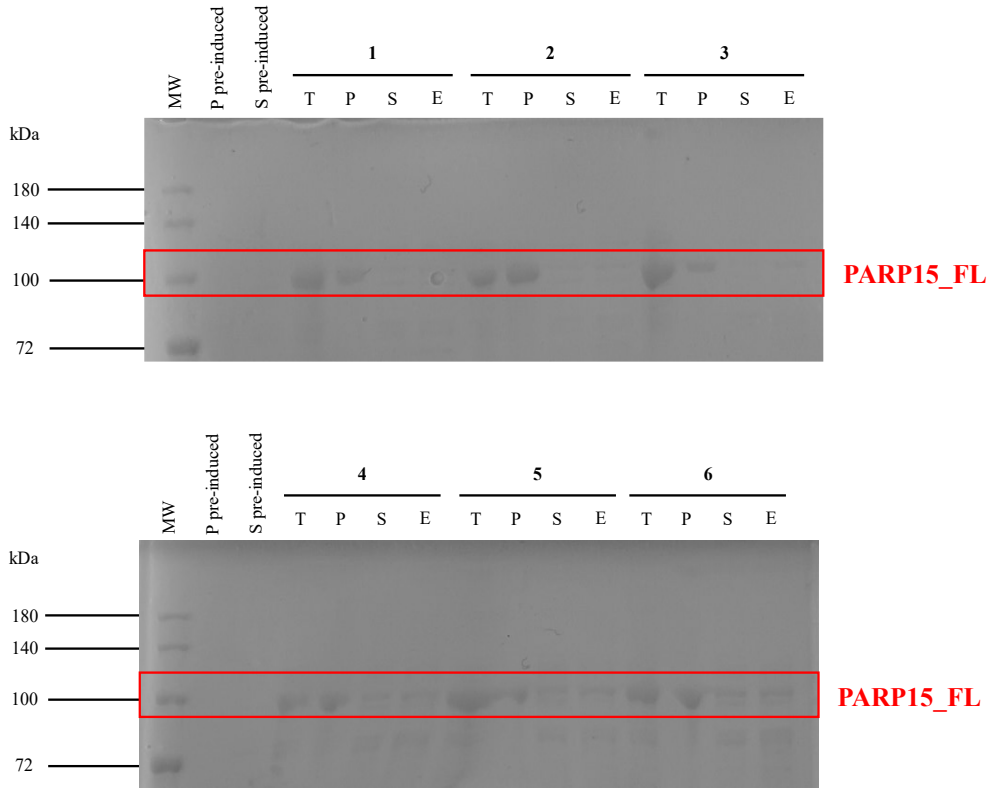

Fig. 13: SDS-PAGE (12% polyacrylamide gels) referring to T (total), P (pellet), S (soluble) and E (elution) samples of the pre and post induced cultures of PARP15_FL, using 2X YT medium and induction with IPTG. Expression conditions from 1 to 6 are described in *Material and Methods* chapter.

Results of the expression tests for PARP15_78-678 and PARP15_FL indicate that proteins are undoubtedly the most highly expressed. However, their abundance relative to other proteins was not significantly higher in the soluble fractions, that's why we decided to analyse the enriched proteins of interest coming from elution of small-scale Ni-NTA affinity chromatography for each expression condition. Furthermore, both proteins show a strong tendency to remain in their respective insoluble fractions (pellets), indicating high insolubility, despite the addition of the SUMO tag in the plasmid constructs. Their retention in the insoluble fractions prevents efficient recovery for subsequent purification steps.

One possible solution could be to increase the number of sonication cycles applied to the samples, although this might lead to excessive stress, causing protein denaturation and other adverse effects.

Nevertheless, it is evident that certain expression conditions yielded better results than others. Specifically, a decent expression of PARP15_78-678 under conditions 5 and 6 (see section 5.6) can be noted, highlighted by the respective elution samples. For this reason, these two expression conditions were identified as the most favourable. An entirely similar result was observed for PARP15_FL, which also found conditions 5 and 6 (see section 5.6) to be the most effective among those tested.

3.2.2 Test expression in autoinduction medium

To define best expression condition for PARP15_78-678 and PARP15_FL, test expression was also performed making bacterial cells growing in autoinduction medium (see *Material and methods* chapter), following the conditions reported in table 3.

Condition	Temperature (°C)	Time of growth
1	37	3 hours
2	37	4 hours
3	37	6 hours
4	37, then 16	3 hours, then 1 hour
5	37, then 16	3 hours, then 3 hours
6	37, then 16	3 hours, then o.n.

Table 3: expression conditions tested with autoinduction medium.

Final SDS-PAGE analyses of the various expression conditions are described below:

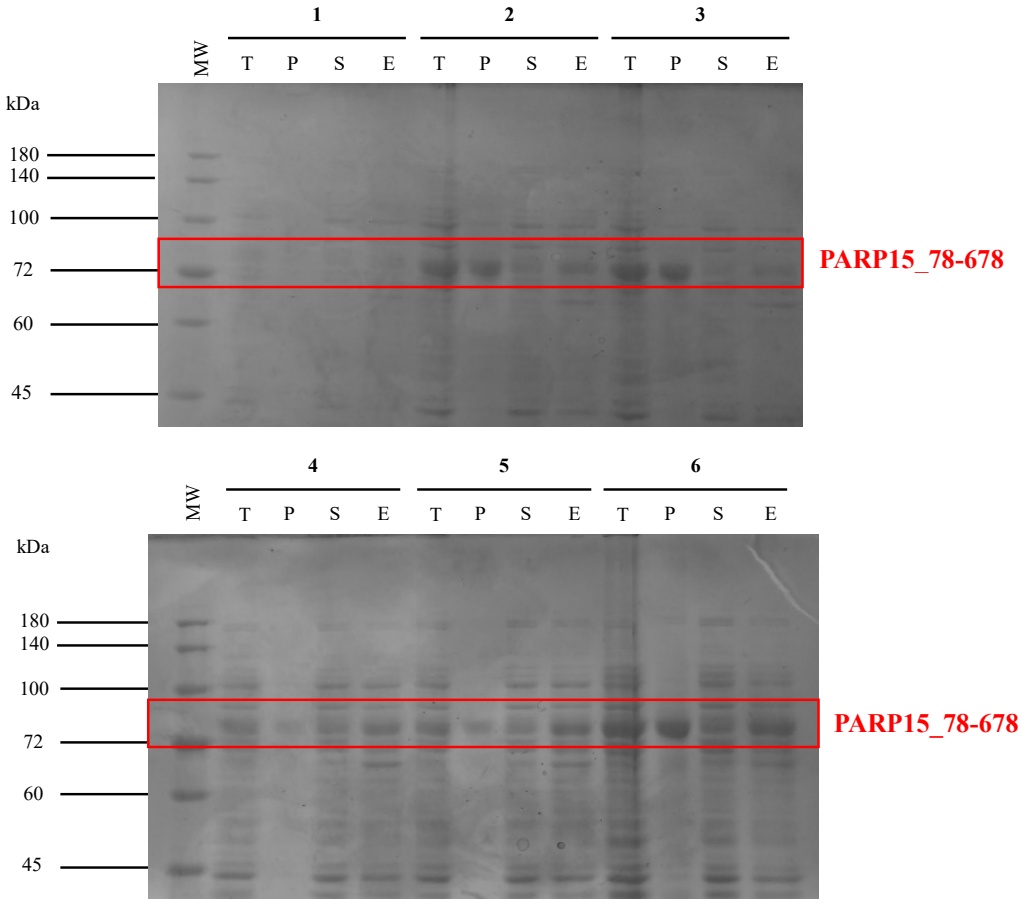

Fig. 14: SDS-PAGE (12% polyacrylamide gels) referring to T (total), P (pellet), S (soluble) and E (elution) samples of the pre (condition 1) and post induced cultures of PARP15_78-678, using autoinduction medium. Expression conditions from 1 to 6 are described in *Material and Methods* chapter.

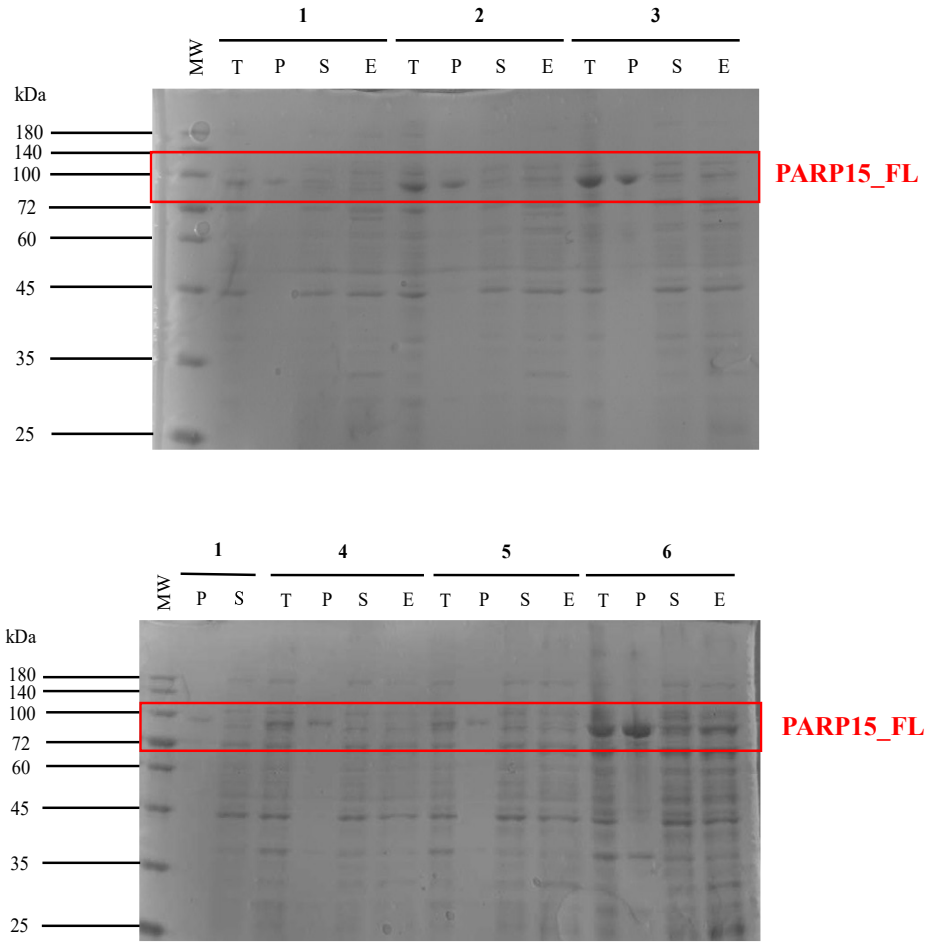

Fig. 15: SDS-PAGE (12% polyacrylamide gels) referring to T (total), P (pellet), S (soluble) and E (elution) samples of the pre (condition 1) and post induced cultures of PARP15_FL, using autoinduction medium. Expression conditions from 1 to 6 are described in *Material and Methods* chapter.

Once again, as previously highlighted, both proteins are the most highly expressed but tend to remain in their respective insoluble fractions, thus presenting the same issues as noted earlier. Therefore, changing the culture medium did not provide any advantages in this regard. However, the use of the autoinduction medium, which is richer in nutrients compared to 2X YT medium, leads to an overall increase in protein expression levels, and consequently, PARP15_78-678 and PARP15_FL are relatively more abundant. Furthermore, it is evident that the expression conditions involving a temperature switch to 16°C (conditions 4 to 6)

yield significant improvements over those maintained only at 37°C (conditions 1 to 3). Best expression conditions selected for both proteins were conditions 5 and 6 indeed.

By cross-referencing all data obtained from the expression tests, condition number 5 with autoinduction medium was chosen as the overall best condition to proceed with the subsequent purification trials for PARP15_78-678. While for PARP15_FL, condition number 6 with autoinduction medium was selected as the most suitable for possible purification trials.

3.3 Purification trials of PARP15_78-678

3.3.1 First purification

After processing 7.8 g of pellet of transformed cells grown under the best expression conditions, the protein was purified. The first step involved a batch Ni-NTA affinity chromatography, performed following the protocol described in the *Materials and Methods* section. The final SDS-PAGE analysis yielded the following results:

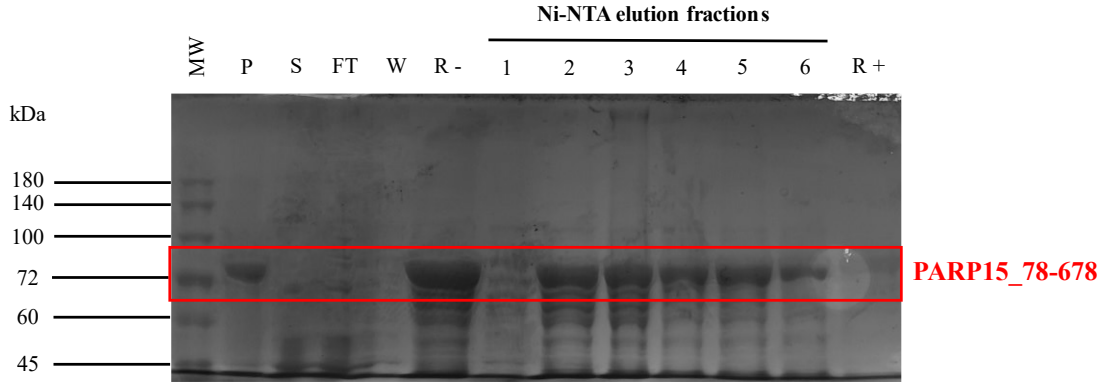

Fig. 16: SDS-PAGE (12% polyacrylamide gel) referring to Ni-NTA purification.

The six elution fractions were quantified using Bradford assay with a spectrophotometer, resulting in the following protein concentration values:

Fraction	Concentration (mg/mL)	mg of proteins
1	0.75	1.12
2	5.00	7.50
3	6.93	10.39
4	2.60	3.90
5	2.56	3.84
6	1.77	2.66
TOTAL	/	29.41

Table 4: results of the Bradford assays performed.

The six elution fractions were pooled and subsequently concentrated using 50 kDa MWCO Vivaspin® 20 ultrafiltration units (Sartorius). During this step, Bradford assays were performed at roughly regular intervals to assess any potential protein loss. Final volume sample was 1 ml at 14.883 mg/mL.

This sample was loaded onto the Superdex® 200 Increase 10/300 GL column (Cytiva) SEC column, and the following chromatogram was obtained:

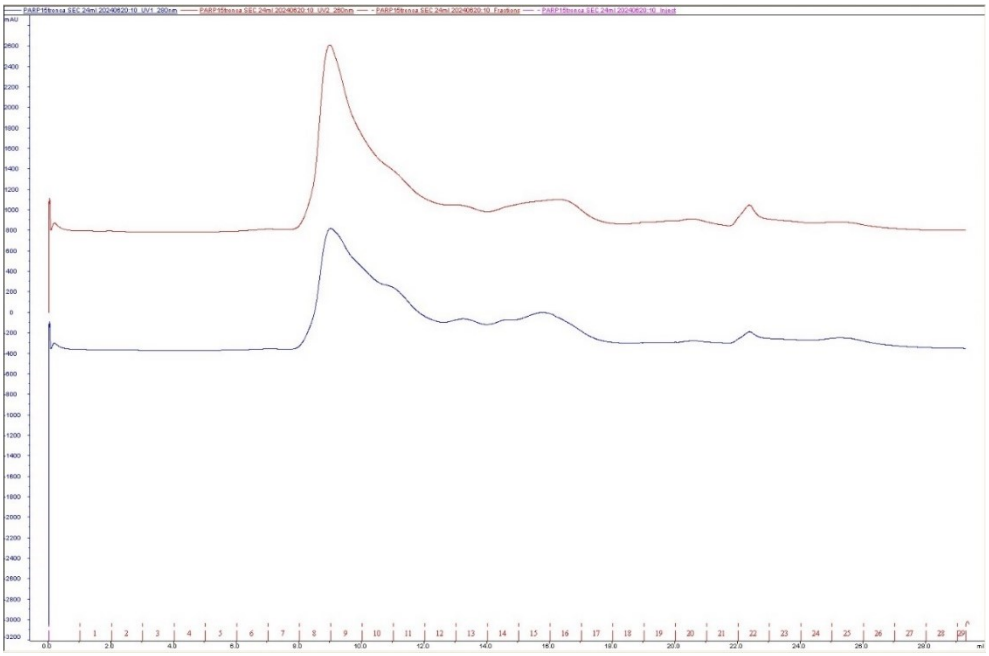

Fig. 17: SEC chromatogram of the first purification trial for PARP15_78-678. The red trace shows the variation in absorbance at 260 nm, while the blue trace shows the variation in absorbance at 280 nm. The numbers at the bottom of the chromatogram indicate the elution fractions.

