Universitat de Girona

Facultat de Ciències

Memòria del Treball Final de Grau

Títol del treball:

Cloning, production and purification of proteins that interact with Apoptin

---

Estudiant: Raquel Vivet Noguer

Grau en Biotecnologia

Correu electrònic: riki_vivet@hotmail.com

Tutor: Dr. Antoni Benito Mundet

Empresa / institució: Universitat de Girona

Vistiplau tutor:

Nom del tutor: Dr. Antoni Benito Mundet

Empresa / institució: Universitat de Girona

Correu(s) electrònic(s): antoni.benito@udg.edu

Data de dipòsit de la memòria a secretaria de coordinació: 26 de maig del 2016

ÍNDEX

LIST OF ABBREVIATIONS

Apop	Apoptin
CAV	Chicken anemia virus
IDP	Intrinsically disordered protein
Bca3	Breast cancer associated gene 3
Ppil3	Peptidyl-prolyl isomerase-like 3
GST	Glutathione S-transpherase
NLS	Nuclear localization sequence
NES	Nuclear export signal
PRS	Proline-rich segment
LRS	Leucine-rich segment
PKCβ	Protein Kinase C
FADD	Fas-associated protein with death domain
PKG-I	Protein Kinase G
PI3K	PI3-Kinase
Hsc70	Heat shock cognate protein 70
PML	Promyelocytic leukemia protein
APC1	Anaphase promoting complex 1
Hippi	Hip-1 protein interactor
DEDAF	Death effector domain-associated factor
Nmi	N-myc interacting protein
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
LB	Luria-Bertani
His	Histidine
DNA	Deoxyribonucleic acid
TAE	Tris-acetate-EDTA
EDTA	Ethylenediaminetetraacetic acid
ddNTP	Dideoxynucleotide
SDS-PAGE	SDS-Polyacrilamide gel electrophoresis
SDS	Sodium dodecyl sulfate
ddH_2O	Bidistilled water
PMSF	Phenylmethanesulfonyl fluoride
psi	Pounds per square inch
rpm	Revolutions per minute
HPLC	High-pressure liquid chromatography
atm	Atmosphere
PBS	Phosphate-buffered saline
Ni-NTA	Nickel-nitrilotriacetic acid
MPa	Megapascal
kDa	Kilodalton
PCR	Polymerase Chain Reaction
Tm	Melting temperature
λ	Wavelength

RESUM

El càncer és un conjunt de patologies d'origen clonal que apareix per una acumulació de mutacions en un organisme. Es caracteritza per un creixement anòmal de cèl·lules que tenen tendència a proliferar de forma incontrolada donant lloc a una acumulació de massa o teixit que s'anomena tumor. El càncer és una de les principals causes de mortaldat en l'àmbit mundial. Els tractaments actuals del càncer poden ser molt invasius i no sempre són eficaços. És per això que avui en dia s'estan investigant altres tipus de teràpies més dirigides com per exemple el tractament amb l'Apoptina que mata específicament les cèl·lules canceroses i no afecta les altres cèl·lules.

L'Apoptina és una proteïna codificada pel genoma del virus de l'anèmia del pollastre que indueix específicament la mort cel·lular de cèl·lules tumorals. Aquesta proteïna té una alta tendència a formar agregats i forma part del grup de les proteïnes intrínsecament desplegades, un conjunt de proteïnes sense una estructura definida que poden acoblar-se a diferents proteïnes mitjançant canvis de conformació. L'objectiu d'aquest Treball Final de Grau s'emmarca dins d'un projecte més ampli que consisteix en estudiar els canvis estructurals que pateix l'Apoptina en interaccionar amb proteïnes implicades en la senyalització cel·lular per tal de conèixer millor les bases moleculars del mecanisme de citotoxicitat d'aquesta proteïna.

En aquest treball s'ha dut a terme la clonació de dues proteïnes que prèviament s'ha descrit que interaccionen amb l'Apoptina. Es tracta de la *Breast cancer associated gene 3* i la *Peptidyl-prolyl isomerase-like 3* les quals s'han fusionat a la Glutatió S-transferasa. Per això s'ha dissenyat cada un dels dos gens optimitzant l'ús de codó i introduint les dianes de restricció adequades per permetre el seu clonatge en el vector pGEX-4T-2. El vector conté la seqüència d'ADN de la Trombina que servirà per poder separar la proteïna recombinant de la Glutatió S-transferasa. Un cop duta a terme la clonació, s'han produït les proteïnes recombinants en la soca d'*Escherichia coli* XL1Blue i s'han purificat mitjançant una cromatografia d'afinitat amb glutatió. Paral·lelament s'ha produït i purificat una variant de l'Apoptina a la qual se li han eliminat els residus 1 fins al 43, anomenada *H_6-ApoptinaΔProΔLeu,* disminuint-li la tendència a formar agregats sense que per això s'alterin les seves propietats citotòxiques sobre cèl·lules tumorals. Un cop purificades s'han realitzat experiments de *pull-down* per tal d'investigar el grau d'interacció *in vitro* d'aquestes proteïnes amb l'Apoptina.

Paraules clau: Càncer, Apoptina, proteïna antitumoral, agregació de proteïnes, *Breast cancer associated gene 3, Peptidyl-prolyl isomerase-like 3,* citotoxicitat, apoptosi, senyalització cel·lular.

RESUMEN

El cáncer es un conjunto de patologías de origen clonal que tienen lugar debido a una acumulación de mutaciones en un organismo. Está caracterizado por un crecimiento anómalo de células que presentan una tendencia a proliferar de forma descontrolada, dando lugar a una acumulación de masa o tejido llamada tumor. El cáncer es una de las principales causas de mortalidad en el ámbito mundial. Los actuales tratamientos de cáncer pueden llegar a ser muy invasivos y no siempre son eficaces. Es por este motivo que actualmente se están investigando otro tipo de terapias más dirigidas, tales como el tratamiento con Apoptina que mata específicamente las células cancerígenas sin afectar a las otras células.

La Apoptina es una proteína codificada en el genoma del virus de la anemia del pollo que induce específicamente la muerte celular de células tumorales. Esta proteína tiene una elevada tendencia a formar agregados y forma parte del grupo de proteínas intrínsecamente desestructuradas, una serie de proteínas que no presentan una estructura definida y pueden acoplarse a distintas proteínas mediante cambios en su conformación. El objetivo de este Trabajo de Fin de Grado forma parte de un proyecto más amplio que consiste en estudiar los cambios estructurales que padece la Apoptina al interaccionar con otras proteínas implicadas en la señalización celular, con la intención de conocer mejor las bases moleculares del mecanismo de citotoxicidad de esta proteína.

En este trabajo se ha llevado a cabo la clonación de dos proteínas de las cuales previamente se había descrito que interaccionan con la Apoptina. Concretamente se trata de las proteínas *Breast cancer associated gene 3* y *Peptidyl-prolyl isomerase-like 3* que se han fusionado a la Glutatión S-transferasa. Por esta razón se han diseñado cada uno de los genes optimizando el uso de codón e introduciendo las dianas de restricción correspondientes para permitir su clonaje en el vector pGEX-4T-2. Este vector contiene la secuencia de ADN de la Trombina que permite separar la proteína recombinante de la Glutatión S-transferasa. Una vez realizada la clonación, se han producido las proteínas recombinantes en la cepa de *Escherichia coli* XL1Blue y se han purificado mediante una cromatografía de afinidad con glutatión. Paralelamente se ha producido y purificado una variante de la Apoptina a la cual se le han eliminado los residuos 1 hasta el 43, llamada *H_6-ApoptinaΔProΔLeu,* disminuyéndole la tendencia a agregar pero sin alterar sus propiedades citotóxicas en células tumorales. Después de la obtención de las proteínas purificadas se han realizado ensayos de *pull-down* con la intención de investigar el grado de interacción *in vitro* de estas proteínas con la Apoptina.

Palabras clave: Cáncer, Apoptina, proteína antitumoral, agregación de proteínas, *Breast cancer associated gene 3, Peptidyl-prolyl isomerase-like 3,* citotoxicidad, apoptosis, señalización celular.

ABSTRACT

Cancer is a group of pathologies of clonal origin that appears due to an accumulation of mutations in an organism. It is characterized by an abnormal growth of cells that tend to proliferate in an uncontrolled way leading to the formation of a mass or tissue accumulation commonly known as tumor. Cancer is one of the main causes of mortality worldwide. Current cancer treatments can be invasive and sometimes fail at being efficient. For this reason, research is currently focused on targeted therapies. For instance, treatment with Apoptin, which specifically kills cancer cells leaving non-cancer cells unharmed.

Apoptin is a protein coded by the genome of the chicken anemia virus that specifically kills tumor cells. This protein has a strong tendency to aggregate and is a member of the intrinsically disordered proteins, a group of proteins that do not have a defined structure and can interact with several proteins by conformational changes. The aim of this Final Grade Project forms part of a wider project that consists in studying the structural changes that Apoptin suffers when interacting with proteins related to cellular signaling in order to find out more about the molecular basis of the mechanism of cytotoxicity of this protein.

In this project I have cloned two proteins that are known to interact with Apoptin. These proteins are Breast cancer associated gene 3 and Peptidyl-prolyl isomerase-like 3 which have been fused to Glutathione S-transferase. For this reason, each gene has been designed choosing the optimal codon usage and introducing the appropriate restriction sites with the aim of cloning them in pGEX-4T-2 vector. This vector contains the DNA sequence that codes for Trombine, which can be used for the removal of Glutathione S-transferase from the recombinant protein. Once the clonation has been finished, the recombinant proteins have been produced in *Escherichia coli* strain XL1Blue and a variant of Apoptin has been purified. This variant is called H_6-ApoptinΔProΔLeu and its residues 1 to 43 have been eliminated in order to reduce its tendency to aggregate without affecting its cytotoxic effect on cancer cells. Once purified, pull-down assays have been performed with the aim of investigating the *in vitro* interaction of these proteins with Apoptin.

Keywords: Cancer, Apoptin, antitumor protein, protein aggregation, Breast cancer associated gene 3, Peptidyl-prolyl isomerase-like 3, cytotoxicity, apoptosis, cellular signaling.

INTRODUCTION

1. Cancer

Cancer is one of the main causes of death worldwide. It is due to an abnormal and uncontrolled growth of a group of cells that lead to the formation of a tumor. Tumors can metastasize, which means that tumor cells have the capacity to invade the surrounding tissue and break off from the main tumor to enter into the bloodstream or into the lymphatic system. Thus metastatic tumor cells can spread to other parts of the body and become malignant tumors. In fact, metastasis is one of the main medical hindrances to treat cancer. On the other hand, cells that form benign tumors cannot metastasize (Mantovani, 2009).

2. Current treatments

Cancer treatment depends on the type of cancer and the stage of the disease. Current treatment for benign tumors is surgery whereas treatments for malignant tumors are radiation and chemotherapy. Sometimes, these therapies complement each other. Even though their efficacy and specificity for tumor cells have been improved in the last decades, they still present different important drawbacks that must be bypassed:

Firstly, surgery is intended to remove the tumor but it can be invasive and it is only useful if the cancer is detected on the early stages. Secondly, radiation consists of directing energy to the tumor, which leads to the formation of oxygen radicals that end up damaging the cancer cells. However, radiation can also affect non-cancer cells and therefore the levels of radiation have to be cautiously calculated. Thirdly, chemotherapy affects cells that present high rates of proliferation by interfering with vital processes like DNA synthesis or cell proliferation. However, there are other cells, apart from tumor cells, that proliferate a lot, like cells lining the intestines, cells of the hair follicles or bone marrow cells, which can be harmed leading to nausea, hair loss, etc.