The elution fractions corresponding to the peaks observed in the SEC chromatogram were collected and finally analysed on SDS-PAGE:

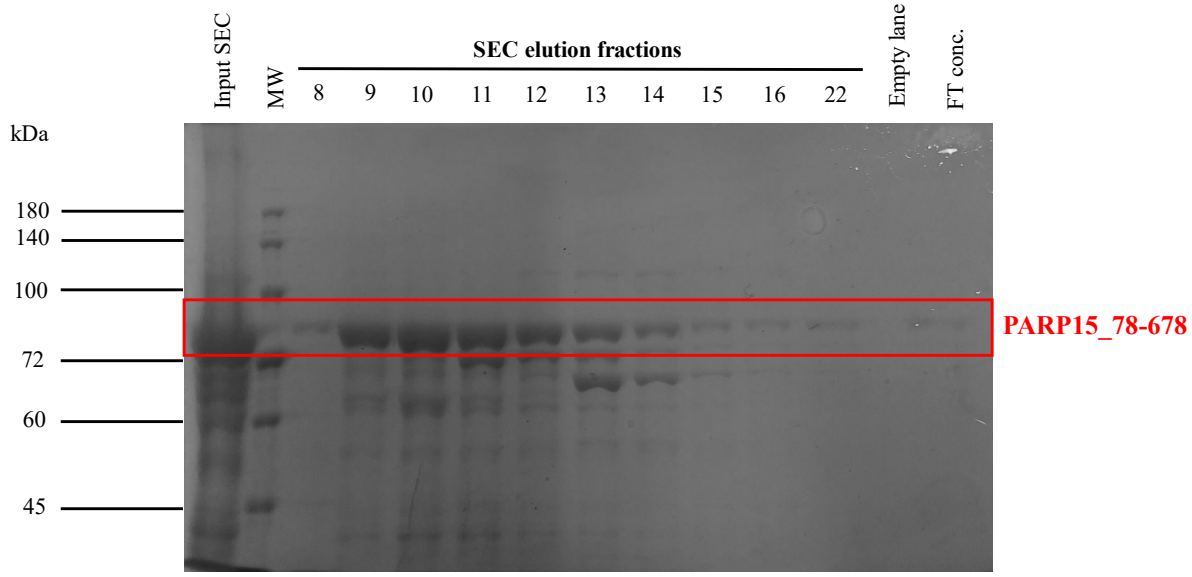

Fig. 18: SDS-PAGE (12% polyacrylamide gel) referring to SEC purification.

As can be seen in fig. 16, the protein is effectively eluted using Ni-NTA affinity chromatography; this result highlights the significant utility of the His-tag included in the expression plasmid construct. PARP15_78-678 responds well to Ni-NTA affinity chromatography, making this process an ideal initial purification step in the overall purification protocol. The protein is undoubtedly the most abundant among those present in the sample, indicating that the plasmid construct enables the overexpression of the recombinant protein in *E. coli*. After this stage, an additional chromatographic process is necessary to increase the sample's purity level, as there are still many contaminants that have also interacted with the nickel resin.

The protein concentrations of the Ni-NTA elution fractions are consistent with what is observed in the SDS-PAGE analysis. During concentration step, a significant amount of protein is lost; this may be due to the substantial stress placed on the protein component in the sample during this operational step, leading to breakdown and passage through the filter mesh, and possibly some adhesion to the concentrator filter as well. Nevertheless, the quantity of protein in milligrams at the end of the sample concentration process remains substantial.

However, the SEC chromatogram indicates that this first purification trial was unsuccessful, as the protein tends to elute immediately after the column’s exclusion volume rather than at the expected volume based on its molecular weight, suggesting protein aggregation. The SDS-PAGE analysis shows the presence of PARP15_78-678 in all elution fractions corresponding to the chromatogram peaks, confirming the observations mentioned above.

3.3.2 Second purification

The second purification trial (9.31 g of pellet) for PARP15_78-678 also included an initial step of Ni-NTA affinity chromatography, yielding the following results in terms of SDS-PAGE analysis and protein concentration measurement using the Bradford assay:

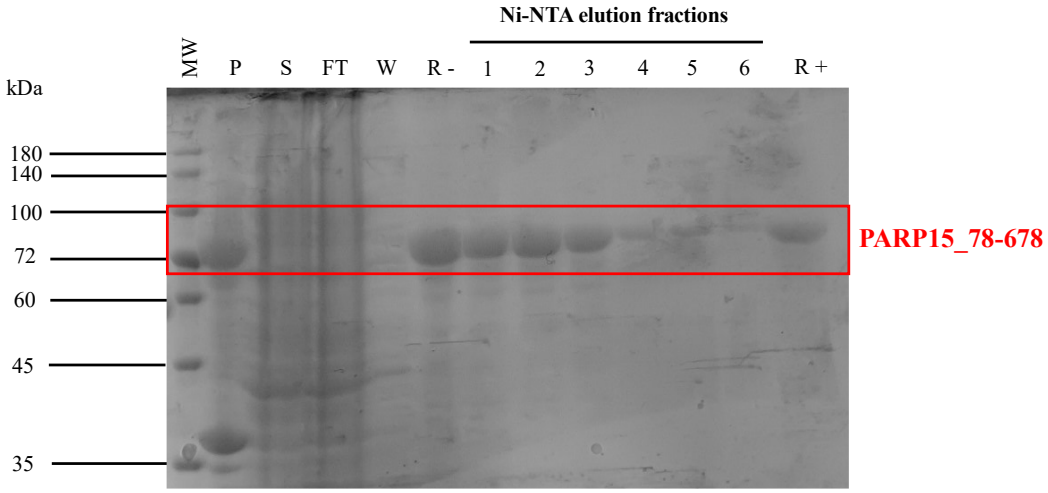

Fig. 19: SDS-PAGE (12% polyacrylamide gel) referring to Ni-NTA purification.

Fraction	Concentration (mg/mL)	mg of proteins
1	0.74	2.96
2	0.87	3.48
3	0.35	1.4
4	0	0
5	0	0
6	0	0
TOTAL	/	7.84

Table 5: results of the Bradford assays performed.

Subsequently, the Ni-NTA elution fractions were pooled and concentrated to a final volume of 5 mL, and the sample was injected into a desalting column to change the buffer, removing the imidazole and reducing the NaCl concentration. The resulting chromatogram is shown below:

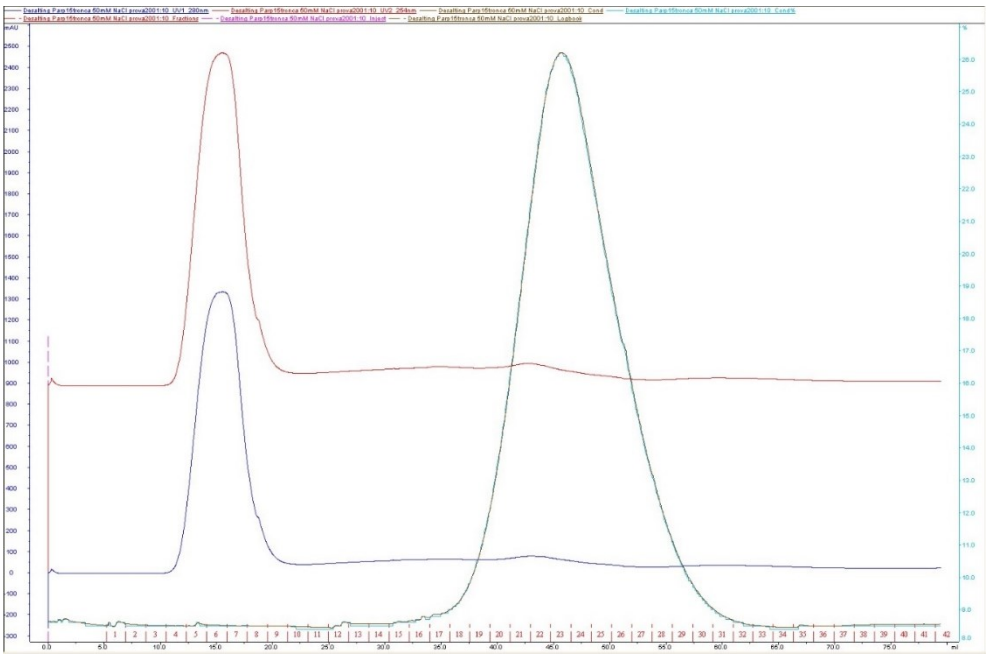

Fig. 20: desalting chromatogram of the second purification trial of PARP15_78-678. The red trace describes the variation in absorbance at 260 nm; the blue trace describes the variation in absorbance at 280 nm; the brown trace describes the variation in conductivity. The numbers at the bottom of the chromatogram indicate the elution fractions.

Then, the elution fractions from 4 to 9 obtained after the desalting step were pooled, and the sample was directly injected into 1 mL a HiTrap Q column (anion exchange chromatography). Elution was performed using a gradient strategy, gradually increasing the NaCl concentration (from 50 mM to 2 M) over 10 CV with HiTrap Q buffer. The following chromatogram was obtained from this process:

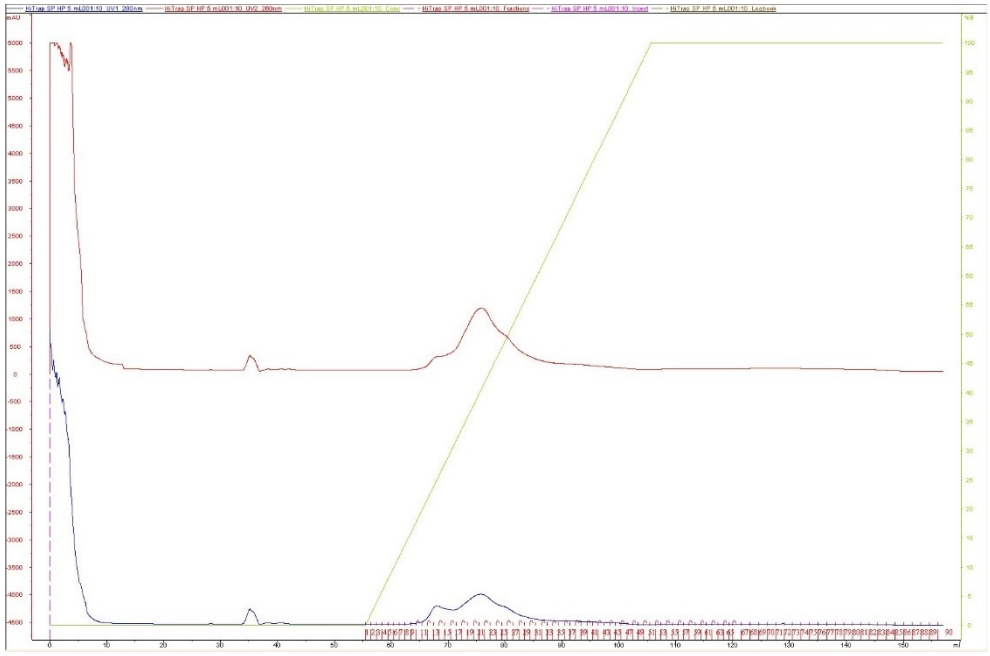

Fig. 21: anion exchange chromatogram of the second purification trial of PARP15_78-678. The red trace describes the variation in absorbance at 260 nm; the blue trace describes the variation in absorbance at 280 nm; the green trace describes the NaCl concentration gradient, from 50 mM to 2 M. The numbers at the bottom of the chromatogram indicate the elution fractions.

The 1 mL elution fractions corresponding to the chromatogram peaks were collected and loaded onto a gel for SDS-PAGE analysis:

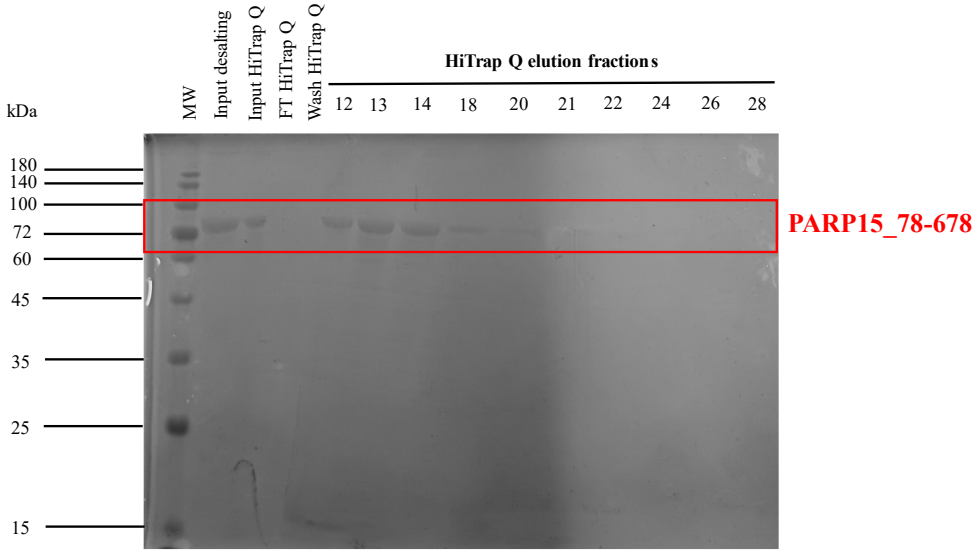

Fig. 22: SDS-PAGE (12% polyacrylamide gel) referring to anion exchange purification.

Examining the gel from the SDS-PAGE analysis, the elution fractions containing the protein of interest were quantified using a Bradford assay, yielding the following results:

Fraction	Concentration (mg/mL)	mg of proteins
11	0	0
12	0.19	0.19
13	0.23	0.23
14	0.35	0.35
15	0.19	0.19
TOTAL	/	0.96

Table 6: results of the Bradford assays performed.

Finally, these fractions were pooled and concentrated to a final volume of 0.5 mL, and a final Bradford assay was performed, indicating a final sample concentration of 0.3544 mg/mL (equivalent to 0.1772 mg of protein).

The results obtained provide some useful insights. First, the Ni-NTA affinity chromatography step is confirmed to be an effective procedure for the first step of the purification of PARP15_78-678; once again, the protein of interest proves to be the most highly expressed. However, based on the results from the Bradford assays, significantly less protein was obtained from the Ni-NTA affinity chromatography step compared to the first purification trial. Nevertheless, the elution fractions in this case are much purer, with a considerably lower number of contaminants present. This might suggest that, in the first purification attempt, the protein concentration values were “inflated” due to the presence of various contaminants. Additionally, comparing the results from these two purification trials highlights that increasing the imidazole concentration in the wash buffer (from 40 mM to 70 mM) positively impacts the outcome of the process, allowing more contaminants to be removed and resulting in elution fractions with a higher degree of purity. For this reason, the wash buffer used in this second purification trial has been chosen as the definitive one for the purification protocol.

The subsequent desalting step enabled an extremely efficient buffer exchange, as evidenced by the chromatogram in fig. 20. Two distinct peaks can be observed: the first one, associated with the curve describing absorbance measured at 280 nm, corresponds to the elution of the protein component in the sample. The second peak, associated with the curve showing the change in conductivity, corresponds to the elution of salts and imidazole. This indicates an optimal removal of these latter components from the protein, resulting in PARP15_78-678 being in a buffer that is optimal for the subsequent chromatographic step and free from imidazole.

The anion exchange chromatography step enabled the removal of additional protein contaminants from the sample. As clearly shown by the chromatogram in fig. 21, referring to the absorbance curve at 280 nm, two consecutive, low-intensity peaks, corresponding to the elution fractions from 11 to 16 and from 18 to 29, can be observed, which are partially overlapping. However, this partial overlap of peaks reflects an intrinsic resolution limit of the column used; with a different column offering higher resolving power, it might have been possible to separate PARP15_78-678 more distinctly from the contaminants. Nonetheless, sample appears to have reached a satisfactory level of purity, suggesting that this chromatographic strategy could be a valuable resource for the final purification step of PARP15_78-678. Bradford assays conducted on the corresponding elution fractions confirmed relatively low protein concentration values. Since obtaining protein samples suitable for subsequent studies of activity and, particularly, structural analysis by X-ray crystallography requires extremely pure and highly concentrated samples, the elution fractions were pooled, and a final concentration step was carried out.

The results from this second purification trial have demonstrated that this purification protocol could be a promising strategy for obtaining PARP15_78-678 in a pure form. However, a significant amount of protein is lost during the various processing steps, ultimately resulting in a low-concentration sample that is not suitable for further activity and structural studies. Trying to concentrate the sample to a very little volume (50 μL) means concentrate also contaminants; indeed, this procedure, applied to reach as higher concentration and purity as possible, turned out to be ineffective (as shown in fig. 23).

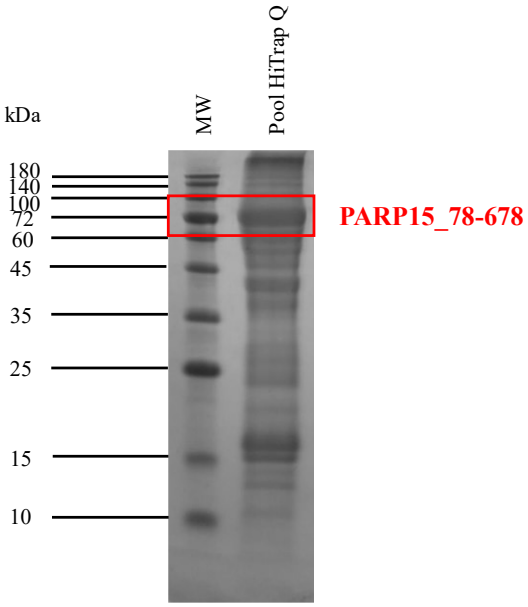

Fig. 23: SDS-PAGE (15% polyacrylamide gel) referring to pooled fractions after anion exchange purification and concentration.

3.3.3 Third and fourth purification

In this purification trial (6.13 g of pellet), the same steps described previously in the second purification attempt were followed, with the difference being that, in this case, the final chromatographic steps involved cation exchange chromatography and heparin chromatography instead of anion exchange chromatography. For this reason, to ensure the protein remained in a positively charged state, the buffers used after Ni-NTA purification step had a pH of 6 in this trial.

The elution fractions from the Ni-NTA affinity chromatography were pooled and quantified using the Bradford assay, which resulted in a protein concentration of 0.49 mg/mL in 18 mL of sample (corresponding to 8.82 mg of protein). The sample was then concentrated to a volume of 5 mL, and this concentrated sample was also quantified using the Bradford assay, yielding a concentration of 1.21 mg/mL (6.05 mg of protein). The sample was finally injected into a desalting column, which produced the following result:

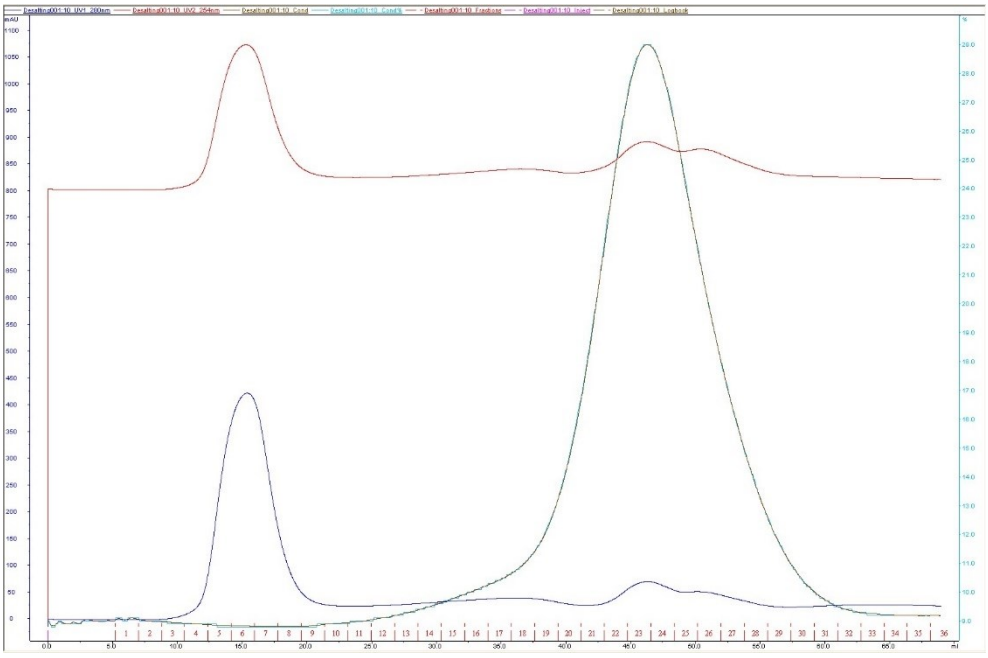

Fig. 24: desalting chromatogram of the third and fourth purification trial of PARP15_78-678. The red trace describes the variation in absorbance at 260 nm; the blue trace describes the variation in absorbance at 280 nm; the brown trace describes the variation in conductivity. The numbers at the bottom of the chromatogram indicate the elution fractions.

Fractions from 4 to 9 were recovered, pooled, and quantified using the Bradford assay, which indicated a protein concentration of 0.25 mg/mL in approximately 10.8 mL of pooled sample (corresponding to about 2.7 mg of protein). 5 mL of the pooled sample were injected into a cation exchange column, while another 5 mL were injected into a heparin column. In both cases, a gradient of increasing ionic strength was applied, up to 2 M NaCl and up to 500 mM respectively. The following chromatograms were obtained from these trials:

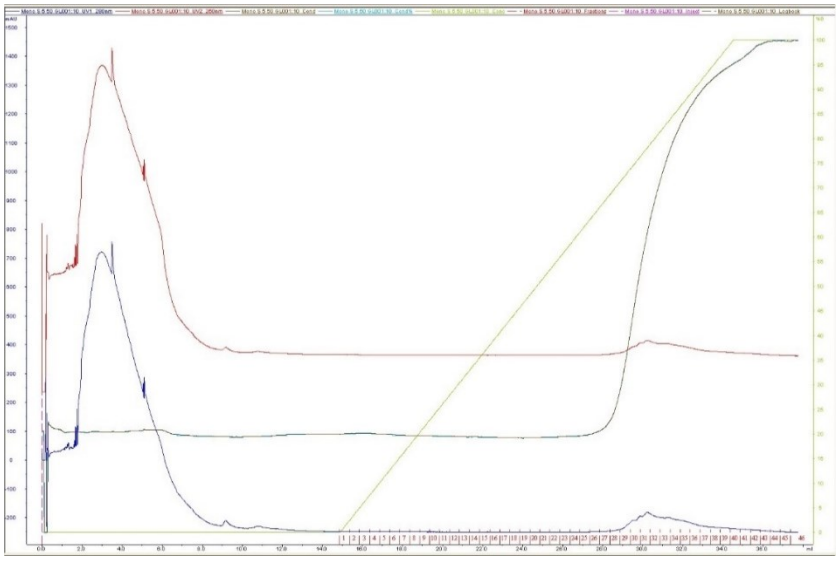

Fig. 25: cation exchange chromatogram of the third purification trial of PARP15_78-678. The red trace describes the variation in absorbance at 260 nm; the blue trace describes the variation in absorbance at 280 nm; the green trace describes the NaCl concentration gradient, from 50 mM to 2 M. The numbers at the bottom of the chromatogram indicate the elution fractions.

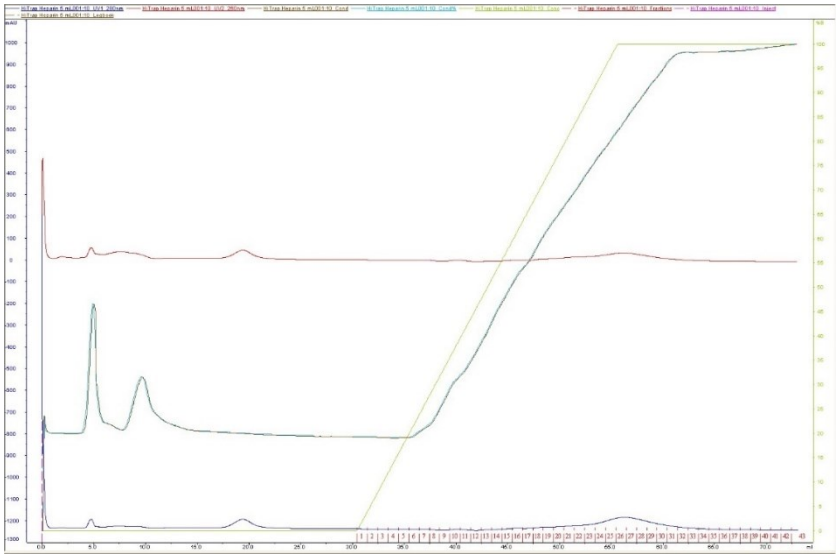

Fig. 26: heparin chromatogram of the fourth purification trial of PARP15_78-678. The red trace describes the variation in absorbance at 260 nm; the blue trace describes the variation in absorbance at 280 nm; the green trace describes the NaCl concentration gradient, from 50 mM to 500 mM. The numbers at the bottom of the chromatogram indicate the elution fractions.

Finally, an SDS-PAGE analysis was conducted on the elution fractions corresponding to the peaks from the cation exchange and heparin chromatograms, yielding the following results:

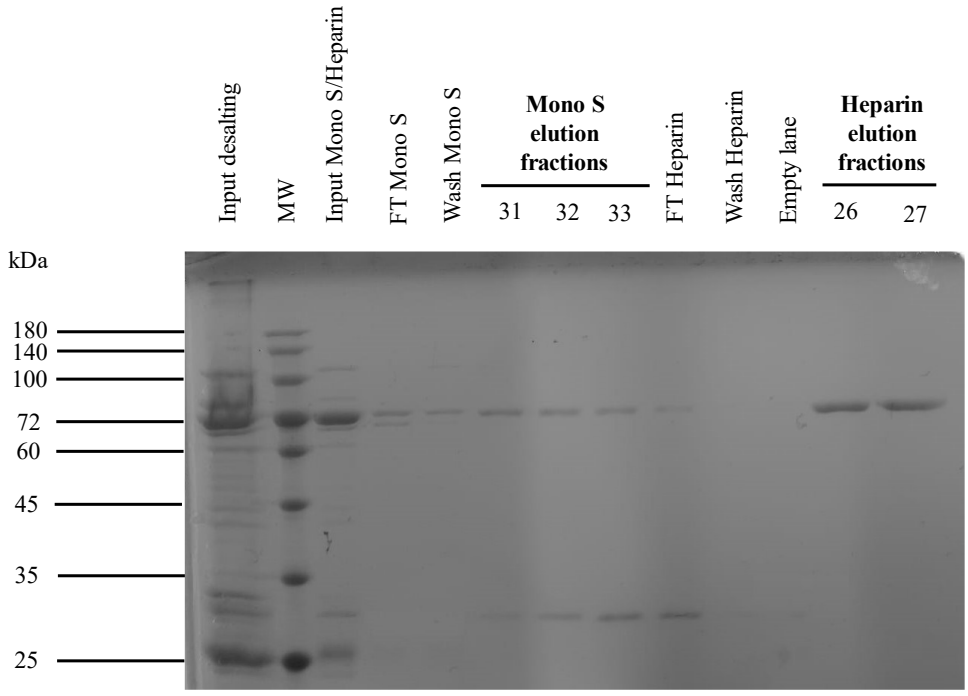

Fig. 27: SDS-PAGE (12% polyacrylamide gel) referring to cation exchange and heparin purifications.

Considering that Ni-NTA affinity chromatography was performed in the same way as the first and second purification trials and that the only difference is that the subsequent chromatographic steps buffer pH was lowered to 6, we may assume that this change is responsible for the almost complete absence of the protein of interest in the eluting fractions.

The final steps, namely cation exchange chromatography and heparin chromatography, yielded quite similar chromatograms. In both cases, a low-intensity peak can be observed, corresponding to the point of maximum ionic strength, as indicated by the absorbance curve at 280 nm. SDS-PAGE analysis of the elution fractions corresponding to these peaks confirmed this observation: for both cation exchange chromatography and heparin chromatography, distinct bands are evident at the marker level of 72 kDa. However, the cation exchange chromatography fractions also displayed several lower molecular weight

contaminants, which is why this protocol was not considered further, especially since the anion exchange chromatography used in the second trial produced more promising results.

The situation was different with heparin chromatography, as the SDS-PAGE analysis showed the exceptionally high purity achieved. However, it was immediately noticed that, as previously mentioned, the bands appear at the 72 kDa marker, while PARP15_78-678 has a molecular weight of approximately 80 kDa. This raised the question of whether the bands observed corresponded to the recombinant protein of interest or to a contaminant. To resolve this question, the heparin elution fractions corresponding to the peak of the chromatogram were pooled and concentrated to a final volume of 1.5 mL. The concentrated sample was quantified using a Bradford assay (0.1 mg/mL → 0.15 mg of protein), and a western blot analysis was performed following the protocol described in section 5.9 of the *Materials and Methods* chapter, using 4 µg of protein. The following result was obtained:

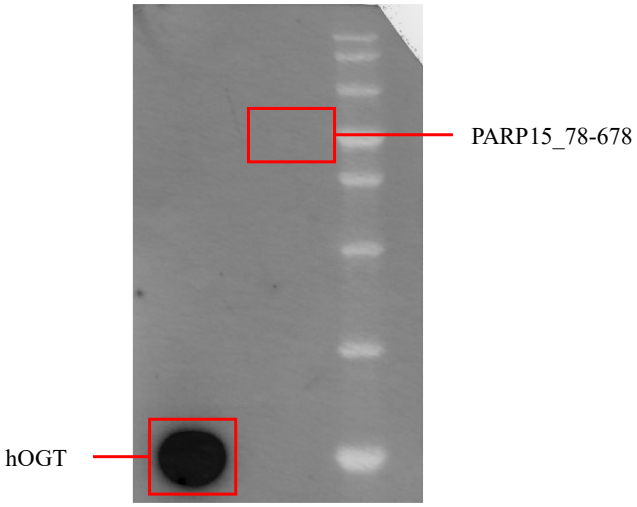

Fig. 28: western blot for the identification of PARP15_78-678. hOGT is the His-tagged protein control used.

As can be seen, the western blot confirmed that the bands observed in the SDS-PAGE were not associated with PARP15_78-678. The His-tagged hOGT protein control used demonstrated that the western blot was conducted correctly, so the absence of a signal corresponding to the recombinant protein of interest indicates that it was a contaminant. Based on this conclusion, neither of the strategies described were considered effective for the purification of PARP15_78-678.