3. Novel treatments

To overcome these drawbacks, new strategies are being developed for the design of new drugs for cancer treatment. These novel treatments are intended to be targeted. This term refers to drugs or other substances that block the proliferation and spread of cancer by interfering with specific molecules that are involved in the growth, progression and spread of the disease. Among these targeted therapies we can include hormone therapies, signal transduction inhibitors, gene expression modulators, apoptosis inducers, angiogenesis inhibitors, immunotherapies, viruses and monoclonal antibodies that deliver toxic molecules (Ramaswami et al., 2016).

4. Virus-based cancer therapies

Virus-based cancer therapies are found on viruses that specifically infect tumor cells. These viruses are called oncolytic viruses and could be used with therapeutic proposes in order to stimulate the immune system to attack the malignant cells. Therefore, normal cells would not be affected. Advantages of this treatment are reaching organs that are difficult to access and

the fact that the patient would manage it to confront the illness by itself (Ledford, 2015). On the one hand, there are viruses that block apoptosis with the aim of protecting infected cells from early death by producing specific proteins that inhibit pro-apoptotic factors. On the other hand, some viruses can induce multiple mechanisms that kill the infected cells including apoptosis. This is an important characteristic because many cancer cells avoid apoptosis due to an accumulation of alterations in the apoptotic machinery (Ruiz-Martínez, 2016).

Moreover, viruses are professional gene delivery vehicles. For instance, genes coding for specific toxins (Seow & Wood, 2009). Finally, some viral products can be independently used as recombinant proteins to induce apoptosis to cancer cells. This is the case for example of Apoptin, a protein obtained from the Chicken anemia virus (CAV).

5. Apoptin

CAV, a member of the *Gyrovirus* genus in the *Circoviridae* family, is an icosahedral virus with a negative-sense, single-stranded circular DNA genome of 2.298 – 2.319 nucleotides in length. The genome contains three partially overlapping reading frames that encode the viral proteins VP1, VP2 and VP3 (Zhang et al., 2012). VP3, also known as Apoptin, is a 13.6 kDa protein needed for replication. It comprises a C-terminal bipartite nuclear localization sequence (NLS) as well as a nuclear export signal (NES) that are required for the transport of Apoptin (Grimm & Noteborn, 2010). Within the NLS1 there is a Proline-rich segment (PRS) from residues 81 to 86. At the N-terminal domain, Apoptin possesses a Leucine-rich segment (LRS) spanning residues 33 to 46 that promotes aggregation (Leliveld et al., 2003) and a second PRS from residues 8 to 28 (Noteborn et al., 1994) (see Figure 1).

1 MNALQEDT**PP GPSTVFRPPT SSRPLETP**HC RE**IRIGIAGI TITLSL**CGCA NARAPTLRSA
61 TADNSESTGF KNVPDLRTDQ PKPPSKKRSC DPSEYRVSEL KESLTTTTPS RPRTARRCIR L

Figure 1. Primary structure of Apoptin. The localization of key functional domains of the protein like nuclear localization sequence (NLS) and nuclear export signal (NES) are indicated. The lower panel shows Apoptin's amino acid sequence with the proline rich sequence (PRS) and leucine rich sequence (LRS) in bold (adapted from Grimm & Noteborn, 2010).

Even though Apoptin's function in the viral cycle has not yet been discovered, it is known that Apoptin induces apoptosis in chicken thymocytes and chicken lymphoblastoid cell lines (Jeurissen et al., 1992). Moreover, long term expression of Apoptin in normal human fibroblasts has revealed that it does not produce toxic or transforming activities in these cells and that Apoptin does not interfere in normal cell proliferation (Backendorf et al., 2008). Apoptin also induces apoptosis in many types of mammalian transformed and malignant cell lines without affecting primary and non-transformed cells (Burek et al., 2006; Poon et al., 2005).

Apoptin-induced apoptosis is independent of p53. This is an important feature for its potential as a possible anti-cancer therapeutic agent since many tumors contain a mutated, non-functional p53 (Grimm & Noteborn, 2010). In fact, Apoptin seems to induce programmed cell

death of cells after translocation from the cytoplasm to the nucleus and arresting the cell cycle at G2/M (Moore et al., 2005).

Even though the mode of action of this protein is currently unknown, several mechanisms have been proposed (Tavassoli et al., 2005):

The tumor-specific nuclear translocation and phosphorylation of Apoptin by a protein kinase seem to be required for its function (Grimm & Noteborn, 2010). Firstly, in tumor cells or transformed cells, Apoptin is mainly expressed in the nucleus whereas in normal cells it is mainly localized in the cytoplasm. Nevertheless, forced expression of Apoptin into the nucleus of normal cells *per se* does not result in apoptosis, showing that additional factors are required for the activation of Apoptin. The translocation of Apoptin to the nucleus is regulated by the protein's NLS and NES. Additionally, localization of Apoptin can also be affected by the interaction with cellular proteins resulting in its nuclear or cytoplasmatic retention (Danen-Van Oorschot et al., 2003). Secondly, in cancer cells but not in normal cells Apoptin is phosphorylated on threonine 108 by a cellular kinase (Zhang et al., 2004).

It has been recently described that Apoptin is a member of the intrinsically disordered proteins (IDPs) (Ruiz-Martínez et al., 2016a submitted). IDPs are a class of proteins that almost a significant portion of their sequence lack well-structured three-dimensional fold presenting high flexibility and structural instability (Oldfield & Dunker, 2014). However, IDPs display different signaling or regulatory functions because of this distinctive property. Firstly, these proteins provide a larger interaction surface area than globular proteins of similar size. Secondly, they have a great conformational flexibility and the exposure of their peptide motifs allows IDPs to scaffold and interact with numerous other proteins. Thirdly, different post-translational modifications facilitate regulation of their function and stability in a cell (Dyson & Wright, 2005). Interestingly, many disordered segments fold on binding to their biological targets, whereas others constitute flexible linkers that have a role in the assembly of macromolecular structures.

Although much remains to be studied, it has been described that more than a third of the eukaryotic proteins contain intrinsically disordered regions of more than 30 residues in length (Ward et al., 2004). Additionally, a deregulated expression of IDPs is associated with several diseases such as cancer and neurodegeneration. Taking all this information into consideration, the presence of IDPs in cells has to be tightly regulated (Uversky et al., 2008).

To exert its tumor-specific cytotoxicity Apoptin needs to be activated to be able to trigger cellular signaling pathways. Several proteins have been identified to interact or co-localize with Apoptin and influence its cytotoxic activity or its cellular localization. These proteins are Protein Kinase C β (PKCβ), Fas-associated protein with death domain (FADD), Bcl-10, Protein Kinase G (PKG-I), PI3-Kinase (PI3K) and Akt, Heat shock cognate protein 70 (Hsc70), Promyelocytic leukemia protein (PML), Anaphase promoting complex 1 (APC1), Hip-1 protein interaction (Hippi), Death effector domain-associated factor (DEDAF), Peptidyl-prolyl isomerase-like 3 (Ppil3), N-myc interacting protein (Nmi) and Breast cancer associated gene 3 (Bca3) (Grimm & Noteborn, 2010). The study of how Apoptin interacts with these partners

may help to clarify the molecular determinants of the anticancer activity of Apoptin. Among the proteins described to interact with Apoptin, this work is focused on Ppil3 and Bca3.

Bca3 enhances Apoptin phosphorylation and cytotoxic function by an unknown mechanism. Bca3 binds to Apoptin in normal human fibroblasts. Overexpression of Bca3 has shown to enhance Apoptin phosphorylation and increase apoptosis induction in some types of cells like Saos-2. However, the significance of their interaction remains unknown. Bca3 interacts with proteins like protein kinase A and Tap73 to modify cellular signaling pathways (Harada & Berk, 1999).

Ppil3 is part of the cyclophilin family which catalyzes the cis-trans isomerization of peptide bonds at proline residues and can regulate protein stability. Proteins of this family are associated with the mitochondrial maintenance, cell cycle and apoptosis progression (Flotte et al., 1996). Binding of Ppil3 to Apoptin regulates Apoptin localization in both normal and tumor cells. In fact, an overexpression of this protein can promote cytoplasmic localization of Apoptin (Bergmann et al., 2001).

Due to the fact that the use of Apoptin in cancer therapy is limited by its strong tendency to aggregate, the research group where I have carried out my Final Grade Project had previously constructed a variant of Apoptin that lacks residues 1 to 43 suppressing the PRS and LRS regions (see Figure 1). This variant is a soluble non-aggregating protein of 10.8 kDa that maintains most of the biological properties of wild type Apoptin when transfected into cells. Specifically, the deleted residues are at the N-terminus and consist of a PRS and a LRS coding for the aggregation-prone segment (see Figure 1). Furthermore, this variant which is called H_6-ApopΔProΔLeu maintains its cytotoxic effect when is added exogenously or when its gene is transfected (Ruiz-Martínez et al., 2016b submitted). It has also an N-terminal His-tag to facilitate the process of purification. This variant is the one used in this project.

OBJECTIVES

Nowadays, cancer is one of the principal causes of death worldwide. Apoptin represents a very interesting alternative for the treatment of cancer. However, the molecular mechanism of its anticancer activity is only partially known. It has been recently described that Apoptin is an IDP. This kind of proteins can adapt its conformation to bind different types of partners. However, a drawback that makes difficult to obtain this knowledge is that Apoptin has a significative tendency to aggregate, hampering for example its structure determination. The research group where I have performed this Final Grade Project has constructed a new modified variant of this protein that has no tendency to aggregate but is still cytotoxic for cancer cells. This new variant can be used to study the structural changes that Apoptin suffers when it interacts with proteins that are involved in the signaling pathways.

As part of this project, my Final Grade Project has consisted in the selection, cloning and production of different binding partners for Apoptin whose interaction could be lately studied by NMR. Hence, the specific objectives of my project are:

- Design and construction of the genes coding for Breast cancer associated gene 3 (Bca3) and Peptidyl-prolyl isomerase-like 3 (Ppil3) fused to the gene of Glutathione S-transferase (GST) in the expression vector pGEX-4T-2.
- Production and purification of the two recombinant proteins coded by these genes.
- Purification of the H_6-ApopΔProΔLeu variant.
- Analysis through pull-down assays of the interaction of these two proteins with the H_6-ApopΔPro variant.

MATERIALS AND METHODS

1. Biological material

a) Escherichia coli strains

XL1-Blue. These cells are ideal for routine cloning needs such as propagation of plasmid vectors. The genotype is: *rec*A1 *end*A1 gyrA96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 *lac* [F' *proAB lacI*^{q}*ZΔM*15 Tn*10* (Tet^r)]. XL1-Blue cells are nalidixic acid resistant and tetracycline resistant.