3.3.4 Fifth purification

Finally, for the last purification trial (9.22 g of pellet), only one Ni-NTA affinity chromatography was performed, but following a different strategy described in detail in *Materials and Methods* chapter. Briefly, a stepwise imidazole concentration gradient was employed to remove as many contaminants as possible in a single chromatographic step. The final SDS-PAGE analysis of the various eluate fractions yielded the following results:

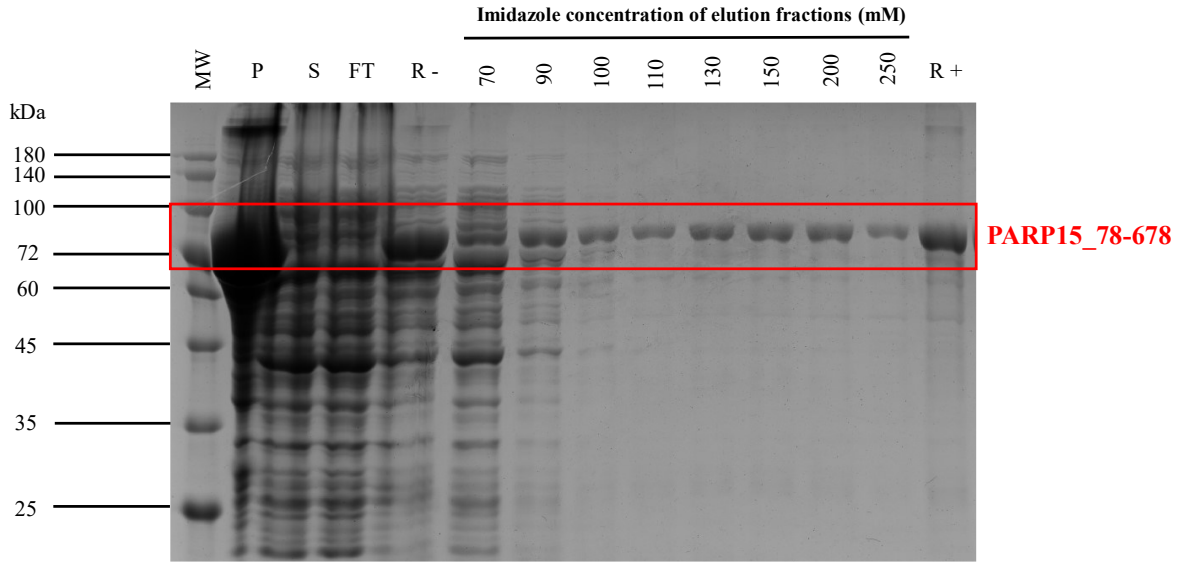

Fig. 29: SDS-PAGE (12% polyacrylamide gel) referring to Ni-NTA purification.

The 10 mL elution fractions from 110 mM imidazole to 250 mM imidazole were pooled and concentrated to a final volume of 1.2 mL. A Bradford assay was performed on the concentrated sample, yielding a protein concentration of 6.06 mg/mL (corresponding to 7.27 mg of protein). The concentrated sample was dialyzed overnight and then 4 µg of protein were loaded onto a gel for a final SDS-PAGE analysis:

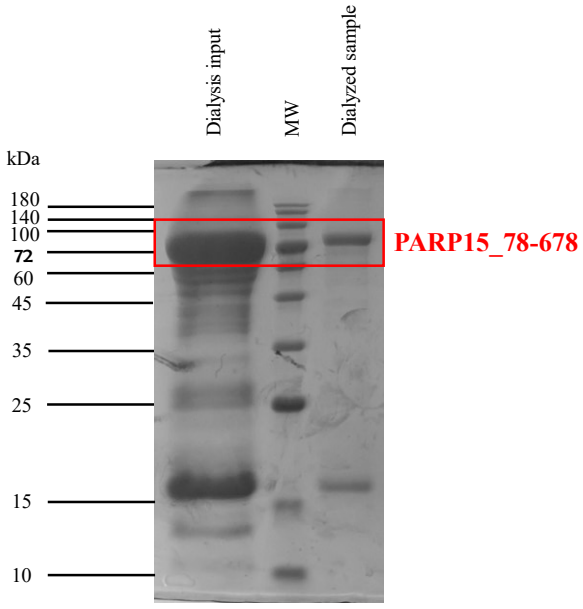

Fig. 30: SDS-PAGE (15% polyacrylamide gel) after dialysis step.

As shown in fig. 29, using this approach most contaminants elute at imidazole concentrations between 70 mM and 100 mM, which allows for a much more effective removal of contaminants compared to traditional Ni-NTA affinity chromatography. While it is true that during these elution fractions, PARP15_78-678 also partially elutes, it is acceptable to lose a small amount of the target protein to efficiently separate out the contaminants. Therefore, the fractions selected for pooling and concentration are those eluted between 110 mM and 250 mM imidazole. Despite this, these elution fractions still appear to be somewhat impure, as some contaminants, particularly those with lower molecular weights, are clearly present. Additionally, a significant amount of PARP15_78-678 remains bound to the resin at the end of the elution, indicating the potential for recovering further elution fractions to maximize the amount of protein available. In any case, this result has demonstrated that the Ni-NTA affinity chromatography step can be improved by increasing the imidazole concentration in the wash buffer up to 100 mM, effectively removing more contaminants and yielding PARP15_78-678 in a relatively pure form.

After concentrating, quantifying, and dialyzing the sample to remove imidazole, the final SDS-PAGE analysis of the sample shows that PARP15_78-678 still contains some

contaminants, one of which is particularly prominent around the 15 kDa marker (representing approximately 50% of the total protein content). The fact that this band was not visible in the gel in fig. 29 is associated with that gel being 12% polyacrylamide, while the one in fig. 30 is 15% polyacrylamide. This result suggests using gels with a polyacrylamide concentration of 15% to be able to visualize lower molecular weight contaminants that might otherwise go undetected.

Anyway, it is possible to conclude that the study has led to the development of an experimental protocol for obtaining PARP15_78-678 with a good level of purity and at a concentration suitable for inhibition studies and structural investigations using cryo-EM.

3.4 Crystallization of PARP15 Catalytic Domain

The final work step involved a crystallization trial of the Catalytic Domain (CD) of PARP15 in both ligand-free form and co-crystallized with the dual inhibitor PARP10/15-IN-3, as described in section 5.9 of the *Materials and Methods*. Currently, there are no results available, data from literature suggest that crystals will form after two weeks incubation at 20°C at set up conditions.

However, this crystallization trial can provide various useful insights; first, performing this test allows to verify the information reported in the literature and have an internal version of what has been described in other studies. Additionally, this procedure may lead to obtaining the crystallized structure of the catalytic domain in both the ligand-free and co-crystallized forms with the inhibitor, information that could be useful in the future to determine if the same conformation and interaction will be observed on PARP15_78-678. Furthermore, the availability of this inhibitor enables its use both during the purification phase of PARP15_78-678 to stabilize it and as a potential tool employed in biochemical characterizing assays.

4. Conclusions and future perspectives

The thesis project conducted at the hosting biochemistry and structural biology lab focused on the recombinant expression and purification of the PARP15 (BAL3) protein, a novel potential pharmacological target in oncology that has been scarcely studied.

Results highlight several obstacles that still need to be overcome. In particular, the expression tests conducted revealed significant insolubility of the protein, which posed a major limitation on the yield of recombinant protein available for subsequent purification steps. Purification trials encountered additional challenges in obtaining a sufficiently pure and concentrated sample. The results show the difficulty in developing a truly effective purification protocol for obtaining a high purity and concentrated sample. Thus far, the levels of purity and concentration achieved with a single Ni-NTA affinity chromatography step are not sufficient for structural studies using X-ray crystallography. Despite the difficulties encountered, it can be stated that the work has outlined the experimental protocol for obtaining PARP15_78-678 with a good level of purity and at a concentration suitable for inhibition studies and structural investigations using cryo-EM.

To achieve better yields, some alternatives remain viable: for example, it may be possible changing the expression system by using insect cells or redesigning the expression plasmid by adding a different tag (for instance, a GST-tag) to potentially perform a double affinity chromatography.

For PARP15_FL, the optimal expression condition was identified. No purification trials were conducted, as it was assumed that the protocol established for the truncated variant will be also applied to the full-length protein.

Regarding crystallization, promising results are expected, which would allow for a series of considerations regarding PARP15_78-678 and the use of the PARP10/15-IN-3 dual inhibitor on its purification procedure and biochemical analysis.

This study provides a critical foundation in the PARP15 research project since obtaining the recombinant protein can be considered as the very first step of various analyses. Building on the groundwork established by this thesis, functional and structural studies on PARP15 can be pursued, allowing for a clearer understanding of its physiological role at the cellular level and its potential involvement in various pathological contexts.

5. Materials and methods

5.1 Recombinant protein production in *E. coli*

Recombinant proteins are biomolecules produced from their respective cloned genes through recombinant DNA technologies, typically in a heterologous expression system (i.e., the introduction of a gene encoding a protein of interest from one species into the cell of another species).

The process begins with the DNA segment containing the gene coding for the protein of interest; this gene is inserted into a vector (such as plasmids, bacteriophages, artificial chromosomes, etc.) using recombinant DNA technologies. The insertion of the gene into the vector is achieved by first using restriction enzymes that cut at specific target sequences, followed by ligase enzymes that facilitate the insertion of the DNA fragment of interest into the vector. Vectors are then introduced into host cells, where they can self-replicate by exploiting the host's molecular machinery. Recombinant proteins are typically retained in the cytoplasm within inclusion bodies, which in rare cases can be directed to the periplasmic space or the extracellular matrix (Gräslund *et al.,* 2008; Correa and Oppezzo, 2011).

Host cells include yeasts, filamentous fungi, mammalian cells, and bacterial cells, with the latter being the most widely used. *Escherichia coli* is favoured due to its rapid growth, ability to reach high cell densities, minimal nutritional requirements (making it relatively inexpensive), and status as the living organism with the most well-characterized genome and metabolism (Zhu, 2012; Ward, 2012; Nielsen, 2013; Vincentelli *et al.,* 2011; Kazemi Seresht *et al.,* 2013; Çelik and Çalik, 2012).

The applications of recombinant proteins today are vast, ranging from the food industry to gene therapy, genetically modified organisms (GMOs), recombinant antibodies for diagnostic purposes, and even therapeutic uses.

However, during the characterization of these recombinant proteins, several challenges must be addressed. For instance, recombinant proteins are subjected to chemical and physical stresses at every stage of production (Zhao and Chen, 2014), which can lead to the formation of degradation products, resulting in significant heterogeneity in the final product. In some cases, the expression system chosen plays a crucial role, as various proteins require specific

PTMs to be functional (Walsh, 2010; Jenkins *et al.,* 2008). Additionally, for therapeutic proteins intended for human use, there are concerns about immunogenicity if they are produced in expression systems other than mammalian cells (Walsh, 2010; Wurm, 2004).

5.2 Chromatographic techniques for the purification of recombinant proteins

A purification step is required to recover the recombinant protein of interest from the cell lysate. Once the expressing cells are collected, they are lysed in a specific buffer. This process results in the target molecule being mixed with residual cell membranes, cytoplasm, and nuclei, all of which need to be removed through centrifugation that separates the soluble fraction, which is expected to contain the recombinant protein, from the insoluble fraction (or pellet) composed of various cellular debris. Soluble components subsequently undergo purification steps that separate the protein of interest from any remaining contaminants.

Typically, protein purification protocols are well-defined and include at least one chromatographic step. There are various chromatographic techniques available for protein purification, based on different physicochemical principles, some of which will be described below. Chromatographic purification can be carried out using either handmade columns or commercially available precast columns, which are ready to use and can be coupled with automated systems. The associated software provides a chromatogram as the output, a graph that displays peaks corresponding to the elution of biomolecules. The height of the peak directly indicates protein concentration: the higher the peak, the greater the protein concentration (Coskun, 2016; Milewski *et al.,* 2016).

In this work, various chromatographic techniques were employed, which are briefly described below.

Please note that PARP15_78-678 is particularly unstable at room temperature, which is why all purification steps must be carried out at 4°C. Moreover, a high salt concentration in the buffer is required, as it was observed that this recombinant protein has a strong tendency to aggregate.

5.2.1 Ni-NTA affinity chromatography

Ni-NTA affinity chromatography is a metal-chelate affinity chromatography (MCAC). A Ni-NTA resin is characterized by bivalent Ni^{2+} ions covalently attached to NTA molecules. This chromatographic method takes advantage of the 6-histidine tag located at the N-terminal end of PARP15_78-678 and PARP15_FL. The imidazole rings in the histidine residues interact with the Ni^{2+} ions in the resin, forming coordination bonds. This interaction allows the target proteins to remain bound to the resin, enabling the easy removal of contaminating molecules (fig. 31) (Petty *et al.,* 2001). In this work, Ni-NTA His • Bind® Resin (Novagen) was used.

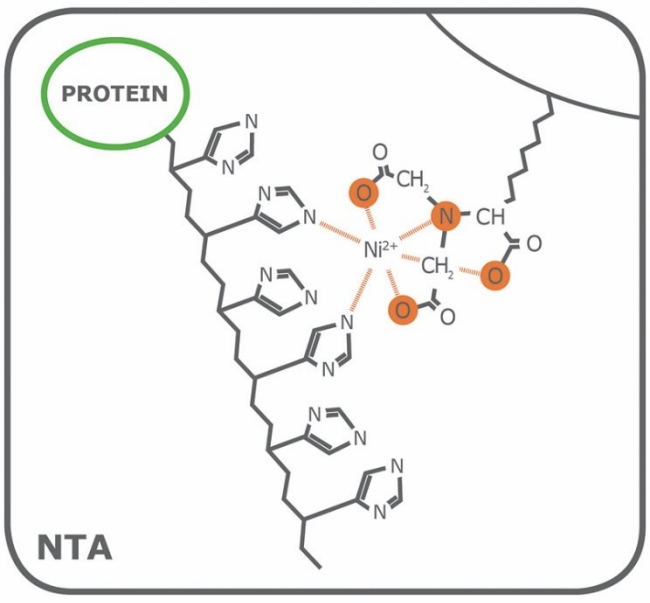

Fig. 31: Ni-NTA resin, interaction between nickel and poly-histidine tag. (Cube Biotech)

5.2.2 Size exclusion chromatography (SEC)

Gel filtration chromatography, also known as size exclusion chromatography (SEC), is a technique used to separate molecules within a mixture based on their different sizes. This method is particularly useful for identifying oligomeric forms of proteins. Moreover, SEC is frequently employed to exchange the buffer in which the sample is contained (Doung-Ly and Gabelli, 2014).

Gel filtration columns are composed of beads that act as molecular sieves. Large molecules, such as proteins and polymers, may or may not enter the spaces between the beads depending on their size, which results in larger entities eluting first, followed by smaller ones (fig. 32).

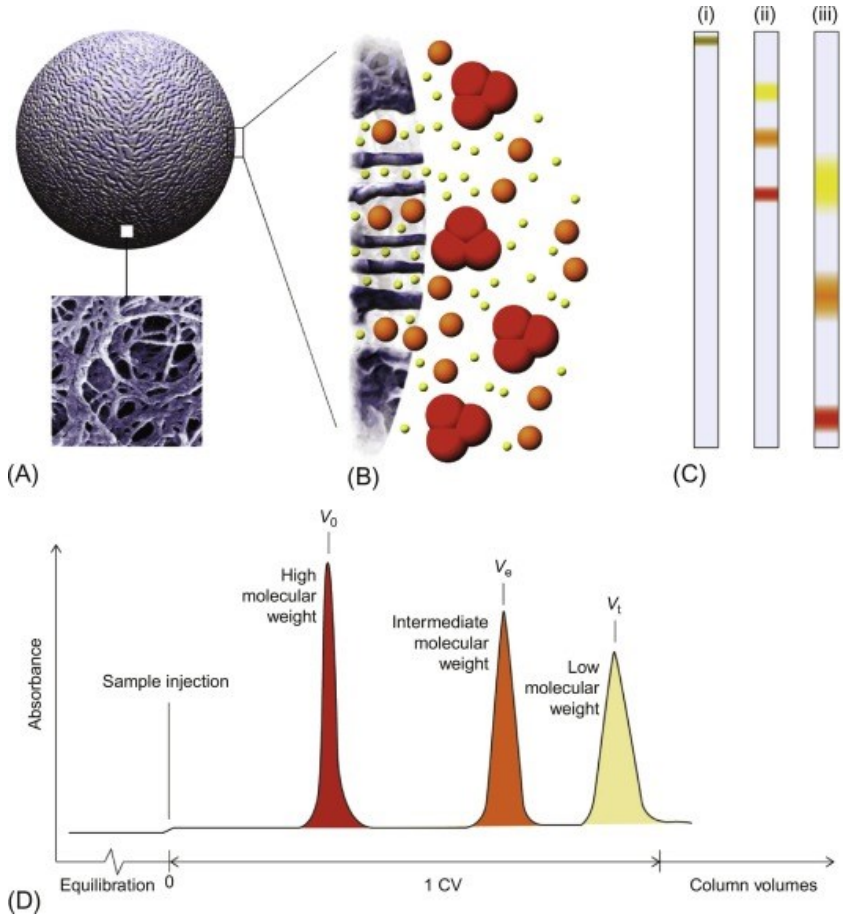

Fig. 32: process of SEC. (A) Schematic picture of a particle. (B) Schematic drawing of sample molecules diffusing into the pores of the particle. (C) Graphical description of separation: (i) sample is applied to the column; (ii) the smallest molecule (yellow) is more delayed than the largest molecule (dark orange); (iii) the largest molecule is eluted first from the column. Band broadening causes significant dilution of the protein zones during chromatography. (D) Schematic chromatogram (size exclusion chromatography handbook, Cytiva).

A Superdex® 200 Increase 10/300 GL column (Cytiva), featuring a bed volume of 24 mL, was used for this process. Additionally, a UV detector set to a wavelength of 280 nm was employed to produce a chromatogram that illustrates the sample's elution profile.

5.2.3 Ion-exchange chromatography (IEX)

Ion-exchange chromatography (IEX) is a chromatographic technique widely used in various fields, proving especially useful for the separation of ions or molecules with different charges, which are pH dependent. In IEX, reversible interactions occur between the ions or charged molecules and the chromatographic support, which has an opposite charge to the analyte of interest. These interactions can be easily modulated to favour either the binding or elution of specific molecules (Cummins *et al.,* 2017).

When it comes to proteins, these interactions may not occur if the pH of the environment matches the isoelectric point (pI) of the protein. At this point, the number of positive charges equals the number of negative charges. Therefore, it is essential to ensure that the biomolecule is maintained in a buffer that preserves its charged form. Specifically, a negatively charged protein will bind to an anion-exchanger, while a positively charged protein will bind to a cation-exchanger (fig. 33). The following general rule applies: a protein will have a net negative charge if the environmental pH is above its pI; conversely, it will have a net positive charge if the pH is below its pI.

Separation occurs because the ions or biomolecules in the sample do not all have the same charge. As a result, they interact with the stationary phase with varying strengths.

Protein elution is achieved by increasing the ionic strength (i.e., the salt concentration) of the mobile phase or by adjusting the pH. Increasing the ionic strength raises the concentration of charged ions (Na^+ and Cl^-), which then compete with the protein for binding to the counterions in the matrix. The higher the net charge of the protein, the greater the ionic strength required for elution (Cummins *et al.,* 2017).

In this work, 5 mL HiTrap Q HP and 1 mL Mono S 5/50 GL columns (Cytiva) were used.

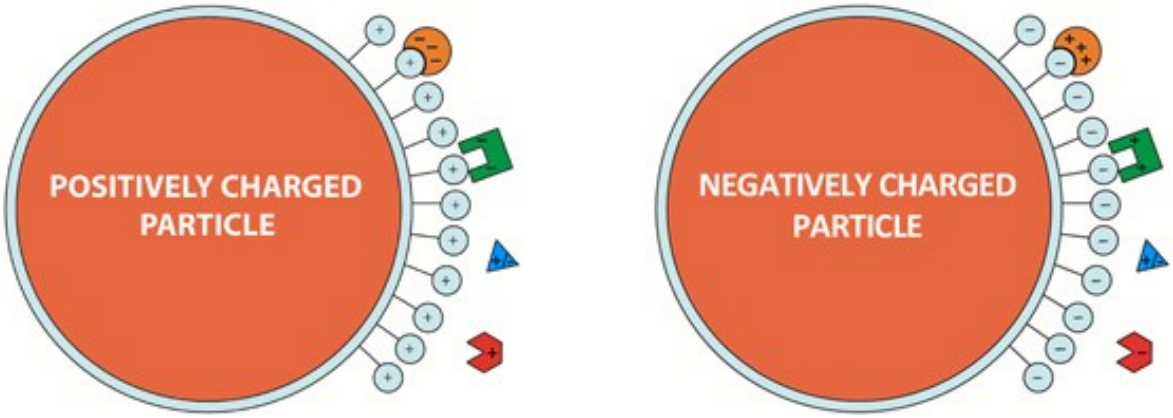

Fig. 33: diagrams of (left) anion-exchange and (right) cation-exchange media particles. (Bio-Rad)

5.2.4 Heparin chromatography

Heparin chromatography is an extremely versatile technique that can be used to purify various types of proteins; it does not require any tags and is compatible with different types of oxidizing, reducing, and chelating agents (Farooqui, 1980; Xiong *et al.,* 2008). Heparin naturally carries a negative charge, which is why it is often considered an analog of DNA. For this reason, heparin chromatography is one of the most effective strategies for purifying DNA-binding proteins (Bolten *et al.,* 2018).

Since heparins carry an electric charge, heparin chromatography is positioned between affinity chromatography and IEX (Xiong *et al.,* 2008).

In this work, a HiTrap Heparin HP affinity column (Cytiva) was used.

5.3 Amplification and isolation of high-purity vectors encoding PARP15_78-678 and PARP15_FL proteins

5.3.1 Expression clones

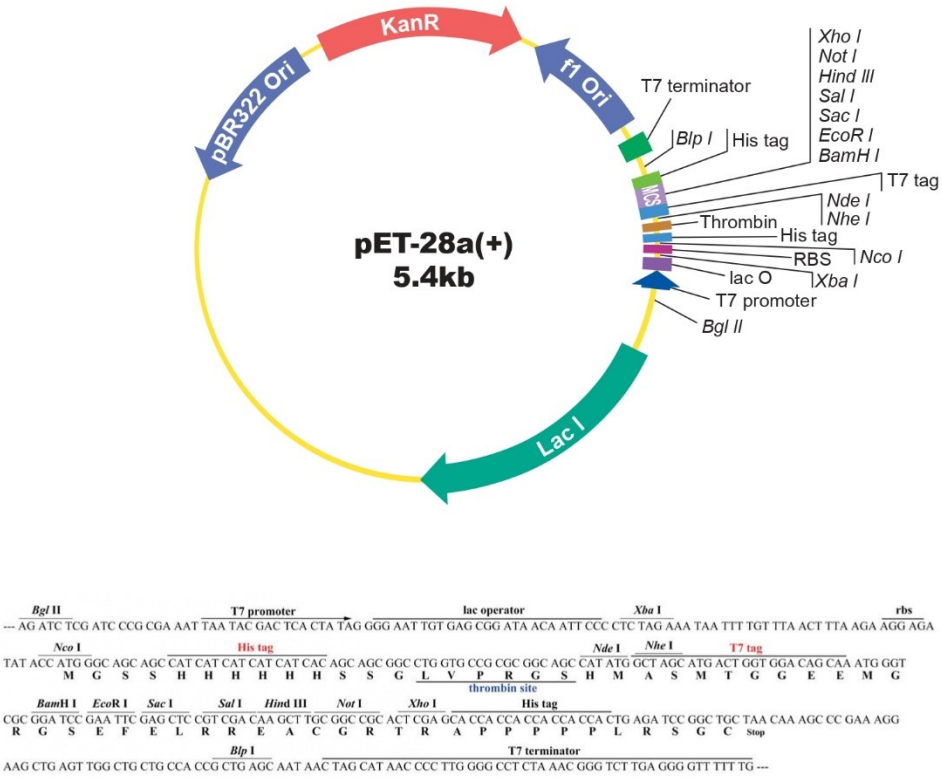

Fig. 34: schematic and linear maps of pET-28a (+) vector. (GenScript Biotech)

The coding sequences for the recombinant PARP15_78-678 and PARP15_FL proteins were synthesized and subcloned into pET-28a (+) vectors by GenScript Biotech, using the Nde I and Xho I restriction sites (fig. 34). Both vectors:

- Include the kanamycin resistance gene, which is crucial for selecting properly transformed colonies.
- Contain the *lac I* gene, which encodes the *lac I* repressor, necessary to inhibit recombinant protein expression until it is triggered by IPTG.

- Feature the T7 promoter, a DNA sequence recognized by T7 RNA polymerase, commonly used to control the expression of recombinant proteins.

Both proteins will include a poly-histidine tag (6X His tag), essential for purification via affinity chromatography, and a SUMO (Small Ubiquitin-like Modifier) tag, a 100-amino-acid sequence (14 kDa) that enhances the expression and solubility of recombinant proteins.

5.3.2 Isolation of high-purity pET-28a (+) plasmids from *E. coli* DH5α cells for subsequent transformations

Plasmid preparation is a DNA extraction and purification technique used for various molecular biology procedures, essential for the effective use of plasmids in biotechnology. In this case, the Fast-n-Easy Mini-Prep Kit by Jena Bioscience was used. The resulting high-quality plasmid DNA is ready for a variety of downstream applications.

To obtain pure pET-28a (+) plasmids, *E. coli* DH5α cells were transformed using the heat shock method (see 5.4.3 paragraph). The day after transformation, a pre-culture was prepared by picking a single colony and placing it in 3 mL of 2X YT liquid medium with 1X kanamycin. The culture was incubated overnight at 37 °C in a shaking incubator to allow cell growth. The next day, the bacterial culture was harvested by centrifugation at 6000 rpm for 10 minutes. The pelleted cells were then resuspended in 300 µL of lysis buffer by pipetting for 1 minute. Following this, 300 µL of neutralization buffer (containing RNase A) were added, and the mixture was gently inverted 4-6 times. The solution was then centrifuged at 10000 rpm for 5 minutes at room temperature to obtain the supernatant; the mixture should turn bright yellow, indicating the required pH of 7.5 for optimal DNA binding. Next, a binding column was placed into a 2 mL collection tube, and 100 µL of activation buffer were added. The column was centrifuged at 10000 rpm for 30 seconds. The supernatant was then applied to the activated binding column, centrifuged at 10000 rpm for 30 seconds, and the flow-through was discarded. Subsequently, 500 µL of washing buffer (containing ethanol) were added to the binding column, followed by centrifugation at 10000 rpm for 30 seconds and discarding the flow-through. A secondary wash was performed by adding 700 µL of

washing buffer to the column, then centrifuging at 10000 rpm for 30 seconds, discarding the flow-through, and centrifuging again for 2 minutes to remove any remaining buffer. Finally, the binding column was placed in a 1.5 mL tube, and 40 µL of elution buffer were added to the centre of the column membrane. After incubating for 1 minute at room temperature, the column was centrifuged at 10000 rpm for 1 minute to elute the DNA. The ultra-pure plasmid DNA was then ready for use.

The plasmid samples concentrations were measured using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific) and diluted to final concentrations of 10 ng/µL.

5.4 Transformation of *E. coli* cells

5.4.1 *E. coli* DH5α competent cells

E. coli DH5α is one of the most widely used *E. coli* strains for plasmid amplification, chosen for its enhanced transformation efficiency. These competent cells are characterized by different key mutations, among which recA1 and endA1, which facilitate plasmid insertion (Wertz J. *The Coli Genetic Stock Center*, CGSC; Strain - DH5α).

5.4.2 *E. coli* BL21 (DE3) competent cells

Competent *E. coli* BL21 (DE3) cells are commonly used for expressing heterologous proteins because they are specifically designed to produce bacteriophage T7 polymerase. This RNA polymerase is faster and more efficient than the standard bacterial RNA polymerase, and these cells have fewer proteases that might degrade the recombinant proteins. T7 polymerase is essential for expressing proteins using systems like pET plasmids, which include a T7 promoter that directs the production of the recombinant protein. The expression of T7 polymerase and the heterologous proteins is regulated by the *lac* promoter, which is typically controlled by the *lac* repressor. The repressor can be removed using IPTG or lactose, which

are non-hydrolysable analogues of allolactose. When IPTG or lactose binds to the *lac* repressor, it causes the repressor to release, thereby initiating the expression of both T7 polymerase and the recombinant protein.

5.4.3 Transformation of *E. coli* cells via heat-shock protocol

A 200 µL aliquot of competent *E. coli* cells, stored at -80 °C, was slowly thawed on ice. Then, 1 µL of the plasmid solution (10 ng/µL, encoding either PARP15_78-678 or PARP15_FL) was added and incubated on ice for 15 minutes. The cells were then subjected to a heat shock by placing the sample in a 42 °C water bath for 90 seconds, creating pores in the bacterial membrane to facilitate plasmid entry. Afterwards, the cells were returned to ice for 2 minutes. Next, 800 µL of sterile 2X YT medium were added, and the cells were incubated in a 37 °C water bath for 45 minutes to allow recovery and proliferation. Following this incubation, 200 µL of the transformed *E. coli* cells were spread onto a Petri dish containing solid 2X YT agar medium (15 g/L) supplemented with kanamycin (50 µg/L) to prevent contamination and select for plasmid-containing cells. The plate was incubated overnight at 37 °C to promote optimal growth.

5.5 Growth of transformed *E. coli* BL21 (DE3) bacterial cells

5.5.1 2X YT medium and induction of recombinant proteins expression

2X YT is a nutrient-rich culture medium used for growing bacteria and expressing recombinant proteins (Sagen, 2023). The medium is prepared by combining:

- 16 g/L of tryptone (a casein digest that provides essential amino acids and is particularly rich in tryptophan, crucial for bacterial growth).
- 10 g/L of yeast extract (a primary nitrogen source that also supplies vitamins, carbohydrates, and nucleotides).
- 5 g/L of NaCl (which helps maintain the medium's water-salt balance).

- Demineralized water to achieve the final volume.

After dissolving all components, the medium was autoclaved at 121°C for 20 minutes. Then, antibiotic (1X kanamycin) was added.

To induce the expression of the recombinant proteins of interest, IPTG was used. This compound is a lactose analogue that functions like it by initiating the transcription of the *lac* operon (fig. 35).

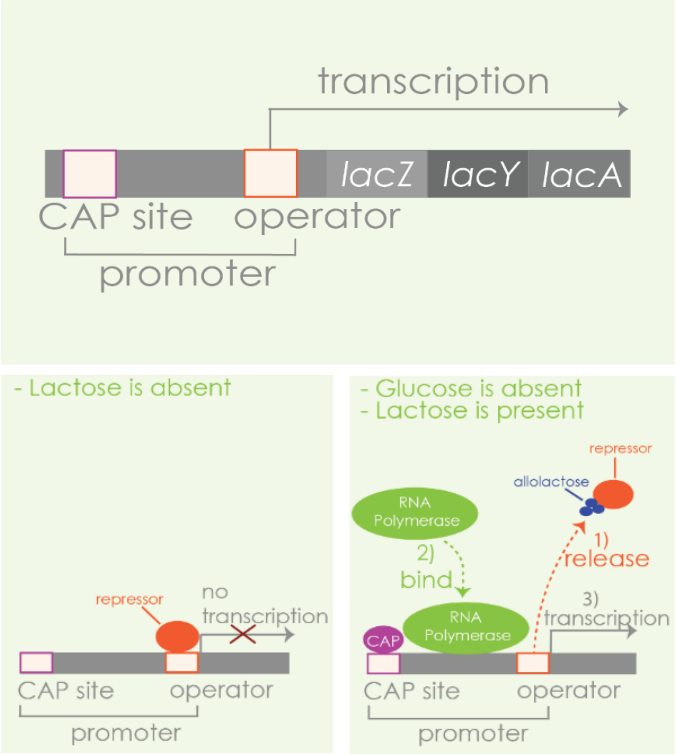

Fig. 35: structure of the *lac* operon and mechanism of action with lactose/IPTG. (Gold Biotechnology®)

5.5.2 Autoinduction medium for recombinant proteins expression

In autoinduction expression medium, IPTG is not required. Instead, the medium contains glucose and an induction sugar (lactose). Initially, the cells use glucose to grow. Once the glucose is depleted, the cells switch to using lactose, which triggers the expression of the target protein (Studier, 2005). This method boosts biomass production and eliminates the

need to monitor culture OD. The autoinduction medium is selective for antibiotics (1X kanamycin) and consists of:

- ZY medium: 1% tryptone (v/w) and 0.5% yeast extract (v/w);
- 50x 5052 buffer: 25% glycerol (v/w), 2.5% glucose (v/w), and 10% α-lactose (v/w), which must be filtered.
- 1 M $MgSO_4$ • 7 H_2O, which must be autoclaved.
- 20x NPS: 1.25 M Na_2HPO_4, 1.25 M KH_2PO_4, 2.5 M NH_4Cl, and 0.25 M Na_2SO_4, which should be filtered and autoclaved.