BL21 (DE3). This strain has been used for the expression of pET28a vector containing the H_6-ApopΔProΔLeu gene. Gene *int* from the bacterial genome has been inactivated through the insertion of the prophage DE3 (derived from phag λ). This bacteriophage contains *lac*UV5 promoter controlling the expression of the initial region of *lac*Z gene as well as the gene that codes for T7 RNA polimerase. Consequently, gene *int* stays inactivated and DE3 prophage cannot either excise from the genome, resulting in a stable lisogen. In these conditions, *lac*UV5 is the only promoter that is able to induce the T7 RNA polymerase transcription by the addition of isopropyl-β-D-1-thiogalactopyranoside (IPTG). The genotype is: F$^-$ *ompT gal dcm lon hsdS*$_B$($r_B^- m_B^-$) λ(DE3 [*lacI lacUV5-T7 gene 1ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S).

b) Medium

The medium used has been Luria-Bertani (LB) medium. It is a rich medium commonly used for the growth of bacteria. The composition used was 10 g of tryptone, 10 g of NaCl and 5 g of yeast extract, dissolved in 1 L of distilled water. The mix was autoclaved on liquid cycle for 20 minutes at 121ºC. The solution was cooled to 55ºC and the antibiotic was added (50 µg/mL for ampicillin or kanamycin). The antibiotics were previously sterilized by filtration. For preparing LB agar-plates 15 g of agar were added to the medium before autoclaving. The solution was cooled to approximately 55ºC, the antibiotic ampicillin (50 µg/mL) was added and then, the medium was poured into petri dishes. Once the medium was solid, the plates were inverted and stored at 4ºC in the dark.

c) Plasmids

The vector used for cloning the genes coding for Bca3 and Ppil3 has been pGEX-4T-2 (General Electric Healthcare Life Sciences, Sweden) (see Figure 2). It is an expression vector that contains the Ptac promoter for chemical induction with IPTG and a gene that codes for the GST intended to create a fusion protein in order to help on the purification and stabilization of the

recombinant proteins. The vector also contains a sequence that codes for the Thrombin cleavage signal that allows the removal of GST from the recombinant protein. This vector confers resistance to ampicillin.

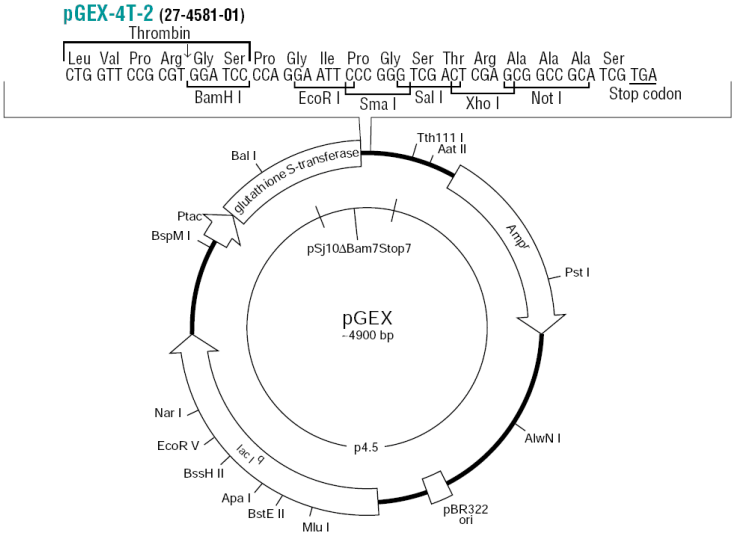

Figure 2. Map of the pGEX-4T-2 vector showing the reading frame of its multicloning site and its main features. The stop codon is also included. Obtained from: http://www.biofeng.com/zaiti/dachang/pGEX-4T-2.html

The vector carrying the H_6-ApopΔProΔLeu gene variant is pET-28a(+) (Novagen, USA) (see Figure 3). It is an expression vector that contains the T7 promoter for the expression using the T7 RNA polymerase under the control of T7lacUV5 inducible with IPTG encoded in the BL21 (DE3) genome. It also contains stretches coding for Histag/thrombin/T7-Tag and for a His-Tag sequence at the 5' and 3' ends of the polylinker, respectively. This vector confers resistance to kanamycin.

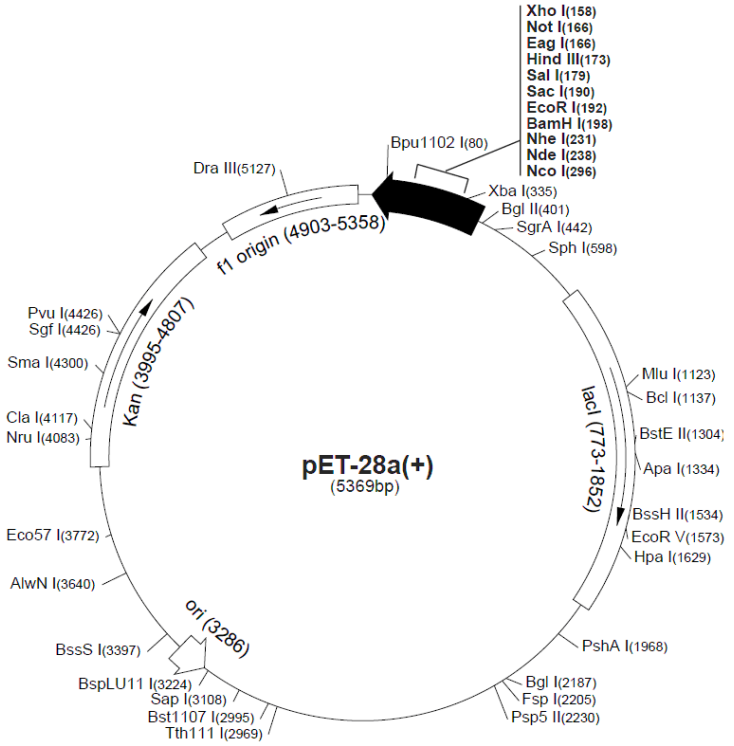

Figure 3. Map of the pET-28a(+) vector. The different genes and restriction sites are indicated. Obtained from: http://www.biofeng.com/zaiti/dachang/pet28a.html

2. Recombinant DNA techniques

a) Plasmid DNA purification (miniprep)

Plasmid DNA was purified using the commercial Kit *GeneJET™ Plasmid Miniprep Kit* (Thermo Fisher Scientific, Lithuania). The process of plasmid DNA purification was carried out following the protocol recommended by the supplier. The mechanism is based on the alkaline lysis method and uses spin columns based on the silica-based membrane technology with high affinity for DNA. Firstly, there are some steps to retain the plasmid in the silica and finally the plasmid is eluted using an elution buffer.

b) Digestion of the DNA fragments

The digestion of DNA fragments was performed incubating 45 µl of DNA with 5 µl of the corresponding 10-fold concentrated restriction enzyme's buffer and 1 µl of the desired corresponding restriction enzyme. Finally, the resulting mix was incubated 1 hour at 37ºC. It is important to add the exact volume because too much buffer could modify the conductivity or the pH, altering the enzyme's specificity for its target. This phenomenon is called star activity. The restriction enzymes used have been *Bam*HI and *Xho*I (Roche, Germany). These enzymes must be stored at -20ºC and therefore the suppliers keep them diluted with a 50% of glycerol. During the reaction, the quantity of glycerol cannot be higher than 5%, otherwise it could also induce star activity.

c) Purification of DNA fragments through agarose gel electrophoresis (GeneClean)

Digested DNA was subjected to an agarose gel (1%) electrophoresis to separate and identify the DNA fragment of interest. Firstly, 0.5 g of agarose were melted in 50 mL of Tris-acetate-EDTA (TAE) buffer (40 mM Tris base, 20 mM acetic acid and 1mM EDTA). Then, 5 µl of ethidium bromide were added and the gel allowed polymerize. Samples were mixed with the DNA loading buffer 6-fold concentrated (0.03% bromophenol blue, 0.03% xylene cyanol FF, 50% glycerol, 60 mM EDTA, 10 mM Tris-HCl pH 7.6 (Fermentas, Lithuania)) at a final concentration 1x and loaded inside the gel pockets. The gel was run at 110 volts for 30 minutes.

Once the gel had run, it was irradiated with UV light and the desired DNA fragment was excised and purified using the GeneClean method employing the commercial Kit *GeneJET™ Gel Extraction Kit* (Thermo Scientific, Lithuania). The process was carried out following the protocol recommended by the manufacturer. Firstly, it elutes DNA from agarose by melting the agarose. Then, the kit provides spin columns to isolate DNA based on its binding properties to silica membrane in the presence of a chaotropic salt. Once the salt is removed, the DNA detaches from the silica using an aqueous buffer.

d) Ligation

The ligation reaction was performed at 18ºC overnight by incubating the corresponding DNA fragments with 1 µl DNA ligase from bacteriophage T4 (Roche, Germany) in the buffer supplied

by the manufacturer. A control was included by adding bidistilled water instead of the DNA insert. This control allows you to verify that the vector is completely digested and cannot ligate without the insert. The products of the ligation were transformed in competent cells.

The transformation of competent cells was carried out according to the following procedure. 100 µl of competent cells (Stratagene, USA) were incubated at 4ºC with 2 µl of DNA coming from a miniprep (see section 2.a) or 10 µl of DNA coming from a ligation. The mix was incubated 30 minutes at 4ºC. Then, a thermal shock was induced by submitting the cells at 42ºC for 90 seconds and incubated in ice for 5 minutes. The phenotypic recovery was carried out adding 300 µl of LB medium and incubating at 37ºC during an hour. Then, the cells were transferred to LB plates supplemented with the corresponding antibiotic and incubated at 37ºC overnight.

e) DNA sequencing

In order to check the sequence of the recombinant DNA products, DNA sequencing was performed using the Kit *BigDye Terminator v 3.1,* based on the Sanger sequencing method. The process is based on the incorporation of dideoxynucleotides (ddNTPs) by DNA polymerase during in vitro DNA replication. These ddNTPs do not have the 3' hydroxyl so when they are added the replication is stopped because the phosphodiester bond can not be formed. The four types of ddNTPs are fluorescently labeled in order to be distinguished.

The sequencing mix contained 6 µl of DNA sample, 0.5 µl of the corresponding primer, 1.5 µl of Ready mix buffer containing the DNA polymerase (Roche, Germany) and 2 µl of sequencing buffer. The primers used for DNA sequencing (Thermo Fisher Scientific, USA) are depicted in Table 1. The Polymerase Chain Reaction (PCR) steps were of 30 seconds at 96ºC, 15 seconds at 50ºC and 4 minutes at 60ºC for 25 cycles. The product of the PCR was submitted to a precipitation process as follows: 1 µl of EDTA 125 mM and 1 µl of sodium acetate 3M pH 5.2 were added to the samples. Then, 25 µl of cold ethanol 100% were added and the mix was left 15 minutes at room temperature. The mix was centrifuged 15 minutes at 15,000 rpm at 4ºC. The supernatant was discarded and 35 µl of cold ethanol 70% were added. The mix was centrifuged 10 minutes at 15,000 rpm at 4ºC. Finally, the supernatant was discarded and the pellet was resuspended with 10 µl of formamide. The resuspended samples were analyzed at the Laboratory of Genetic Ictiology of the Universitat de Girona.

Table 1. Oligonucleotides used for DNA sequencing. The sequence, molecular weight and the melting temperature (Tm) of each oligonucleotide is indicated.

Primer	Sequence 5' to 3'	Molecular weight (µg/µmole)	Tm 50 mM Na+ (ºC)
pGEX 5 (forward)	GGG CTG GCA AGC CAC GTT TGG TG	7137,6	57
pGEX 3 (reverse)	CCG GGA GCT GCA TGT GTC AGA GG	7146,6	57

3. Protein detection techniques

a) SDS-Polyacrilamide gel electrophoresis

SDS-Polyacrilamide gel electrophoresis (SDS-PAGE) is a technique that separates molecules by their molecular weight. The molecules are forced through a sieving gel matrix by an electrical current. In this study this technique has been used at different points to identify the protein of interest in different samples.