5.6 Screening of PARP15_78-678 and PARP15_FL expression conditions

To determine the optimal expression conditions for the recombinant proteins, test expressions were initiated the morning after the transformation, once bacterial colonies on the plate were confirmed. Using a sterile L-shaped spatula and sterile 2X YT medium, all colonies were collected and transferred into a 2 mL tube. Next, six 50 mL tubes were prepared, each containing 10 mL of sterile 2X YT medium with 1X kanamycin. Into each tube, 200 µL of the transformed cells were added. After measuring the initial OD at 600 nm, the cultures were incubated at 200 rpm under different conditions, initiated when the OD at 600 nm reached 0.8 at 37 °C:

1. Add IPTG at 0.3 mM and maintain the culture at 37°C for 4 hours.
2. Add IPTG at 0.5 mM and maintain the culture at 37°C for 4 hours.
3. Add IPTG at 0.7 mM and maintain the culture at 37°C for 4 hours.
4. Transfer the culture to 25°C for 1 hour, then add IPTG at 0.3 mM and shift the temperature to 16°C, keeping it overnight.
5. Transfer the culture to 25°C for 1 hour, then add IPTG at 0.5 mM and shift the temperature to 16°C, keeping it overnight.
6. Transfer the culture to 25°C for 1 hour, then add IPTG at 0.7 mM and shift the temperature to 16°C, keeping it overnight.

The OD at 600 nm of the different cultures was measured using a pre-calibrated spectrophotometer. The spectrophotometer was calibrated beforehand by measuring solutions of standard BSA (bovine serum albumin) with known concentrations. After these treatments, the cells were pelleted by centrifugation at 6000 rpm for 10 minutes at 4°C and stored at -20°C until processing.

The procedure followed with the autoinduction medium was the same, with the difference that in this case, the following expression conditions were used after measuring the initial OD:

1. Incubate the culture at 37°C for 3 hours.
2. Incubate the culture at 37°C for 4 hours.
3. Incubate the culture at 37°C for 6 hours.
4. Incubate the culture at 37°C for 3 hours, then shift to 16°C and continue growth for 1 hour.
5. Incubate the culture at 37°C for 3 hours, then shift to 16°C and continue growth for 3 hours.
6. Incubate the culture at 37°C for 3 hours, then shift to 16°C and continue growth overnight.

After these incubation periods, the cells were collected by centrifugation at 6000 rpm for 10 minutes at 4°C and stored at -20°C until further processing.

5.7 Test expression procedure

To identify the best expression conditions for PARP15_78-678 and PARP15_FL, all frozen pellets were slowly thawed on ice and resuspended in 1.5 mL lysis buffer with the following characteristics:

Lysis buffer PARP15_78-678	
Tris-HCl pH 8	50 mM
NaCl	500 mM
Glycerol	10%
Tryton x100	1%
TCEP	0.5 mM
Imidazole	10 mM

Lysis buffer PARP15_FL	
Tris-HCl pH 7	50 mM
NaCl	500 mM
Glycerol	10%
Tryton x100	1%
TCEP	0.5 mM
Imidazole	10 mM

Tris-HCl (tris-(hydroxymethyl)-aminomethane hydrochloride) is a commonly used buffer designed to maintain a stable pH. For this work, pH values of 8.5 and 7 were selected based on the isoelectric points (pI) of the proteins produced. Ionic salts, such as NaCl, help regulate the osmolarity of the lysate and simulate physiological conditions. Glycerol preserves the structure of soluble proteins in solution and maintains enzyme activity. Tryton X-100, a non-ionic detergent, enhances protein solubility without denaturation. TCEP, a reducing agent, was used to reduce disulfide bonds in proteins. Imidazole was added to the lysis buffer for subsequent purification steps using Ni-NTA resin.

Cell lysis was performed using a Vibra-Cell VC130 plate sonicator (Sonics®) at 4% amplitude. To prevent temperature rise during sonication, samples were kept on ice. The lysis process involved 10 cycles of 5 seconds of sonication followed by 30 seconds off, with the plate remaining on ice throughout.

For each lysate, 20 µL were collected to obtain the total (T) fractions. To separate soluble (S) and pellet (P) fractions, 50 µL of each sample were centrifuged at 13000 rpm for 10 minutes at 4°C. The supernatants, representing the soluble fractions, were transferred to separate tubes, while the cellular debris, representing the pellet fractions, were resuspended in 50 µL of lysis buffer. Additionally, 1 mL of each lysate was centrifuged at 13000 rpm for 10 minutes

at 4°C, and the soluble fractions were incubated with 50 µL of Ni-NTA resin for 60 minutes. The Ni-NTA resin, stored in 20% ethanol, washed with ultrapure water and equilibrated with lysis buffer, was used to capture the recombinant proteins. The resin contains immobilized nickel ions (Ni^{2+}) that form coordination bonds with the 6 histidines in the N-terminal tag of the recombinant proteins. Afterwards, proteins were eluted from the resin with 100 µL of elution buffer with the following compositions:

Elution buffer PARP15_78-678	
Tris-HCl pH 8	50 mM
NaCl	500 mM
Glycerol	10%
TCEP	0.5 mM
Imidazole	300 mM

Elution buffer PARP15_FL	
Tris-HCl pH 7	50 mM
NaCl	500 mM
Glycerol	10%
TCEP	0.5 mM
Imidazole	300 mM

Imidazole was used to elute PARP15_78-678 and PARP15_FL from the Ni-NTA resin. It competes with the recombinant proteins for the same Ni^{2+} ions, thereby displacing the proteins from the resin. The elution process lasted 15 minutes, with occasional gentle shaking to help release the proteins. After allowing the resin to naturally settle, 20 µL of each eluate were collected to obtain the elution (E) samples. Finally, total, pellet, soluble and elution fractions were loaded on polyacrylamide gels for SDS-PAGE analyses.

5.8 Purification trials of PARP15_78-678

5.8.1 Large-scale cultures for bacterial pellets production

After determining the optimal conditions for the expression of PARP15_78-678, larger-scale cultures were conducted in 5 L flasks. The process followed the established steps from the expression tests: preparing the culture medium, transforming *E. coli* BL21 (DE3), transferring the transformed cells into sterile liquid medium, and allowing them to grow under the optimal conditions. The cultures were then centrifuged at 6000 rpm for 10 minutes at 4 °C using a Beckman Coulter Avanti Centrifuge J-26 XP with rotor JA-14 to collect the cell pellets, which were stored at -20 °C.

5.8.2 Cell lysis and clarification

The cell pellets were thawed on ice and then resuspended in lysis buffer at a ratio of 1:8 (g pellets/mL buffer), which have the following composition:

Lysis buffer PARP15_78-678	
Tris-HCl pH 8 or 8.5	50 mM
NaCl	500 mM
Glycerol	10%
Tryton x100	1%
TCEP	0.5 mM
Imidazole	10 or 30 mM
Benzonase	1 µL/80 mL of lysis buffer
Protease Inhibitors	100 µL/40 mL of lysis buffer

Tris-HCl (tris-(hydroxymethyl)-aminomethane hydrochloride) is a commonly used buffer designed to maintain a stable pH. For this work, pH values of 8 or 8.5 were selected based on the isoelectric point (pI) of the protein produced. Ionic salts, such as NaCl, help regulate

the osmolarity of the lysate and simulate physiological conditions. Glycerol preserves the structure of soluble proteins in solution and maintains enzyme activity. Tryton X-100, a non-ionic detergent, enhances protein solubility without denaturation. TCEP, a reducing agent, was used to reduce disulfide bonds in proteins. Imidazole was added to the lysis buffer for subsequent purification steps using Ni-NTA resin. The lysis buffer was prepared with benzonase to break down nucleic acids in the cell lysate, while protease inhibitors were included to protect the recombinant protein of interest from degradation.

At this stage, the cell pellets, resuspended in lysis buffer, were lysed on ice using sonication (Vibra-Cell VC130 plate sonicator, Sonics®). Specifically, the sonication was carried out in 10 cycles at 40% amplitude, with each cycle comprising 30 seconds of sonication followed by 1 minute of rest. After the sonication process, the cell lysates were centrifuged at 15000 rpm for 45 minutes at 4 °C using a Beckman Coulter Avanti Centrifuge J-26 XP with rotor JA-25.50. Cellular debris (insoluble fractions) formed a pellet at the bottom of the tubes and were discarded, while the supernatants (clarified or soluble fractions), which contain the recombinant proteins of interest, were collected.

5.8.3 First purification trial

The starting purification step for PARP15_78-678 was Ni-NTA affinity chromatography. The Ni-NTA resin is typically stored at 4°C as 50% slurry in 20% ethanol. Before use, 1.5 mL of dry resin was washed in a column with ultrapure water and equilibrated with the appropriate lysis buffer. The cellular lysate was then incubated with the resin in beaker under agitation for about 2 hours at 4°C to facilitate the binding of the recombinant protein to the resin. Following incubation, the mixture was transferred back to the column, and the flow-through was collected. The resin was subsequently washed with 10 CV (i.e., 15 mL) of wash buffer to remove any contaminants that weakly interacted with it. The composition of the wash buffer used was as follows:

Wash buffer	
Tris-HCl pH 8	50 mM
NaCl	500 mM
Glycerol	10%
TCEP	0.5 mM
Imidazole	40 mM

The low concentration of imidazole aids in removing any substances that have non-specifically interacted with the bivalent Ni^{2+} ions of the resin.

Finally, the recombinant protein of interest, which exhibited strong interactions with the resin, was eluted using the corresponding elution buffer, which had the following characteristics:

Elution buffer	
Tris-HCl pH 8	50 mM
NaCl	500 mM
Glycerol	10%
TCEP	0.5 mM
Imidazole	300 mM

The elution buffer is a solution similar in composition to the respective lysis and wash buffers, but with an imidazole concentration of 300 mM. This high imidazole concentration competes with the imidazole rings of the histidine residues in the His-tag for binding to the same Ni^{2+} ions, thereby promoting the release (or elution) of PARP15_78-678. Six elution fractions, each 1.5 mL in volume, were collected in 5 mL tubes. 10 µL aliquots from each fraction were prepared for SDS-PAGE analysis by loading onto polyacrylamide gels.

The six elution fractions from the Ni-NTA chromatography were dosed using the Bradford assay. After quantification, the six elution fractions were combined to achieve a total volume of 9 mL. To proceed with the subsequent chromatographic steps, it was necessary to concentrate the pooled Ni-NTA fractions. The concentration process was carried out using Vivaspin® 20 ultrafiltration units (Sartorius) with a 50 kDa MWCO, suitable for the

molecular weight of the recombinant protein, through multiple centrifuge cycles at 6000 rpm for 6 to 8 minutes, until a final volume of 1 mL was obtained. After the concentration step, a Bradford assay was repeated to verify that the sample had been successfully concentrated.

After Ni-NTA and concentration, a size exclusion chromatography was performed. The Superdex® 200 Increase 10/300 GL column (Cytiva) was integrated with an ÄKTA™ explorer purification system, which enables the monitoring of various parameters, such as flow rate and pressure. The column was previously conditioned with the following degassed running buffer, called SEC buffer.

SEC buffer	
Tris-HCl pH 8	50 mM
NaCl	500 mM
Glycerol	10%
TCEP	0.5 mM

The concentrated sample containing PARP15_78-678 was finally loaded onto the equilibrated column. The elution fractions were collected in 2 mL tubes placed in an automated rotor, and 10 µL of each fraction corresponding to the peak of the protein of interest were recovered and loaded onto a polyacrylamide gel for SDS-PAGE analysis. The elution fractions from the SEC chromatography of PARP15_78-678 were finally dosed using the Bradford assay.

5.8.4 Second purification trial

In the second attempt to purify PARP15_78-678, the initial steps followed were identical to those used in the first trial, including Ni-NTA affinity chromatography and pooling the elution fractions. For Ni-NTA affinity chromatography step, different buffers were used compared to the first purification test:

Wash buffer	
Tris-HCl pH 8.5	50 mM
NaCl	500 mM
Glycerol	10%
TCEP	0.5 mM
Imidazole	70 mM

Elution buffer	
Tris-HCl pH 8.5	50 mM
NaCl	500 mM
Glycerol	10%
TCEP	0.5 mM
Imidazole	300 mM

High concentrations of imidazole could lead to protein degradation. This issue is usually addressed by performing a buffer exchange using a SEC column. However, since the SEC step was critical for this recombinant protein, a buffer exchange was carried out using the HiPrep 26/10 Desalting column (Cytiva).

The column was integrated with an ÄKTA™ explorer purification system, which enables the monitoring of various parameters, such as flow rate and pressure. The column was previously conditioned with 5 CV of the following degassed running buffer, called desalting buffer.

Desalting buffer	
Tris-HCl pH 8.5	50 mM
NaCl	50 mM
Glycerol	10%
TCEP	0.5 mM

The NaCl concentration in this buffer was deliberately kept low, as it is crucial for the next step, which involves anion-exchange chromatography. The sample containing PARP15_78-678 with 300 mM imidazole was added into the equilibrated column; 2 CV of desalting buffer at 15 mL/min were flowed through the column, the elution fractions were collected in 2 mL tubes placed in an automated rotor and those of interest were finally pooled. This process yielded approximately 11 mL of sample containing PARP15_78-678 with no imidazole and low concentration of NaCl.

Finally, the sample obtained was injected into a 5 mL HiTrap Q HP column (Cytiva), previously integrated with an ÄKTA™ explorer purification system, which enables the monitoring of various parameters, such as flow rate and pressure, washed with 5 CV of MilliQ water and equilibrated with 5 CV of start buffer (i.e., desalting buffer). In this chromatographic technique, the pH of the buffer is critical because it greatly influences the protein's charge. PARP15_78-678 has an isoelectric point (pI) of 7.06; a pH of 8.5 was selected to ensure that the protein remains in its negatively charged state. Elution was performed using a gradient strategy, gradually increasing the NaCl concentration (from 50 mM to 2 M) over 10 CV with HiTrap Q buffer with the following specifications:

HiTrap Q buffer	
Tris-HCl pH 8.5	50 mM
NaCl	2 M
Glycerol	10%
TCEP	0.5 mM

The elution fractions were collected in 2 mL tubes placed in an automated rotor, 10 µL of each fraction corresponding to the peak of the protein of interest were recovered and loaded onto a polyacrylamide gel for SDS-PAGE analysis and dosed with the Bradford assay. Then, elution fractions of interest were pooled and concentrated until a finale volume of 0.5 mL

After completing all these steps, protein samples were stored at -80 °C until further use.

5.8.5 Third purification trial

In this purification attempt, the same steps from the second test were followed, except for the final anion exchange chromatography step: cation exchange chromatography was performed using the Mono S 5/50 GL column, with the goal of improving the final purity compared to the results obtained in the second trial. Another important difference is that to successfully carry out cation exchange chromatography the protein must retain a net positive charge; for

this reason, the pH of buffers was lowered to 6, compared to the value of 8.5 used in the second test.

5.8.6 Fourth purification trial

The same procedures as the previous two trials were followed, with the difference that, instead of performing an IEX, a run was carried out using a 5 mL HiTrap Heparin HP column (Cytiva) In this trial, buffers at pH 6 were used to ensure the protein had a net positive charge, which is essential to promote binding to heparin, which is negatively charged.

Buffers used are the same as those employed in the third purification trial, with the only difference being that the heparin buffer used had a NaCl concentration of 500 mM.

5.8.7 Fifth purification trial

Finally, a single Ni-NTA affinity chromatography step was carried out. In this case, after collecting the flow-through, wash steps of the resin were performed using various wash buffers with increasing concentrations of imidazole. The aim was to implement a step gradient elution to maximize the achievable level of purity, allowing the target protein to be eluted. Specifically, the imidazole concentrations used were: 70 mM, 90 mM, 100 mM, 110 mM, 130 mM, 150 mM, 200 mM, and 250 mM. Finally, the various elution fractions were collected and loaded onto a gel for SDS-PAGE analysis. Subsequently, the eluate fractions of interest were pooled, dialyzed to remove imidazole, and quantified using a Bradford assay. Finally, 4 µg of protein of dialyzed sample were loaded into a gel for SDS-PAGE analysis

5.9 Western blot for protein detection

Western blot is a biochemical technique used to detect a protein of interest, perform subsequent quantitative analysis, and determine its molecular weight from a heterogeneous mixture of proteins extracted from cells or tissues (Taylor and Posch, 2014).

In this case, a primary antibody capable of recognizing the His-tag was used, along with a labeled secondary antibody that binds to the primary, and a mixture of hydrogen peroxide and luminol was used as the substrate. To ensure the success of the experiment, another protein containing the His-tag was employed as control.

After the electrophoretic run, the transfer sandwich was assembled using a nitrocellulose membrane, ensuring the complete absence of air bubbles and that the polarity of the electrodes was correct to facilitate transfer onto the membrane. Then, the sandwich was inserted into the appropriate tank containing an ice block to keep the temperature low. Transfer was performed at a constant current of 250 mA for 2 hours. Then, the membrane was stained with Ponceau S staining solution (Ponceau S 0.1% w/v + 5% acetic acid in milli Q H_2O) for one minute and immediately washed with ultrapure water to remove the excess stain. Subsequently, the membrane was incubated in a milk solution (5% w/v milk in TBS 1X) for one hour to block nonspecific antibody binding sites. Excess milk solution was removed, and the membrane was incubated with the primary antibody overnight at 4°C. To remove any remaining excess, 5 minutes washes were performed with TBST 1X (0.3% v/v Tween 20 in TBS 1X), and one-hour incubation with the secondary antibody took place at room temperature. Any remaining excess was removed by washing 5 times for 5 minutes with TBST 1X. Substrate was applied across the entire membrane (H_2O_2 and luminol 1:1) and chemiluminescence reaction was finally captured by Chemidoc imaging system (Bio-Rad).

This procedure was performed exclusively after the fourth purification trial to confirm the presence of PARP15_78-678.

5.10 SDS-PAGE analysis

SDS-PAGE is a widely used analytical technique for separating proteins or protein fragments based on their size.

The principle of SDS-PAGE is straightforward: large proteins encounter more resistance as they move through the polyacrylamide gel, while smaller proteins and peptides pass through the gel more easily and travel further. This size-dependent migration results in protein separation. In SDS-PAGE, proteins are denatured and coated with a uniform negative charge by sodium dodecyl sulfate (SDS), which ensures that all proteins have similar shapes and charges. The negatively charged proteins migrate towards the positive pole, with their speed dependent on their size, leading to their separation into distinct bands.

SDS-PAGE uses a discontinuous gel system: the stacking gel has a low acrylamide concentration (4%) and aligns the proteins before they enter the lower resolving gel, which has a higher acrylamide concentration tailored to the sample's needs.

In SDS-PAGE, the samples were prepared by mixing them with 1X reducing Laemmli buffer. This buffer consists of 375 mM Tris-HCl, 9% SDS, 50% glycerol, 9% β-mercaptoethanol and 0.03% bromophenol blue. Bromophenol blue is a tracking dye that carries a negative charge and migrates quickly in the electric field, allowing the progress of the electrophoresis to be monitored.

During the loading phase, one well was consistently loaded with BlueClassic Prestained Protein Marker (Jena Bioscience), a mixture of blue-coloured proteins with known molecular weights, which serves as a reference.

Runs were performed at a constant voltage of 180 V. After the electrophoresis was complete, the separated proteins were visualized using a gel protein detection method. For this, InstantBlue™ Coomassie Protein Stain (Generon) was employed. Following the staining, the gels were treated with a destaining solution (25% acetic acid, 15% MeOH, 60% H_2O) to remove excess stain. Bands associated with PARP15_78-678 and PARP15_FL were observed in the gels at roughly the positions of the 80 kDa and 88 kDa marker bands, respectively.

5.11 Bradford's method for protein quantification

The Bradford assay is a method for quantifying proteins in a sample, utilizing the dye Coomassie Brilliant Blue R-250 (fig. 36). Initially, the dye has an absorbance peak at 465 nm, but when it reacts with the basic amino acids, this peak shifts to 595 nm, resulting in a colour change from brown to blue (Bradford, 1976). For the quantification of the recombinant proteins, they were mixed with Coomassie Brilliant Blue R-250 at a 1:1000 ratio, and their concentrations were determined using a pre-calibrated spectrophotometer, with readings at 595 nm.

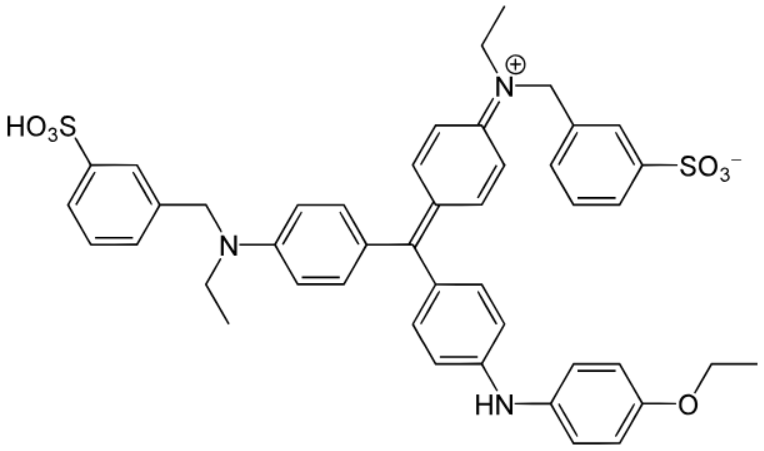

Fig. 36: Coomassie Brilliant Blue R-250 dye structure formula.

5.12 Crystallization procedure for PARP15 Catalytic Domain

To study the 3D structure of the protein in question, crystallization trial was conducted. We decided to start working on a high-purity sample of catalytic domain (CD) of PARP15, already available in the laboratory at a concentration of 1.5 mg/mL. This recombinant form had been produced via expression in *E. coli* BL21 (DE3) competent cells and purified using an optimized protocol during which an on-resin cleavage of the protein tag was performed; pooled eluted fractions were concentrated to a final volume of 2 mL. PARP15_CD was available in the following buffer, called CD buffer.

CD buffer	
Tris-HCl pH 7.5	50 mM
NaCl	150 mM
Glycerol	10%
DTT	2 mM

The CD of PARP15 was subjected to crystallization in both a ligand-free form and co-crystallized with its inhibitor, named PARP10/15-IN-3 (purchased from MedChemExpress as a powder, it was resuspended in 100% DMSO to reach a final concentration of 50 mM), whose structure is shown in fig. 37. This compound had been considered a potential PARP10 inhibitor (IC50 of 0.14 μM). However, as described in the work by Nizi *et al.,* 2022, subsequent testing on other members of the PARP superfamily revealed significant inhibitory activity against PARP15 as well (IC50 of 0.40 μM), identifying IN-3 as a dual inhibitor of PARP10 and PARP15.

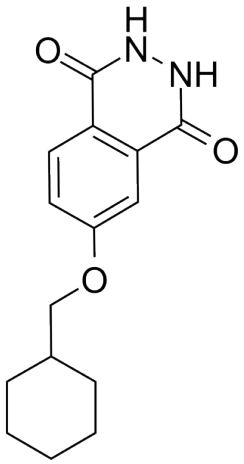

Fig. 37: structure of PARP10/15-IN-3 dual inhibitor.

Regarding the crystallization trial itself, the optimal crystallization conditions for PARP15_CD are already known from the literature (Karlberg *et al.,* 2015; Zhou *et al.,* 2022; Yoneyama-Hirozane *et al.,* 2017):

- PARP15_CD concentration: 10 mg/mL.
- Precipitant (reservoir) solution: 0.2 M NH_4Cl and 20% PEG-3350.

Therefore, for the crystallization trial conducted, the PARP15_CD sample was concentrated to 10 mg/mL as final concentration, a small volume of it was incubated with 1 mM IN-3 (2% DMSO final) at 4°C for 2 hours, and a series of conditions were tested in a VDX plate (Crystalgen), staying within the range of those already described in the literature, as detailed in table 7. The hanging drop method was used in this case.

% PEG-3350					
NH_4Cl concentration (mM)					
16	18	20	22	24	26
180	180	180	180	180	180
16	18	20	22	24	26
190	190	190	190	190	190
16	18	20	22	24	26
200	200	200	200	200	200
16	18	20	22	24	26
210	210	210	210	210	210

Table 7: reservoir compositions for every well of the plate.

After preparing the reservoir solutions in the wells of the plate, two 1 μL drops for each well, corresponding to the ligand-free and protein-ligand forms, were deposited onto 22 mm diameter siliconized Circular Cover Slides (Jena Bioscience) and mixed with an equal volume of the corresponding precipitant solution to allow vapor diffusion and subsequent crystal formation. The plate was finally incubated at 20°C until crystals were observed.

Bibliography

1. Aljabban, J., Syed, S., Syed, S., Rohr, M., Mukhtar, M., Aljabban, H., Cottini, F., Mohammed, M., Hughes, T., Gonzalez, T., Panahiazr, M., Hadley, D., Benson, D. (2023) Characterization of monoclonal gammopathy of undetermined significance progression to multiple myeloma through meta-analysis of GEO data. *Heliyon* **9**, e17298
2. AlphaFold protein structure database: AlphaFold.ebi.ac.uk/
3. Amè, J. C., Pennebaker, C., de Murcia, G. (2004) The PARP superfamily. *BioEssays* **26**, 882-893
4. Aravind, L. (2001) The WWE domain: a common interaction module in protein ubiquitination and ADP ribosylation. *Trends in biochemical sciences* **26**, 273-275
5. Astigiano, C., Benzi, A., Laugieri, M. E., Piacente, F., Sturla, L., Guida, L., Bruzzone, S., De Flora, A. (2022) Paracrine ADP ribosyl cyclase-mediated regulation of biological processes. *Cells* **11**, 2637
6. Beckman Coulter: Avanti J-26S XP, High-performance centrifuge
7. Bellamacina, C. R. (1996) The nicotinamide dinucleotide binding motif: a comparison of nucleotide binding proteins. *FASEB journal* **10**, 1257-1269
8. Best, C. H. (1959) Preparation of heparin and its use in the first clinical cases. *Circulation* **19**, 79-86
9. Bio-Rad: Chemidoc imaging systems
10. Bio Rad: Ion-exchange chromatography
11. Bogan, K. L., and Brenner, C. (2008) Nicotinic acid, nicotinamide, and nicotinamide riboside: a molecular evaluation of NAD^+ precursor vitamins in human nutrition. *Annual review of nutrition* **28**, 115-130
12. Bolten, S. N., Rinas, U., Scheper, T. (2018) Heparin: role in protein purification and substitution with animal-component free material. *Applied microbiology and biotechnology* **102**, 8647-8660
13. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254
14. Caframo Lab Solutions: Protein affinity chromatography
15. Cancer research UK: Types of cancer immunotherapy

16. Cantò, C., Houtkooper, R. H., Pirinen, E., Youn, D. Y., Oosterveer, M. H., Cen, Y., Fernandez-Marcos, P. J., Yamamoto, H., Andreux, P. A., Cettour-Rose, P., Gademann, K., Rinsch, C., Schoonjans, K., Sauve, A. A., Auwerx, J. (2012) The NAD^+ precursor nicotinamide riboside enhances oxidative metabolism and protects against high-fat diet-induced obesity. *Cell metabolism* **15**, 838-847
17. Çelik, E., and Çalık, P. (2012) Production of recombinant proteins by yeast cells. *Biotechnology advances* **30**, 1108-1118
18. Correa, A., and Oppezzo, P. (2011) Tuning different expression parameters to achieve soluble recombinant proteins in *E. coli*: advantages of high-throughput screening. *Biotechnology journal* **6**, 715-730
19. Coskun, O. (2016) Separation techniques: chromatography. *Northern clinics of Istanbul* **3**, 156-160
20. Crystalgen: 24-well plate, super clear
21. Cube Biotech: NTA Versus IDA: what's the difference?
22. Cytiva: Fundamentals of size exclusion chromatography (SEC) handbook
23. Cytiva: HiPrep 26/10 Desalting
24. Cytiva: HiTrap Heparin HP affinity columns
25. Cytiva: HiTrap Q HP anion-exchange chromatography column
26. Cytiva: Mono S 5/50 GL cation-exchange chromatography column
27. Cytiva: Size exclusion chromatography: principles and methods
28. Cytiva: Superdex® 200 Increase 10/300 GL column
29. Cummins, P. M., Rochfort, K. D., O'Connor, B. F. (2016) Ion-exchange chromatography: basic principles and application. *Methods in molecular biology* **1485**, 209-223
30. Dawicki-McKenna, J. M., Langelier, M. F., DeNizio, J. E., Riccio, A. A., Cao, C. D., Karch, K. R., McCauley, M., Steffen, J. D., Black, B. E., Pascal, J. M. (2015) PARP-1 activation requires local unfolding of an autoinhibitory domain. *Molecular cell* 60, 755-768
31. De Vos, M., Schreiber, V., Dantzer, F. (2012) The diverse roles and clinical relevance of PARPs in DNA damage repair: current state of the art. *Biochemical pharmacology* **84**, 137-146

32. Domenighini, M., and Rappuoli, R. (1996) Three conserved consensus sequences identify the NAD-binding site of ADP-ribosylating enzymes, expressed by eukaryotes, bacteria and T-even bacteriophages. *Molecular microbiology* 21, 667-674
33. Duarte-Pereira, S., Matos, S., Oliveira, J. L., Silva, R. M. (2023) Study of NAD-interacting proteins highlights the extent of NAD regulatory roles in the cell and its potential as a therapeutic target. *Journal of integrative bioinformatics* **20**, 20220049
34. Duong-Ly, K. C., Gabelli, S. B. (2014) Gel filtration chromatography (size exclusion chromatography) of proteins. *Methods in enzimology* **541**, 105-114
35. Ekblad, T., Verheugd, P., Lindgren, A. E., Nyman, T., Elofsson, M., Schüler, H. (2018) Identification of Poly (ADP-Ribose) Polymerase macrodomain inhibitors using an AlphaScreen protocol. *SLAS discovery* **23**, 353-362
36. Farmer, H., McCabe, N., Lord, C. J., Tutt, A. N., Johnson, D. A., Richardson, T. B., Santarosa, M., Dillon, K. J., Hickson, I., Knights, C., Martin, N. M., Jackson, S. P., Smith, G. C., Ashworth, A. (2005) Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917-921
37. Farooqui, A. A. (1980) Purification of enzymes by heparin-sepharose affinity chromatography. *Journal of chromatography A* **184**, 335-345
38. Garantziotis, S., Hollingsworth, J. W., Ghanayem, R. B., Timberlake, S., Zhuo, L., Kimata, K., Schwartz, D. A. (2007) Inter-alpha-trypsin inhibitor attenuates complement activation and complement-induced lung injury. *The journal of immunology* **179**, 4187-4192
39. GE Healthcare: ÄKTA™ explorer
40. GenScript Biotech: pET-28a (+) plasmid
41. Gibson, B. A., Kraus, W. L. (2012) New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. *Nature reviews molecular cell biology* **13**, 411-424
42. Gold Biotechnology®: How does IPTG induction work?
43. Gräslund, S., Nordlund, P., Weigelt, J., Hallberg, B. M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C., Hui, R., Ming, J., dhe-Paganon, S., Park, H. W., Savchenko, A., Yee, A., Edwards, A., Vincentelli, R., Cambillau, C., Kim, R., Kim, S. H., Rao, Z., Shi, Y., Terwilliger, T. C., Kim, C. Y., Hung, L. W., Waldo, G. S., Peleg, Y., Albeck, S., Unger, T., Dym, O., Prilusky, J., Sussman, J. L., Stevens, R. C., Lesley, S. A.,