It consists of a discontinuous gel formed by a stacking gel and a separating one. The stacking gel contains a 4% of acrylamide that is insufficient to separate the proteins. Thus all proteins accumulate on the stacking gel and will start migrating on the separating gel at the same time. The separating gel contains a higher concentration of acrylamide (15-20%) that allows the separation of the protein according to their electrophoretic mobility.

The gel used contained a 15% of acrylamide and was 0.75 mm of thickness. Once the gels had polymerised, the samples were mixed with 4-fold concentrated loading buffer (1 M Tris-HCl pH 6.8, ddH_2O, SDS, 0.1% Bromophenol Blue, 50% glycerol and 14.3 M β-mercaptoethanol) and were boiled at 90ºC for 5 minutes. The electrophoresis was carried out at 200 volts. The gel was dyed with Coomassie Brilliant Blue to visualize the bands of the different proteins.

4. Production of recombinant proteins

The desired clones were inoculated in 25 mL of LB medium supplemented with corresponding antibiotic (ampicillin for the PGEX-4T-2 based vectors and kanamycin for H_6-ApopΔProΔLeu). The culture was incubated at 37ºC and 220 rpm overnight.

The next day, 4 mL of the overnight culture were inoculated in 5 Erlenmeyer flasks of 1 L each containing 400 mL of LB medium supplemented with the desired antibiotic. These cultures were incubated at 37ºC and 250 rpm for approximately 3 – 4 hours, until they reached an optical density (λ=550 nm) between 1-2 a.u. Then, 1 mL of IPTG 400 mM was added to each Erlenmeyer flask in order to get a final concentration of 1 mM. The induced culture was further incubated at 37ºC and 250 rpm during 4 hours. The 2 L of culture were centrifuged 10 minutes at 4,000 rpm. The pellet that contains the cells was stored at -80ºC.

Finally, the samples were centrifuged 10 minutes at 4,000 rpm. The bacterial pellets obtained from the production of Bca3 and Ppil3 were resuspended with 30 mL of the buffer Phosphate-buffered saline (PBS) and Phenylmethylsulfonyl fluoride (PMSF) 1 mM. Those pellets resulting from H_6-ApopΔProΔLeu variant's purification were resupended with 30 mL of a buffer containing sodium acetate 50 mM and PMSF 1 mM at pH 5.0.

To check protein expression an aliquot of 1 mL was taken just before the addition of IPTG and another aliquot of 500 µl was taken after 4 hours of induction. These aliquots were centrifuged 30 seconds at 10,000 rpm and the pellets were resuspended in 100 µl of H_2O and stored at -20ºC until their analysis through SDS-PAGE (see section 3.a).

5. Protein purification and concentration techniques

In this section the processes of purification for proteins Bca3 and Ppil3 are explained separately from the one of H_6-ApopΔProΔLeu variant.

Purification of Bca3 and Ppil3

a) French Press

French press is a device based on the use of an external hydraulic pump that drives the piston within a cylinder that contains the sample. When the piston descends, the cells are submitted at high pressure. The apparatus also has a valve so that, as the sample passes through the valve is submitted to atmospheric pressure. This change of pressure causes the disruption of the cells – therefore, the protein of interest is released.

The resuspended pellets were lysed three times in the French press (SLM-Aminco® Spectronic Instruments, UK) at 1100 psi. 100 µl from the lysed sample were taken, centrifuged 30 seconds at 10,000 rpm and equivalent amounts of supernatant and pellet were analysed on a SDS-PAGE (see section 8). The rest of the lysed samples were centrifuged at 9,000 rpm for 30 minutes.

b) High-pressure Liquid Chromatography

The supernatants from the bacterial lysates obtained with the French Press were submitted to a High-pressure liquid chromatography (HPLC). It is a technique used to separate, identify and quantify the different components in a mixture. It is based on the use of pumps that work at high operational pressures (30-600 atm) to generate a flow rate of liquid solvent that contains the sample through a column filled with a solid absorbent material also called the stationary phase. There is a detector incorporated to identify the different components as they elute from the column.

The chromatographic column used to purify Bca3 and Ppil3 was GSTrap™ 4B (General Electric Healthcare, Sweden). The process is based on an affinity chromatography that has glutathione as a ligand. This column is convenient for one-step purification of GST tagged proteins like Bca3 and Ppil3. The proteins are eluted under nondenaturing conditions using reduced glutathione.

The settings chosen for the chromatographic process were according the trademark's recommendations for this column. The maximum pressure was set to 0.5 mPa, the flow was set to 1 mL/min for elution and for equilibration to 0.2 mL/min during the injection. One pump contained the equilibration and loading buffer (PBS (NaCl 137 mM, KH_2PO_4 1.8 mM, KCl 2.7 mM and Na_2HPO_4 10 mM)). The other pump contained the elution buffer (Tris 50 mM pH 8.0, reduced glutathione 20 mM, pH 8.0). These buffers were previously filtered through filters of 0.22 mm of pore diameter and degassed. The column was washed with 5 column volumes of degassed bidistilled water and equilibrated with PBS buffer. Finally, the sample was eluted with 10 column volumes of elution buffer. GST-chimeric proteins eluted as a separate peak from this chromatography and were pure according to SDS-PAGE.

c) Dialysis

Purified samples were dialyzed before its lyophilization. Dialysis is a method used to eliminate unwanted compounds in protein and nucleic acid samples such as reducing agents, non-reacted crosslinking molecules, salts, labeling reagents or preservatives by passive diffusion. The sample was introduced into previously hydrated dialysis sacks and dialyzed against 2,000 mL of bidistilled water under agitation at 4ºC for approximately two days. The beaker containing the bidistilled water was changed approximately every 12 hours.

d) Freeze-drying or lyophilization

A freeze-drying or lyophilization is a process that eliminates the water from a sample under conditions of low temperature and vacuum. Dehydrated samples are more stable and this process avoids loss of activity or other possible damage in heat-sensitive molecules such as proteins.

Samples were covered with a punctured Parafilm^{TM} (Sigma-Aldrich, USA) and frozen at -80ºC. Then, they were lyophilized. The proteins obtained were weighed in order to determine the purification yield. The proteins were stored at -20ºC until their use.

Purification of H_6-ApopΔProΔLeu variant

a) Lysis and urea solubilization

H_6-ApopΔProΔLeu pellets were lysed three times in the French press (SLM-Aminco® Spectronic Instruments, UK) at 1100 psi and centrifuged at 9,000 rpm for 30 minutes. The variant was located in the pellet of the lysed sample. These pellets were solubilized in 10 mL of 50 mM Tris-HCl, 500 mM NaCl, 7 M Urea pH 7.4. Reduced glutathione was added to a final concentration of 100 mM and the pH was adjusted to 7.4 by the addition of solid Tris base. The reduced and denatured sample was incubated at room temperature for 2 hours under N_2 atmosphere to allow its complete reduction. The remaining insoluble material was removed by centrifugation at 10,000 rpm and 4ºC for 30 minutes.

b) Affinity chromatography purification

The chimeric His-tagged solubilized denatured protein was purified through Nickel-nitrilotriacetic acid (Ni-NTA) agarose chromatography. Ni-NTA agarose is an affinity chromatographic matrix used for purifying recombinant proteins carrying a His tag. Histidine residues bind to the immobilized nickel ions with high specificity and affinity. The cleared supernatant was incubated overnight at room temperature and gentle agitation with 15 mL (7.5 mL bed volume) of Ni-NTA agarose (Qiagen, The Netherlands) previously equilibrated with the solution buffer (50 mM Tris-HCl, 500 mM NaCl, 7 M Urea pH 7.4).

After overnight incubation, the resin was loaded onto a column and washed with 90 mL of 50 mM Tris-HCl, 500 mM NaCl, 7 M Urea, 10 mM Imidazole, pH 7.4 to remove unbound proteins. The protein was eluted from the column with 90 mL of 50 mM Tris-HCl, 500 mM NaCl, 7 M Urea, 500 mM Imidazole pH 7.4 and 9 mL fractions were collected. All fractions were then analyzed by SDS-PAGE and those containing the desired protein were gathered.

The unfolded and reduced sample was diluted dropwise tenfold into 50 mM Tris-HCl, 200 mM NaCl, 100 mM L-Arg, 100 mM L-Glu, 1 mM oxidized glutathione, 1 mM reduced glutathione, 1 mM EDTA, pH 8.1 and further incubated for 48 hours at 4ºC to allow protein refolding.

c) Tangential flow filtration

The sample pH was adjusted to 5.0 with HCl and concentrated to 150 mL by tangential flow filtration (Millipore, USA). In tangential flow filtration the sample is applied at high pressure (15 psi) into a feed stream that passes parallel to the membrane face. This avoids the proteins to accumulate onto the filter which would reduce the capacity of the permeate to cross the membrane.

d) Dialysis and lyophilization

Sample still contained some impurities, which could be removed by molecular exclusion chromatography. To use this technique, samples had to be concentrated and therefore, they were dialyzed against bidistilled water as described before and lyophilized.

e) Molecular exclusion chromatography

Gel filtration is a type of chromatography that separates molecules according to their size. There is a stationary phase that consists of porous beads with a defined range of pore sizes. Hence, small proteins can access to the mobile phase inside the beads thus they diffuse more and elute last. On the contrary, big proteins are too large to fit inside the pores so they stay on the mobile phase outside the beads and consequently, elute first.

The device used to perform the gel filtration was the HPLC ÄKTA pure (General Electric Healthcare, Sweden) using the column Superdex^{TM} 75 10/300 GL (General Electric Healthcare, Sweden). The different parameters were defined according to the trademark's instructions for this column and the buffer used was 50 mM Tris-HCl, 150 mM NaCl and 1 mM EDTA at pH 7.4. The buffer was previously filtered through filters of 0.22 mm of pore diameter and degassed. The maximum pressure was set to 1.8 Mpa, and the flow was set to 0.5 mL/minute.

The peak corresponding to the purified H_6-ApopΔProΔLeu was dialyzed against bidistilled water and lyophilized. The protein was stored at -20ºC until its use.

6. Protein-protein interaction analysis

a) Pull-down assay

Pull-down assay is an *in vitro* method used for both confirming the existence of a predicted protein-protein interaction and as an initial screening assay for identifying previously unknown protein-protein interactions. In this project it has been used to investigate the level of interaction between the proteins Bca3 and Ppil3 with H_6-ApoptinΔPro variant. We used this variant instead of H_6-ApoptinΔProΔLeu variant because the purification yield obtained was too low to proceed.

In this assay, a protein that is used as bait is captured on a fixed affinity ligand specific for a tag and is incubated with a protein source that contains prey proteins. In this instance, the bait

protein was H_6-ApoptinΔPro and it was captured in a Ni-NTA agarose through its His tag. The prey proteins were Bca3 and Ppil3.

GST-Bca3, GST-Ppil3 and H_6-ApopΔPro variants were prepared at 1 mg/ml in PBS and filtered through filters of 0.22 mm of pore diameter. The amount of protein was quantified measuring absorbance at 280 nm on the spectrophotometer NanoDrop ND-1000 (NanoDropTechnologies, USA). Then, 250 µl of GST-Bca3 or GST-Ppil3 solutions were mixed with either 250 µl of H_6-ApopΔPro variant or PBS. The mixes were incubated 90 minutes at room temperature. Meanwhile, 2 mL of HisTrap suspension resin (50% suspension) were equilibrated with PBS and resuspended in 1 mL of PBS. The resulting resin was equally distributed (¼) in each sample mix and incubated for 1 hour 30 minutes at room temperature under gentle agitation.