Wilson, I. A., Joachimiak, A., Collart, F., Dementieva, I., Donnelly, M. I., Eschenfeldt, W. H., Kim, Y., Stols, L., Wu, R., Zhou, M., Burley, S. K., Emtage, J. S., Sauder, J. M., Thompson, D., Bain, K., Luz, J., Gheyi, T., Zhang, F., Atwell, S., Almo, S. C., Bonanno, J. B., Fiser, A., Swaminathan, S., Studier, F. W., Chance, M. R., Sali, A., Acton, T. B., Xiao, R., Zhao, L., Ma, L. C., Hunt, J. F., Tong, L., Cunningham, K., Inouye, M., Anderson, S., Janjua, H., Shastry, R., Ho, C. K., Wang, D., Wang, H., Jiang, M., Montelione, G. T., Stuart, D. I., Owens, R. J., Daenke, S., Schütz, A., Heinemann, U., Yokoyama, S., Büssow, K., Gunsalus, K. C. (2008) Protein production and purification. *Nature methods* **5**, 135-146

44. Guettler, S., LaRose, J., Petsalaki, E., Gish, G., Scotter, A., Pawson, T., Rottapel, R., Sicheri, F. (2011) Structural basis and sequence rules for substrate recognition by Tankyrase explain the basis for cherubism disease. *Cell* **147**, 1340-1354
45. Gupte, R., Liu, Z. Kraus, W. L. (2017) PARPs and ADP-ribosylation: recent advances linking molecular functions to biological outcomes. *Genes & Development* **31**, 101-126
46. He, F., Tsuda, K., Takahashi, M., Kuwasako, K., Terada, T., Shirouzu, M., Watanabe, S., Kigawa, T., Kobayashi, N., Güntert, P., Yokoyama, S., Muto, Y. (2012) Structural insight into the interaction of ADP-ribose with the PARP WWE domains. *FEBS letters* **586**, 3858-3864
47. Hileman, R. E., Fromm, J. R., Weiler, J. M., Linhardt, R. J. (1998) Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. *Bioassays* **20** 156-167
48. Imai, S., Guarente, L. (2014) NAD^+ and sirtuins in aging and disease. *Trends in cell biology* **24**, 464-471
49. Jena Bioscience: 22 mm circular cover slides
50. Jena Bioscience: BlueClassic Prestained Protein MW Marker (10-180kDa)
51. Jena Bioscience: Fast-n-Easy Mini-Prep Kit
52. Jenkins, N., Murphy, L., Tyther, R. Post-translational modifications of recombinant proteins: significance for biopharmaceuticals. *Molecular biotechnology* **38**, 113-118
53. Jiang, S., Ma, W., Ma, C., Zhang, Z., Zhang, W., Zhang, J. (2024) An emerging strategy: probiotics enhance the effectiveness of tumor immunotherapy via mediating the gut microbiome. *Gut microbes* **16**, 2341717

54. Karlberg, T., Thorsell, A. G., Kallas, Å., Schüler, H. (2012) Crystal structure of human ADP-ribose transferase ARTD15/PARP16 reveals a novel putative regulatory domain. *Journal of biological chemistry* **287**, 24077-24081
55. Kazemi Seresht, A., Cruz, A. L., de Hulster, E., Hebly, M., Palmqvist, E. A., van Gulik, W., Daran, J. M., Pronk, J., Olsson, L. (2013) Long-term adaptation of *Saccharomyces cerevisiae* to the burden of recombinant insulin production. *Biotechnology and bioengineering* **110**, 2749-2763
56. Kickhoefer, V. A., Siva, A. C., Kedersha, N.L., Inman, E. M., Ruland, C., Streuli, M., Rome, L. H. (1999) The 193-kD vault protein, VPARP, is a novel poly (ADP-ribose) polymerase. *Journal of cell biology* **146**, 917-928
57. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
58. Lescrinier, E. (2011) Structural biology in drug development. Verhandelingen - Koninklijke Academie voor Geneeskunde van Belgie
59. Li, P., Lei, Y., Qi, J., Liu, W., Yao, K. (2022) Functional roles of ADP-ribosylation writers, readers and erasers. *Frontiers in cell and developmental biology* **10**, 941356
60. Lord, C. J., Ashworth, A. (2017) PARP inhibitors: synthetic lethality in the clinic. *Science* **355**, 1152-1158
61. MedChemExpress: PARP10/15-IN-3
62. Metherell, L. A., Guerra-Assunção, J. A., Sternberg, M. J., David, A. (2016) Three-dimensional model of human Nicotinamide Nucleotide Transhydrogenase (NNT) and sequence-structure analysis of its disease-causing variations. *Human mutation* **37**, 1074-1084
63. Milewski, M. C., Broger, T., Kirkpatrick, J., Filomena, A., Komadina, D., Schneiderhan-Marra, N., Wilmanns, M., Parret, A. H. (2016) A standardized production pipeline for high profile targets from *Mycobacterium tuberculosis*. *Proteomics – clinical applications* **10**, 1049-1057
64. Morales, J., Li, L., Fattah, F. J., Dong, Y., Bey, E. A., Patel, M., Gao, J., Boothman, D. A. (2016) Review of poly-(ADP-ribose) polymerase (PARP) mechanisms of action and rationale for targeting in cancer and other diseases. *Critical reviews in eukaryotic gene expression* **24**, 15-28

65. National Cancer Institute: targeted therapy to treat cancer
66. Nicolae, C. M., Aho, E. R., Vlahos, A. H., Choe, K. N., De, S., Karras, G. I., Moldovan, G. L. (2014) The ADP-ribosyltransferase PARP10/ARTD10 interacts with proliferating cell nuclear antigen (PCNA) and is required for DNA damage tolerance. *Journal of biological chemistry* **289**, 13627-13637
67. Nielsen, J. (2013) Production of biopharmaceutical proteins by yeast: advances through metabolic engineering. *Bioengineered* **4**, 207-211
68. Nizi, M. G., Maksimainen, M. M., Murthy, S., Massari, S., Alaviuhkola, J., Lippok, B. E., Sowa, S. T., Galera-Prat, A., Prunskaite-Hyyryläinen, R., Lüscher, B., Korn, P., Lehtiö, L., Tabarrini, O. (2022) Potent 2,3-dihydrophthalazine-1,4-dione derivatives as dual inhibitors for mono-ADP-ribosyltransferases PARP10 and PARP15. *European journal of medicinal chemistry* **237**, 114362
69. Novagen: Ni-NTA His • Bind® Resin
70. Perina, D., Mikoč, A., Ahel, J., Ćetković, H., Žaja, R., Ahel, I. (2014) Distribution of protein poly-(ADP-ribosyl) ation systems across all domains of life. *DNA repair (Amst)* **23**, 4-16
71. Petty, K. J. (2001) Metal-Chelate Affinity Chromatography. *Current protocols in neuroscience* **5**, 5-10
72. PhRMA: Research & Development policy framework
73. Richard, I. A., Burgess, J. T., O'Byrne, K. J., Bolderson, E. (2022) Beyond PARP1: the potential of other members of the Poly (ADP-Ribose) Polymerase family in DNA repair and cancer therapeutics. *Frontiers in cell and developmental biology* **14**, 801200
74. Riley, R. S., June, C. H., Langer, R., Mitchell, M.J. (2019) Delivery technologies for cancer immunotherapy. *Nature reviews in drug discovery* **18**, 175-196
75. Robu, M., Shah, R. G., Petitclerc, N., Brind'Amour, J., Kandan-Kulangara, F., Shah, G. M. (2013) Role of poly-(ADP-ribose) polymerase-1 in the removal of UV-induced DNA lesions by nucleotide excision repair. *Proceedings of the National Academy of Sciences (PNAS)* **110**, 1658-1663
76. Rosenberg, S. A. (2014) IL-2: the first effective immunotherapy for human cancer. *Journal of immunology* **192**, 5451-5458
77. Sagen, A. (2023) YT medium. Protocols.io: protocols.io/YT medium

78. Sartorius: Vivaspin® 20 ultrafiltration units
79. Schuller, M., Riedel, K., Gibbs-Seymour, I., Uth, K., Sieg, C., Gehring, A. P., Ahel, I., Bracher, F., Kessler, B. M., Elkins, J. M., Knapp, S. (2017) Discovery of a selective allosteric inhibitor targeting macrodomain 2 of polyadenosine-diphosphate-ribose polymerase 14. *ACS chemical biology journal* **12**, 2866-2874
80. Scopes, R. K. *Protein purification: principles and practice* (3rd ed.) Springer, New York
81. Sonics: Vibra-cell VC130
82. Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein expression and purification* **207**, 34
83. Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F. (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* **71**, 209-249
84. Suskiewicz, M. J., Munnur, D., Strømland, Ø., Yang, J. C., Easton, L. E., Chatrin, C., Zhu, K., Baretić, D., Goffinont, S., Schuller, M., Wu, W. F., Elkins, J. M., Ahel, D., Sanyal, S., Neuhaus, D., Ahel, I. (2023) Updated protein domain annotation of the PARP protein family sheds new light on biological function. *Nucleic acids research* **51**, 8217-8236
85. Taylor, S. C., Posch, A. (2014) The design of a quantitative western blot experiment. *BioMed research international* **2014**, 361590
86. Thermo Fisher Scientific: NanoDrop 1000 Spectrophotometer
87. Thermo Fisher Scientific Patheon Pharma Services: Exploring the five phases of drug development
88. Valverde, R., Edwards, L., Regan, L. (2008) Structure and function of KH domains. *FEBS Journal* **275**, 2712-2726
89. Vincentelli, R., Cimino, A., Geerlof, A., Kubo, A., Satou, Y., Cambillau, C. (2011) High-throughput protein expression screening and purification in *Escherichia coli*. *Methods* **55**, 65-72
90. Walsh, G. (2010) Post-translational modifications of protein biopharmaceuticals. *Drug discovery today* **15**, 773-780

91. Wang, C., Yuan, X., Xue, J. (2023) Targeted therapy for rare lung cancers: status, challenges, and prospects. *Molecular therapy* **31**, 1960-1978
92. Ward, O. P. (2012) Production of recombinant proteins by filamentous fungi. *Biotechnology advances* **30**, 1119-1139
93. Wertz J. *The Coli Genetic Stock Center*, CGSC; Strain - DH5α
94. Wu, Q. J., Zhang, T. N., Chen, H. H., Yu, X. F., Lv, J. L., Liu, Y. Y., Liu, Y. S., Zheng, G., Zhao, J. Q., Wei, Y. F., Guo, J. Y., Liu, F. H., Chang, Q., Zhang, Y. X., Liu, C. G., Zhao, Y. H. (2022) The sirtuin family in health and disease. *Signal transduction and targeted therapy* **7**, 402
95. Wurm, F. M. (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature biotechnology* **22**, 1393-1398
96. Xiong, S., Zhang, L., He, Q. Y. (2008) Fractionation of proteins by heparin chromatography. *Methods in molecular biology* **424**, 213-221
97. Yoneyama-Hirozane, M., Matsumoto, S. I., Toyoda, Y., Saikatendu, K. S., Zama, Y., Yonemori, K., Oonishi, M., Ishii, T., Kawamoto, T. (2017) Identification of PARP14 inhibitors using novel methods for detecting auto-ribosylation. *Biochemical and biophysical research communications* **486**, 626-631
98. Zhao, S. S., Chen, D. D. (2014) Applications of capillary electrophoresis in characterizing recombinant protein therapeutics. *Electrophoresis* **35**, 96-108
99. Zhou, X., Yang, Y., Xu, Q., Zhou, H., Zhong, F., Deng, J., Zhang, J., Li, J. (2022) Crystal structures of the catalytic domain of human PARP15 in complex with small molecule inhibitors. *Biochemical and biophysical research communications* **622**, 93-100
100. Zhu, J. (2012) Mammalian cell protein expression for biopharmaceutical production. *Biotechnology advances* **30**, 1158-1170

Ringraziamenti

Giunto al termine del mio percorso di studi, ritengo doveroso dedicare un pensiero a tutti coloro che hanno saputo formarmi, sostenermi e supportarmi nel corso di questi anni.

E parto proprio dalla fine, dalle ultime persone che si sono inserite in questo cammino. Un ringraziamento speciale va al mio tutor e correlatore Matteo Leoncini; negli ultimi mesi hai saputo trasmettermi numerosissime conoscenze e competenze nell'ambito della *protein production* e non solo, arricchendo notevolmente il mio bagaglio culturale, e mi hai permesso di aprire gli occhi verso il futuro. Con la tua professionalità e la tua pazienza sei stato un vero e proprio compagno di viaggio, una spalla su cui poter sempre contare; con la tua simpatia e la tua gentilezza hai contribuito a rendere leggera e piacevole questa esperienza formativa. Condividere la postazione di laboratorio con te è stata una fortuna, non dimenticherò ciò che hai fatto per me. Sarai sempre un punto di riferimento professionale.

Non meno importante, un riconoscimento particolare va anche al mio relatore, prof. Riccardo Miggiano, per avermi dato la possibilità di vivere questa occasione formativa e di vita ospitandomi nel laboratorio di biochimica e biologia strutturale di Fondazione Novara Sviluppo; la ringrazio per il supporto di questi mesi e per aver valorizzato al massimo la mia esperienza di tesi. Inoltre, la ringrazio per la sua cordialità e la sua gentilezza, annullando quel rapporto studente-professore e permettendomi di inserirmi con serenità nel mondo pratico di un laboratorio di ricerca.

Rivolgo il mio grazie anche a tutti gli altri ragazzi del laboratorio: Andrea, Daniele, Erika, Fabio, Francesca, Giulia, Marta e Sonia. Il vostro contributo è stato prezioso, mi avete fatto sentire parte del gruppo come se non fossi un semplice tirocinante. Grazie per le chiacchierate e le risate fatte insieme, è anche grazie a voi che concludo con il sorriso questo fantastico percorso.

Vorrei dedicare un pensiero a tutti i miei compagni di corso, con i quali ho vissuto dei momenti straordinari e che hanno fatto sì che questi due anni siano passati molto, forse troppo, rapidamente. Le risate in classe, i momenti di confronto e tutto ciò che non ha riguardato strettamente il mondo universitario mi hanno riempito il cuore di gioia. Un grazie speciale va ad Alessio, Andrea, Chiara, Elisa, Federico, Francesco, Lorenzo, Sara e Simone;

siete stati dei compagni di viaggio meravigliosi, diventati amici nel giro di pochissimo tempo. So di poter contare sul vostro supporto per qualsiasi cosa e non vedo l'ora di condividere un sacco di avventure con voi.

Un doveroso ringraziamento va anche ai miei compagni di vita di sempre, Davide e Matteo; il vostro supporto è stato essenziale, probabilmente una delle forze maggiori che ho ricevuto in questi anni per andare avanti e raggiungere l'obiettivo; il semplice parlare, confrontarsi e vivere momenti spensierati mi hanno dato l'energia per proseguire nel cammino verso la meta. Sapere di potermi affidare a voi è un tesoro che custodirò per sempre.

Ringrazio di cuore anche Gabriele, Simone e tutto il gruppo di amiche del mio paese: Elena, Francesca, Gaia C., Gaia G., Giulia, Matilde, Sara B., Sara C. e Silvia. La vostra vicinanza, consolidata in questi ultimi anni, è stata un bene prezioso ed essenziale per staccare la spina dallo studio quando necessario e dare spazio a momenti indimenticabili. A volte non mi rendo conto della fortuna che ho ad avervi come sostegno e compagni di viaggio; un posto nel mio cuore è dedicato a tutti voi.

Arrivo alla conclusione di questi ringraziamenti, dicendo il grazie più grande di tutti alla mia famiglia. A mamma Rosella e papà Elio, che fin da piccolo hanno saputo educarmi, sostenermi nei momenti di maggiore difficoltà, guidarmi con i loro saggi consigli, supportarmi economicamente ma soprattutto moralmente, dandomi sempre la motivazione per spingermi oltre. Avere dei genitori pronti a sacrificarsi fino alla fine per il proprio figlio è un bene immenso; per questo e per quanto non ho detto in queste poche righe il grazie per voi non sarà mai sufficiente. Ringrazio anche mio fratello Gianluca, per tutto il supporto mostrato da sempre; sei stato il più grande esempio a cui mi sono sempre ispirato e, ascoltandoti dai tempi delle superiori, hai saputo farmi appassionare al mondo scientifico e poi a quello farmaceutico. Ancora oggi rappresenti un punto di riferimento da cui cerco di imparare il più possibile.

E a tutti coloro che non ho citato in queste pagine, che, seppur inconsapevolmente, hanno saputo darmi una mano nel credere sempre in me stesso e guardare avanti, è rivolto il mio grazie.