After this time, each mix was added in a column. The columns were washed with 5 mL of PBS 20 mM imidazole and samples of 0.5 mL were gathered until de DO_{280} dropped to zero. The 5 mL were added sequentially for an optimal washing. After that, proteins were eluted by the addition of 4 mL of PBS 500 mM imidazole in each column and samples of 0.5 mL were gathered. The 4 mL were added sequentially for an optimal elution. Absorbance at 280 nm was measured of each sample and those with a meaningful D.O. value were analysed through SDS-PAGE, dialyzed and lyophilized.

7. Ethical and sustainable considerations

All the information included in this project has been correctly cited and therefore no copyright has been committed. Furthermore, all the information has been obtained from reliable sources of information such as trademarks' websites, books, thesis, scientific papers, scientific databases, etc.

In the course of this project it has always been kept in mind to use the least amount of materials and substances as possible. However, some parts of this project required to work on sterile conditions which means that most of the material was of single-use. Nonetheless, all kind of materials have been correctly separated and distributed in the appropriate containers. In the case of biologic material, it has been previously submitted to a sterilization process.

The ethical implications of this project have also been taken into account. Such implications consist mainly of bacterial cloning. The bacterial cloning has been performed for research purposes at all times. Furthermore, the experimental conditions were safe for both the environment and the members working in the laboratory.

RESULTS AND DISCUSSION

1. Selection of putative *in vitro* binding partners of Apoptin

The main purpose of this Final Grade Project has been to select different genes whose encoded proteins are able to interact with Apoptin, to clone them in an expression vector and to produce and purify them. This is part of a broader objective consisting on the characterization through nuclear magnetic resonance of the structural changes produced in Apoptin upon binding to proteins that are relevant for its cytotoxic mechanism.

Table 2 depicts a list of protein candidates that have been previously described to interact or colocalize *in vivo* with Apoptin.

Table 2. Apoptin interaction partners with their biological function (adapted from Grimm & Noteborn, 2010).

Protein	Biological function
PKCβ	Interaction of PKCβ with Apoptin results in PKCβ-dependent phosphorylation, activation of Apoptin and translocation of both proteins to the nucleus.
FADD	Overexpression causes co-localization of FADD and Apoptin in cytoplasmic death effector filaments that could interfere with death receptor signaling.
Bcl-10	Apoptin and Bcl-10 co-localize in cytoplasmic filaments but the consequences remain unknown.
PKG-I	High expression levels of PKG-I in normal cells seem to have a protective effect. However, the mechanism is yet unknown.
PI3K and Akt	Apoptin associates with PI3K by binding through its PRS to the SH3 domain of p85. This leads to sustained activation and nuclear translocation of Akt which triggers activation of CDK2 that phosporylates Apoptin.
Hsc70	Apoptin expression triggers translocation of Hsc70 from the cytoplasm to the nucleus and Hsc70 is required for Apoptin-induced down-regulation of p65.
PML	Apoptin is sumoylated and targeted to nuclear PML bodies. Nevertheless, these modifications do not seem to be required for the activation of Apoptin.
APC1	Overexpression of Apoptin and its association with APC1 leads to a G2/M arrest and apoptosis.
Hippi	In normal cells Apoptin and Hippi interact and co-localize in the cytoplasm. On the other hand, in cancer cells Apoptin is translocated to the nucleus while Hippi remained in the cytoplasm.
DEDAF	Apoptin interacts with DEDAF in tumor cells in the nucleus and an overexpression of DEDAF increases the rate of cell death induced by Apoptin.
Ppil3	Overexpression of Ppil3 results in increased retention of Apoptin in the cytoplasm.
Nmi	Nmi is an Apoptin binding partner but their interaction has not been studied.
Bca3	Overexpression of Bca3 has been shown to enhance Apoptin phosphorylation and increase apoptosis induction by an unknown mechanism.

We discarded those proteins whose interaction with Apoptin was only indicated trough cell co-localization, as is the case of FADD, DEDAF and Bcl-10, since this co-localization could be caused by the common interaction with another protein. We also discarded those proteins

that introduced covalent modifications into Apoptin (phosphorylation in the case of PKCβ or sumoylation in the case of PML) since in this instance it would be difficult to understand whether the structural changes observed in Apoptin were due to binding to the protein or to the covalent modification. PI3K was also discarded because it binds to the Pro rich N-terminal region which has been removed in the Apoptin variant used in this study.

From the rest of the proteins depicted in Table 2 we selected those having the lower molecular weight since they would be easily constructed as synthetic genes, that is, Bca3 and Ppil3. These two genes were designed and purchased as synthetic genes.

2. Design of the synthetic genes of Bca3 and Ppil3

For the design of the genes of Bca3 and Ppil3 it was taken into account that these genes had to be cloned in the pGEX-4T-2 vector. This vector contains the coding sequence of Gluthatione S-transferase (GST) that can be used to create fusion proteins in order to help on the purification and stabilization of the resulting recombinant proteins. For the gene design, the amino acid sequence of both proteins (Figures 4 and 5) were back-translated to DNA sequences using the codons more represented in the codon usage of *E. coli*. The "TAA" stop codon was also added because it is the most used in *E.coli.* Then, a restriction map of each gene was performed *in silico* in order to search for restriction sites that were not present in the designed sequence of both genes but were present in the polylinker of pGEX-4T-2 expression vector (see Figure 2). The restriction sites of *Bam*HI and *Xho*I were added at the 5' and 3' ends of each gene to facilitate the cloning process. The use of two different restriction sites avoided the formation of plasmid self-ligation in the subsequent ligation reaction. Finally, a linker coding for the "LVPRGS" hexapeptide was placed between GST and the genes of Bca3 or Ppil. The resulting designed sequences of Bca3 and Ppil3 genes are indicated in Figures 6 and 7 with a length of 645 bp and 498 bp respectively. These synthetic genes were purchased cloned into a plasmid to *Synthetic Genomics DNA and Integrated DNA Technologies* (USA) and these plasmids were named GST-Bca3 and GST-Ppil3.

MDNCLAAAALNGVDRRSLQRSARLALEVLERAKRRAVDWHALERPKGCMGVLAREAPHLEKQPAAGPQR
VLPGEREERPPTLSASFRTMAEFMDYTSSQCGKYYSSVPEEGGATHVYRYHRGESKLHMCLDIGNGQRKDR
KKTSLGPGGSYQISEHAPEASQPAENISKDLYIEVYPGTYSVTVGSNDLTKKTHVVAVDSGQSVDLVFPV

Figure 4. Amino acid sequence of Bca3.

MSVTLHTDVGDIKIEVFCERTPKTCENFLALCASNYYNGCIFHRNIKGFMVQTGDPTGTGRGGNSIWGKK
FEDEYSEYLKHNVRGVVSMANNGPNTNGSQFFITYGKQPHLDMKYTVFGKVIDGLETLDELEKLPVNEKT
YRPLNDVHIKDITIHANPFAQ

Figure 5. Amino acid sequence of Ppil3.

GGATCCATGGATAACTGCCTGGCGGCGGCGGCGCTGAACGGCGTGGATCGCCGCAGCCTGCAGCGCA
GCGCGCGCCTGGCGCTGGAAGTGCTGGAACGCGCGAAACGCCGCGCGGTGGATTGGCATGCGCTGGA
ACGCCCGAAAGGCTGCATGGGCGTGCTGGCGCGCGAAGCGCCGCATCTGGAAAAACAGCCGGCGGCG
GGCCCGCAGCGCGTGCTGCCGGGCGAACGCGAAGAACGCCCGCCGACCCTGAGCGCGAGCTTTCGCA
CCATGGCGGAATTTATGGATTATACCAGCAGCCAGTGCGGCAAATATTATAGCAGCGTGCCGGAAGAA
GGCGGCGCGACCCATGTGTATCGCTATCATCGCGGCGAAAGCAAACTGCATATGTGCCTGGATATTGG
CAACGGCCAGCGCAAAGATCGCAAAAAAACCAGCCTGGGCCCGGGCGGCAGCTATCAGATTAGCGAA

CATGCGCCGGAAGCGAGCCAGCCGGCGGAAAACATTAGCAAAGATCTGTATATTGAAGTGTATCCGG
GCACCTATAGCGTGACCGTGGGCAGCAACGATCTGACCAAAAAAACCCATGTGGTGGCGGTGGATAG
CGGCCAGAGCGTGGATCTGGTGTTTCCGGTGTAA**CTCGAG**

Figure 6. Designed DNA sequence of GST-Bca3. The sequences in bold indicate the position of the restriction sites of *Bam*HI (5') and *Xho*I (3') and the sequence shaded in grey indicates the stop codon.

GGATCCATGAGCGTGACCCTGCATACCGATGTGGGCGATATTAAAATTGAAGTGTTTTGCGAACGCACC
CCGAAAACCTGCGAAAACTTTCTGGCGCTGTGCGCGAGCAACTATTATAACGGCTGCATTTTTCATCGC
AACATTAAAGGCTTTATGGTGCAGACCGGCGATCCGACCGGCACCGGCCGCGGCGGCAACAGCATTTG
GGGCAAAAAATTTGAAGATGAATATAGCGAATATCTGAAACATAACGTGCGCGGCGTGGTGAGCATG
GCGAACAACGGCCCGAACACCAACGGCAGCCAGTTTTTTATTACCTATGGCAAACAGCCGCATCTGGAT
ATGAAATATACCGTGTTTGGCAAAGTGATTGATGGCCTGGAAACCCTGGATGAACTGGAAAAACTGCC
GGTGAACGAAAAAACCTATCGCCCGCTGAACGATGTGCATATTAAAGATATTACCATTCATGCGAACCC
GTTTGCGCAGTAA**CTCGAG**

Figure 7. Designed DNA sequence of GST-Ppil3. The sequences in bold indicated the restriction sites of *Bam*HI (5') and *Xho*I (3') and the sequence shaded in grey indicates the stop codon.

3. Construction of the fusion proteins between Bca3 or Ppil3 and GST

A diagram showing the strategy of construction of both chimeric proteins is shown in Figure 8. Plasmids GST-Bca3, GST-Ppil3 and pGEX-4T-2 were purified using the miniprep method and digested with *Bam*HI. Linearized plasmids were purified using the Geneclean method and further digested with *Xho*I. An example of the resulting digestion can be observed in the agarose gel electrophoresis depicted in Figure 9. Uncut vectors show two bands in these gels because they have two types of conformations: the supercoiled and the relaxed forms. pGEX-4T-2 digested with both restriction enzymes and the Bca3 and Ppil3 genes excised from plasmids GST-Bca3 and GST-Ppil3 were purified using the Geneclean method and subsequently ligated using T4 DNA ligase (Roche, Germany).

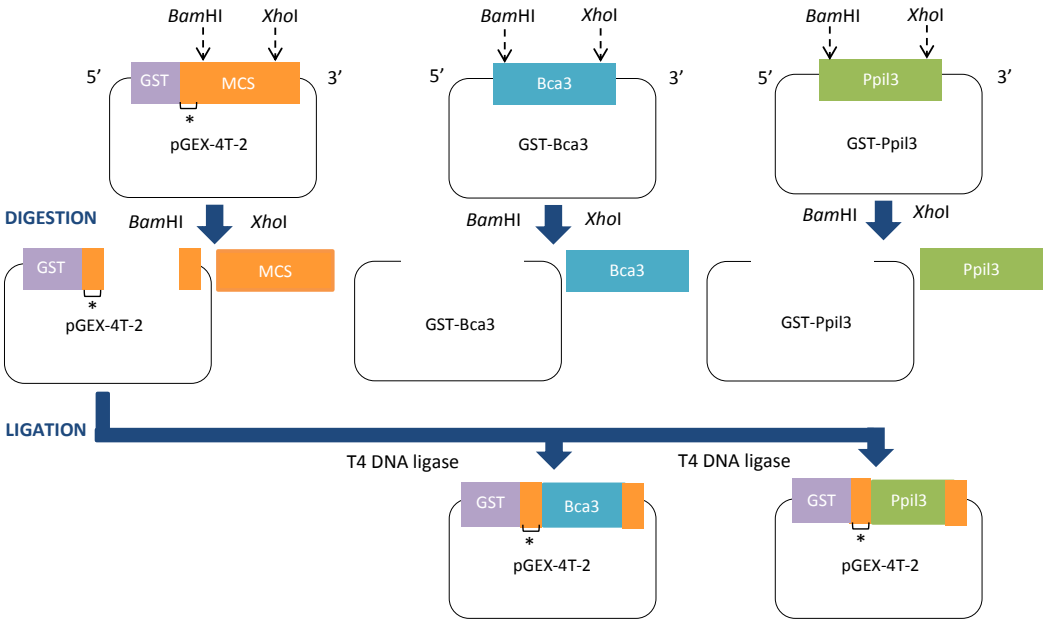

Figure 8. Diagram showing the digestion with the restriction enzymes *Bam*HI and *Xho*I of the vector pGEX-4T-2, the plasmids GST-Bca3 and GST-Ppil3 in conjunction with the ligation of the digested vector with the inserts, leading to the final variant constructions. The asterisk indicates the sequence that codes for the Thrombin cleavage.

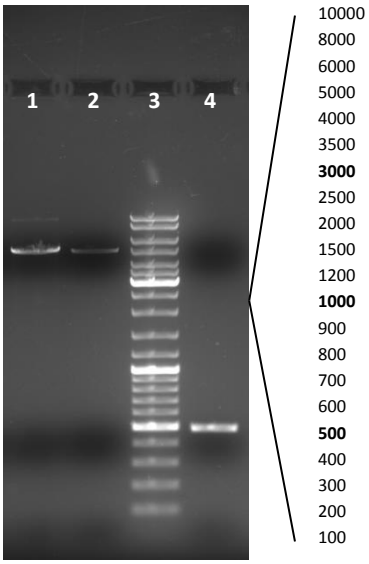

Figure 9. Agarose gel electrophoresis of different DNA fragments. **1 –** Uncut pGEX-4T-2 vector. **2 –** pGEX-4T-2 vector digested with *Bam*HI and *Xho*I. **3 –** Ppil3 digested with *Bam*HI and *Xho*I. **4 –** Molecular weight DNA marker expressed in base pairs (DNA ladder, Fermentas, Lituania). The base pairs of each band are indicated aside.

Competent cells were transformed with the products of the ligations (see section 2.d). Religation controls, in which the digested genes had not been added, were included to detect the completeness of the pGEX-4T-2 digestion. The reverse and complement chains of the recombinant genes from different clones were analyzed by sequencing in order to select a clone that did not present any mutation. Then, these sequences were aligned with the expected sequence of each insert using *Multiple Sequence Alignment (MultAlin)* to detect whether the sequences were fully correct (Figures 10 and 11).

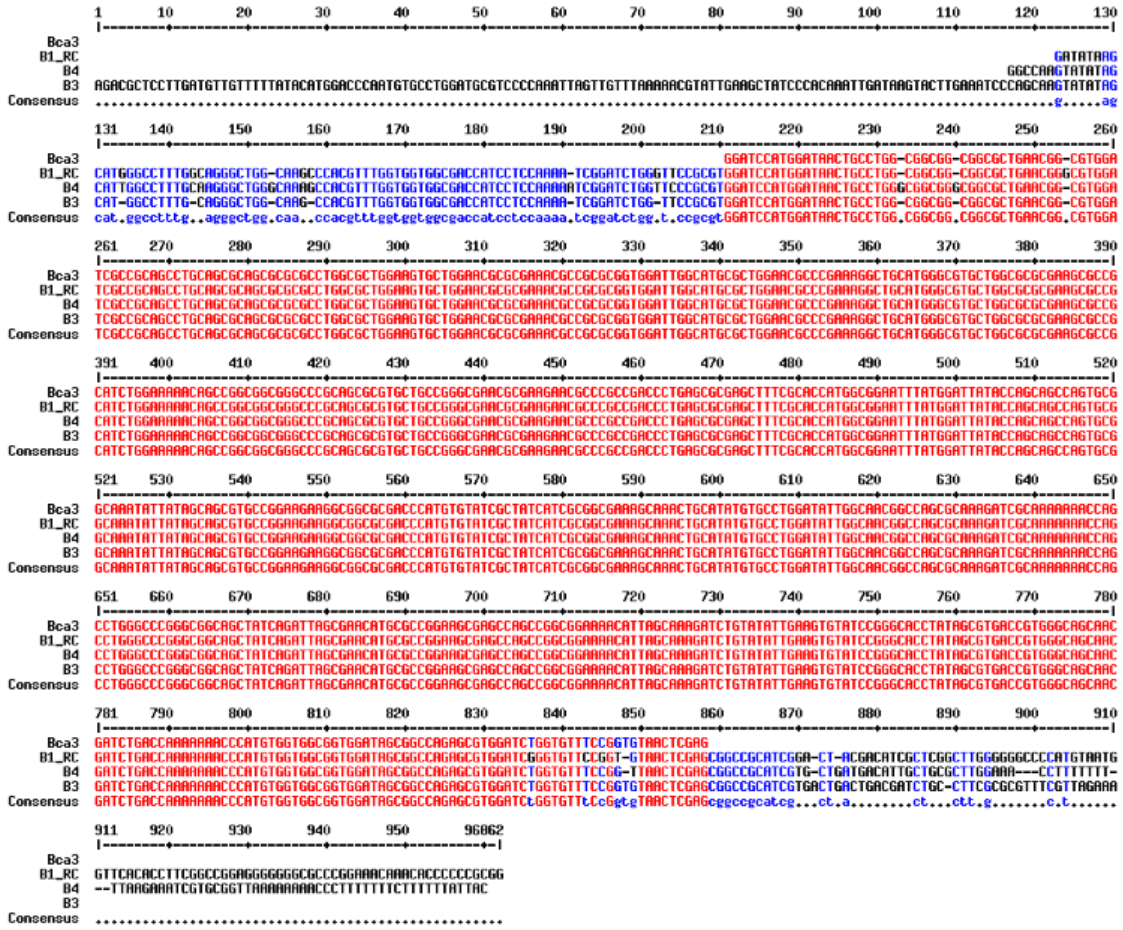

Figure 10. Alignments of the sequences obtained from the different clones (B1_RC, B4 and B3) with Bca3 expected sequence (Bca3) obtained with *Multiple Sequence Alignment (MultAlin)*. The sequence of the expected sequence (Bca3), that of the investigated clones (B1_RC, B4 and B3) and the consensus sequence are indicated. Positions in red indicate that all the sequences are identical at these positions whereas those positions in blue represent changes in the sequence.

In the alignments we can see that clone number 1 (B1_RC) has four mutations: at the positions 836 (G instead of T), 843 (C instead of T), 845 (G instead of C), 847 (T instead of G). Clone number 4 has one mutation at the position 849 (T instead of G). Clone number 3 (B3) does not present any mutation (see Figure 10). Therefore, we chose this one to carry out the recombinant production of Bca3.

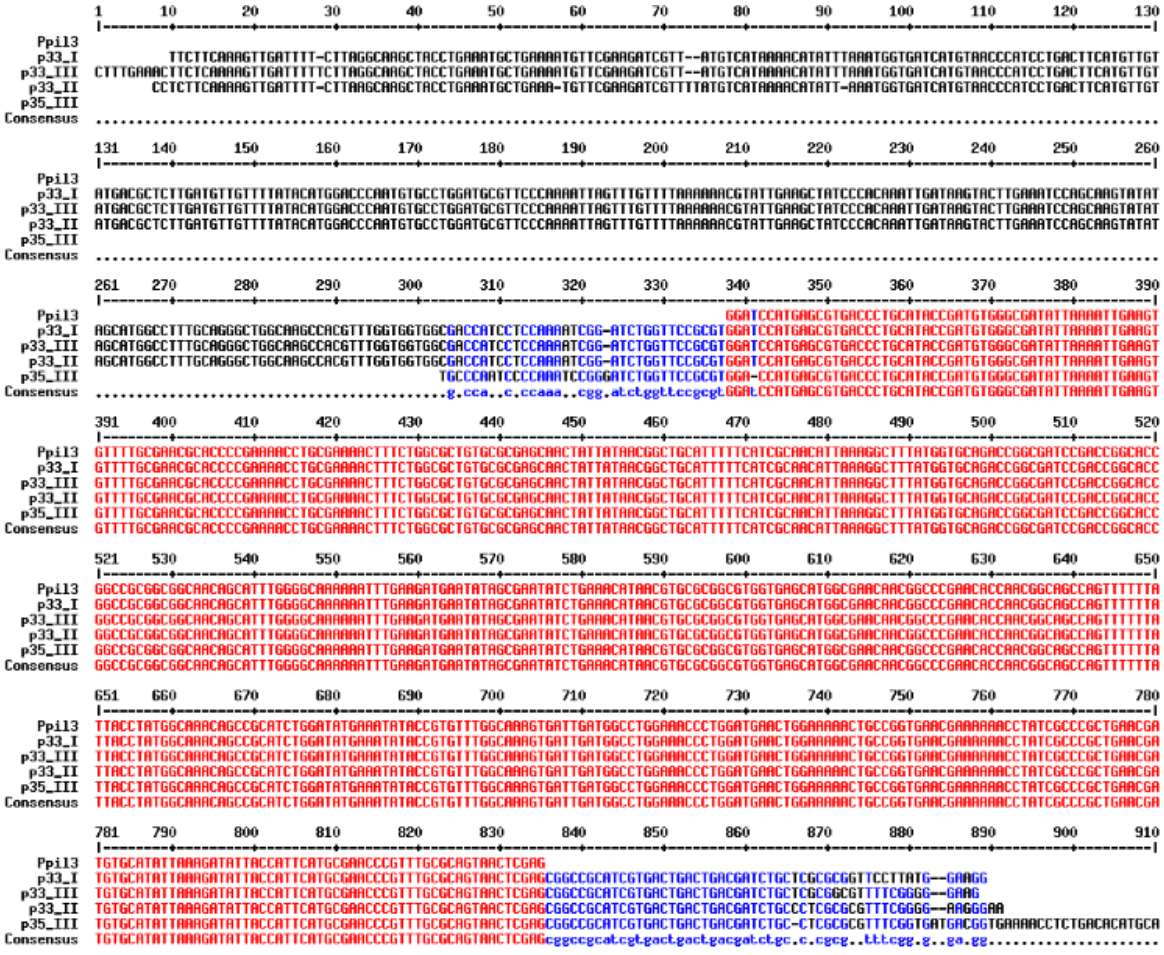

Figure 11. Alignments of the sequences obtained from the different clones (p33_I, p33_III, p33_II and p35_III) with Ppil3 expected sequence (Ppil3) obtained with *Multiple Sequence Alignment (MultAlin)*. The sequence of the expected sequence (Ppil3), that of the investigated clones (p33_I, p33_III, p33_II and p35_III) and the consensus sequence are indicated. Positions in red indicate that all the sequences are identical at these positions whereas those positions in blue represent changes in the sequence.

Clone number 1 (p33_I) and clone number 2 (p33_II) do not have any mutation. Clone number 3 (p33_III) does not present any mutation on the sequence amplified with 3' primer but it has a mutation at the position 341 (A instead of T) when it is amplified with 5' primer (see Figure 11). Hence, clone number 1 was the one used for the subsequent production of Ppil3.

4. Recombinant protein expression

a) Selection of the producing colony

Once the fusion proteins were obtained, we attempted the production and purification of these proteins. XL1blue competent cells were transformed and different colonies were screened in order to choose the most appropriate colony, that which produced a higher yield of recombinant protein. An example study with Bca3 is shown in Figure 12.

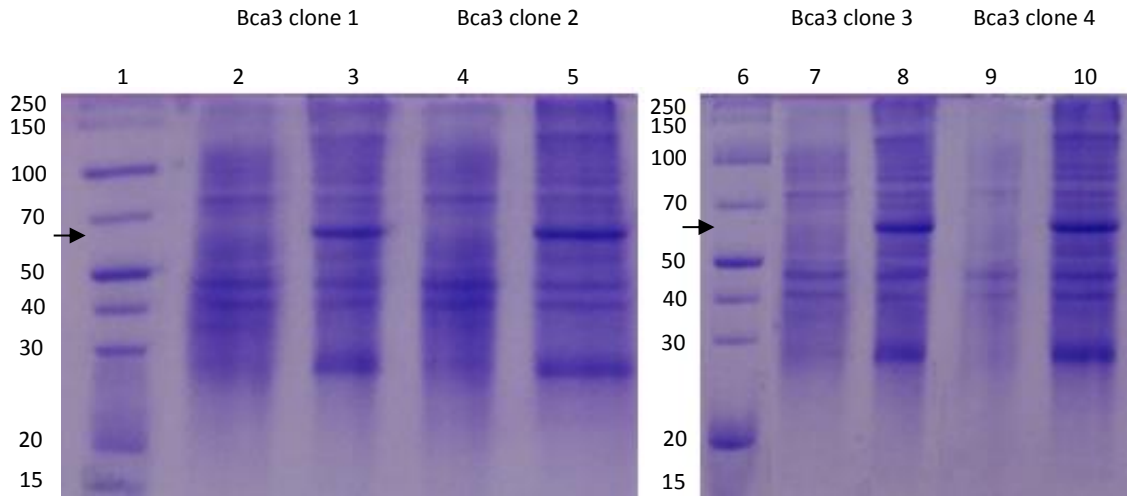

Figure 12. SDS-PAGE showing the recombinant protein yield of different colonies transformed with clone number 3 (B3) and induced for 3 hours with 1 mM IPTG. **1 –** Molecular weight marker (PageRuler™ Prestained Protein Ladder, ThermoFisher Scientific, USA). For each colony, a sample of uninduced (left) and induced (right) cell extract is analyzed. The arrow indicates the position of Bca3 in the gel. The molecular weight of each marker, expressed in kDa, is indicated.

Bca3 has a molecular weight of 51.07 kDa. In this case, all the colonies expressed high levels of Bca3 but we chose number 3.

b) Production and purification

Two L of culture were induced for the production and purification of the proteins (Bca3, Ppil3 and H6-ApopΔProΔLeu). The induced cultures were centrifuged and lysed with the French Press. They were centrifuged to separate the soluble fraction from inclusion bodies.

Soluble fractions from Bca3 and Ppil3 were purified through GSTrap affinity chromatography and the sole elution peaks (Figures 13 and 14) were gathered and analyzed through SDS-PAGE together with the other fractions of the purification process. This analysis showed that Bca3 was retained in the insoluble fraction likely because it is produced in form of inclusion bodies (Figure 15). The peak eluted from the GSTrap (lane 5, Figure 15) had the same molecular weight as GST. It can be postulated that part of the recombinant protein is cleaved during the productions and this GST fraction remained in the soluble fraction and was purified. Hence, purification process of Bca3 should include solubilization and refolding steps that should be set up. Regarding Ppil3, SDS-PAGE analysis confirmed that the recombinant protein was retained in the soluble fraction (Figure 16) and could be purified. The purification yield obtained was of 2 mg per liter of induced culture. The purified protein was dialyzed at 4ºC against 2 L of bidistilled water and lyophilized.

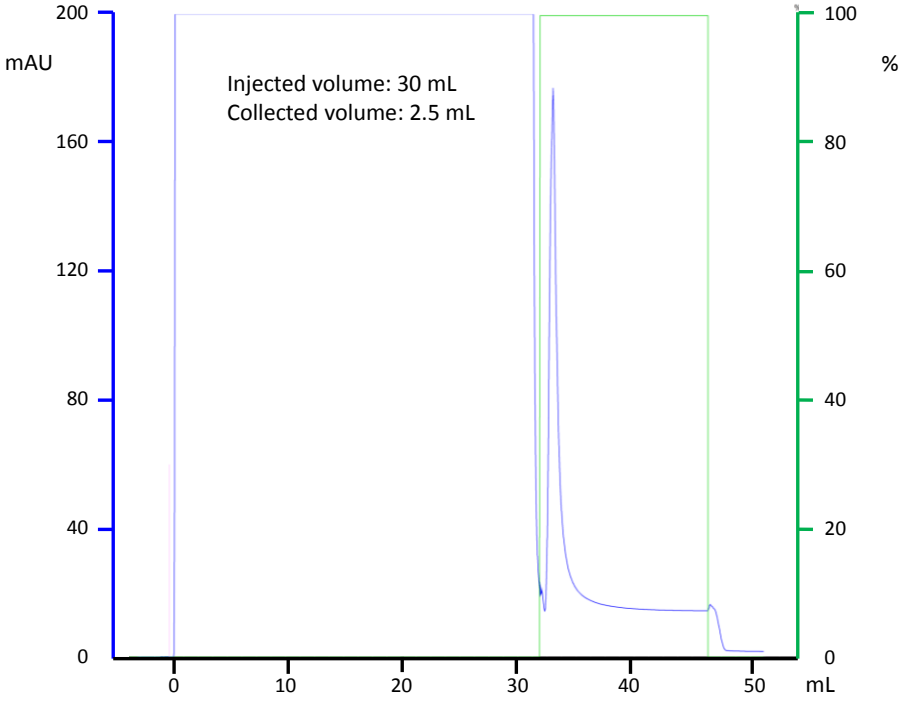

Figure 13. Chromatographic profile of Bca3 (blue line) obtained with the GSTrap affinity chromatography. After the injection of 30 mL of concentrated sample the column was washed until the D.O. was below 0.02 and the protein was eluted with 20 mM reduced glutathione. Line green sharply increases to 100% when the elution buffer is injected into the column. Chromatogram extracted from *Unicorn 6.4.1* software.

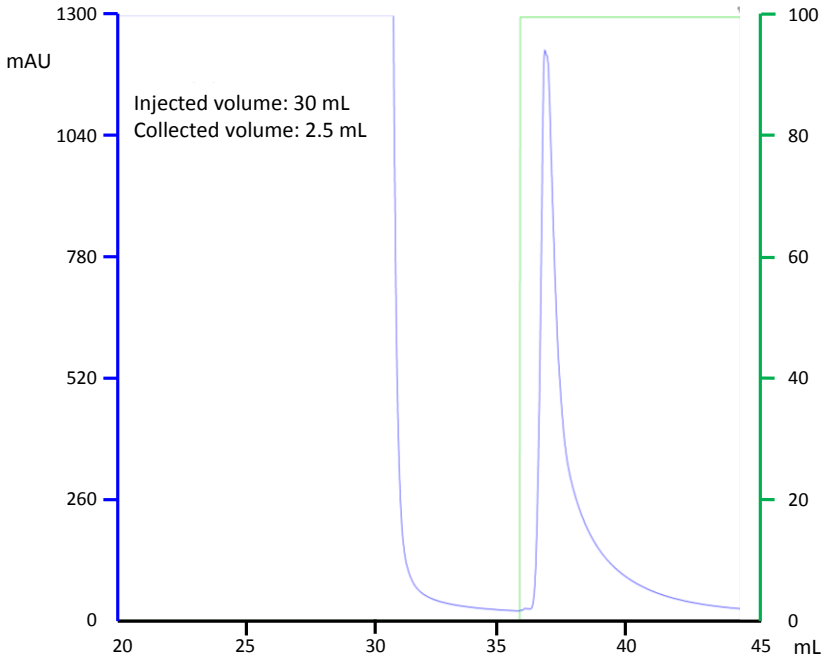

Figure 14. Chromatographic profile of Ppil3 (blue line) obtained with the GSTrap affinity chromatography. After the injection of 30 mL of concentrated sample the column was washed until the D.O. was below 0.02 and the protein was eluted with 20 mM reduced glutathione. Line green sharply increases to 100% when the elution buffer is injected into the column. Chromatogram extracted from *Unicorn 6.4.1* software.

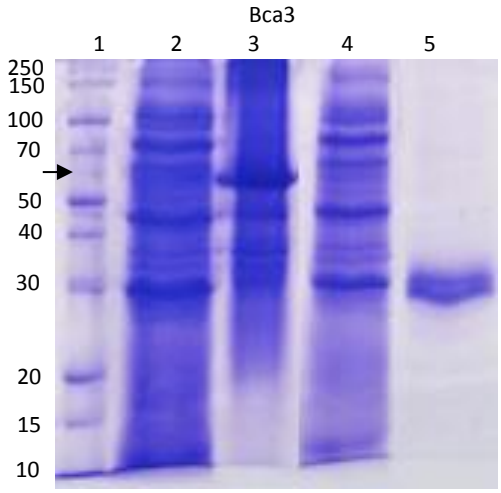

Figure 15. SDS-PAGE analysis of the different steps of the purification of Bca3. **1 –** Molecular weight marker (PageRuler^{TM} Prestained Protein Ladder, ThermoFisher Scientific, USA). **2 –** French Press soluble fraction. **3 –** French Press insoluble fraction. **4 –** Flow through obtained from HPLC. **5 –** Bca3 peak eluted with HPLC. The arrow points Bca3 molecular weight (51.07) that is expressed in kDa.

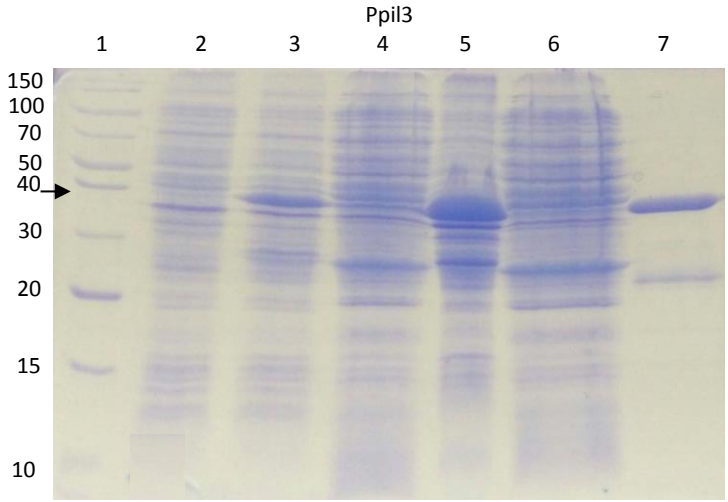

Figure 16. SDS-PAGE analysis of the purification of Ppil3. **1 –** Molecular weight marker (PageRuler^{TM} Prestained Protein Ladder, ThermoFisher Scientific, USA). **2 –** Uninduced cell extract. **3 –** Induced cell extract. **4 –** French Press insoluble fraction. **5 –** French Press soluble fraction. **6 –** Flow through from HPLC. **7 –** Ppil3 peak eluted with HPLC. The arrow indicates Ppil3 molecular weight (46.11) that is expressed in kDa.

Finally, regarding the H_6-ApoptinΔProΔLeu variant, it was also retained in the insoluble fraction, likely in form of inclusion bodies (Figure 17). The protein was solubilized with a buffer containing 7 M Urea to completely denature the aggregates and 0.1 M reduced glutathione to reduce all the possible intermolecular and intramolecular disulfide bonds. After purification of solubilized H_6-ApoptinΔProΔLeu through Ni-NTA affinity chromatography, the sample was refolded by removing through dilution the denaturant and allowing the proper disulfide bonds by the addition a the oxidized glutathione as a redox pair (see section 5b. H_6-ApoptinΔProΔLeu variant). The 1 L sample had to be further purified through molecular-exclusion chromatography which required the use of lower volumes (3-4 mL). The sample was

concentrated to 150 mL through tangential flow filtration and then, it was dialyzed against bidistilled water and lyophilized. Once lyophilized, it was resuspended with 1.5 mL of a buffer containing 50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 7.4 and purified on a Superdex-75 10/300GL (General Electric Healthcare, UK) column. H_6-ApoptinΔProΔLeu variant eluted at 11.30 mL (Figure 18). The purification yield obtained was too low to be calculated. Therefore, new attempts should be made in order to obtain higher purification yields.

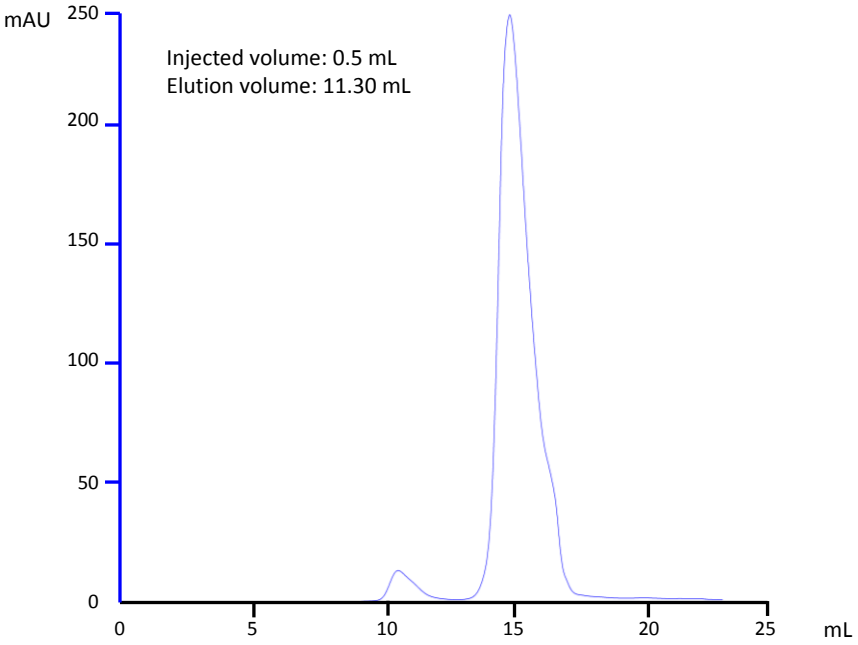

Figure 17. Chromatographic profile of H6-ApopΔProΔLeu variant obtained with the gel filtration column. The volume injected was 0.5 mL and the protein eluted at 11.30 mL. The chromatogram is extracted from *Unicorn 6.4.1* software.

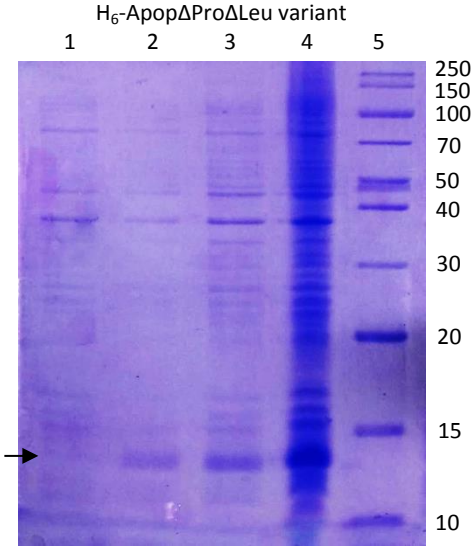

Figure 18. SDS-PAGE analysis of the purification process of H_6-ApopΔProΔLeu. **1 –** Cell extract of uninduced culture. **2 –** Cell extract of induced culture. **3 –** French Press supernatant. **4 –** French Press pellet. **5 –** Molecular weight marker (PageRuler^{TM} Prestained Protein Ladder, ThermoFisher Scientific, USA). The arrow indicates H_6-ApopΔProΔLeu molecular weight (10.8) that is expressed in kDa.

5. Pull-down assay

A pull-down assay was performed in order to investigate the *in vitro* interaction of Ppil3 with H_6-ApopΔPro variant. Since we had not enough H_6-ApoptinΔProΔLeu protein, to carry out this experiment we used another variant of Apoptin lacking residues 1 to 28 that forms soluble aggregates (H_6-ApoptinΔPro). This variant was used as bait and as prey protein we used Ppil3. 1 mg/mL solution of GST-Ppil3 was mixed with either 1 mg/mL solution of H_6-ApopΔPro variant or with PBS. Mixes were loaded to an Ni-NTA agarose which allowed the Apoptin variant to be captured through its His tag. Each chromatographic column was washed and eluted with imidazole. Samples were dialyzed and lyophilized to concentrate them. The presence of both baits together with the Apoptin prey was investigated through SDS-PAGE but no band of protein could be observed. We concluded that the assay had not been properly implemented as the results obtained. A possible explanation is that H_6-ApopΔPro variant has tendency to aggregate and in this case, the His tag of this variant could not be exposed and then, the protein would not interact with the Ni-NTA-Agarose.

CONCLUSIONS

According to the objectives of this project, the following conclusions can be drawn:

- We have designed and constructed the genes coding for the proteins Breast cancer associated gene 3 (Bca3) and Peptidyl-prolyl isomerase-like 3 (Ppil3) fused to the gene of Glutathione S-transferase (GST) in the expression vector pGEX-4T-2.
- We have successfully produced and set up a method for purification of GST-Ppil3 obtaining a yield of 2 mg /L of induced culture.
- GST-Bca3 is produced as an insoluble protein which cannot be directly purified through an affinity chromatography. To purify this protein a new protocol which must include solubilization and refolding steps should be stablished.
- The pull-down analysis of the interaction of GST-Ppil3 with the H_6-ApopΔPro variant was not carried out successfully. Further studies will be required to optimize the pull-down assay protocol.

REFERENCES

Backendorf, C., Visser, A., de Boer, A., Zimmerman, R., Visser, M., Voskamp, P., Zhang, Y., & Noteborn, M. (2008). Apoptin: therapeutic potential of an early sensor of carcinogenic transformation. *Annual Review of Pharmacology and Toxicology*, *48*, 143–169.

Bergmann, M., Romirer, I., Sachet, M., Fleischhacker, R., García-Sastre, A., Palese, P., Wolff, K., Pehamberger, H., Jakesz, R., & Muster, T. (2001). A genetically engineered influenza A virus with ras-dependent oncolytic properties. *Cancer Research*, *61* (22), 8188–8193.

Burek, M., Maddika, S., Burek, C., Daniel, P., Schulze-Osthoff, K., & Los, M. (2006). Apoptin-induced cell death is modulated by Bcl-2 family members and is Apaf-1 dependent. *Oncogene*, *25* (15), 2213–2222.

Danen-Van Oorschot, A., Zhang, Y., Leliveld, S., Rohn, J., Seelen, M., Bolk, M., Van Zon, A., Erkeland, S., Abrahams, J., Mumberg, D., & Noteborn, M. (2003). Importance of nuclear localization of apoptin for tumor-specific induction of apoptosis. *The Journal of Biological Chemistry*, *278* (30), 27729–27736.

Dyson, H., & Wright, P. (2005). Intrinsically unstructured proteins and their functions. *Nature Reviews. Molecular Cell Biology*, *6* (3), 197–208.

Flotte, T., Carter, B., Conrad, C., Guggino, W., Reynolds, T., Rosenstein, B., Taylor, G., Walden, S., & Wetzel, R. (1996). A phase I study of an adeno-associated virus-CFTR gene vector in adult CF patients with mild lung disease. *Human Gene Therapy*, *7* (9), 1145–1159.

Grimm, S., & Noteborn, M. (2010). Signalling of Apoptin. In *Anticancer Genes: inducers of tumour-specific cell death signalling* (pp. 23–33). Philadelphia: Springer.

Harada, J., & Berk, A. (1999) p53-Independent and –dependent requirements for E1B-55K in adenovirus type 5 replication. *Journal of Virology,* 73, 5333-5344.

Jeurissen, S., Wagenaar, F., Pol, J., van der Eb, A., & Noteborn, M. (1992). Chicken anemia virus causes apoptosis of thymocytes after in vivo infection and of cell lines after in vitro infection. *Journal of Virology*, *66* (12), 7383–7388.

Ledford, H. (2015). Cancer-fighting viruses win approval. *Nature*, *526* (7575), 622–623.

Leliveld, S., Zhang, Y., Rohn, J., Noteborn, M., & Abrahams, J. (2003). Apoptin induces tumor-specific apoptosis as a globular multimer. *The Journal of Biological Chemistry*, *278* (11), 9042–9051.

Mantovani, A. (2009). Cancer: Inflaming metastasis. *Nature*, *457*, 36–37.

Moore, S., Knudsen, B., True, L., Hawley, S., Etzioni, R., Wade, C., Gifford, D., Coleman, I., & Nelson, P. (2005). Loss of stearoyl-CoA desaturase expression is a frequent event in prostate carcinoma. *International Journal of Cancer*, *114* (4), 563–571.

Noteborn, M., Todd, D., Verschueren, C., de Gauw, H., Curran, W., Veldkamp, S., Douglas, A., McNulty, M., van der, E., & Koch, G. (1994). A single chicken anemia virus protein induces apoptosis. *Journal of Virology*, *68*(1), 346–351.

Oldfield, C., & Dunker, A. (2014). Intrinsically disordered proteins and intrinsically disordered protein regions. *Annual Review of Biochemistry*, *83*, 553–584.

Poon, I., Oro, C., Dias, M., Zhang, J., & Jans, D. (2005). A Tumor Cell-Specific Nuclear Targeting Signal within Chicken Anemia Virus VP3/Apoptin. *Journal of Virology*, *79* (2), 1339–1341.

Ramaswami, R., Harding, V., & Newsom-Davis, T. (2016). Novel Cancer Therapies. *Postgraduate Medical Journal*, *89* (1057), 652–658.

Ruiz Martínez, S. (2016). *Structural and functional characteriztion of Apoptin and design of a monomeric variant with antitumor activity*. Doctoral Thesis Universitat de Girona.

Ruiz-Martínez, S., Pantoja-Uceda, D., Castro, J., Vilanova, M., Ribó, M., Bruix, M., Benito, A., & Laurents, D. V. (2016a). A Cytotoxic Apoptin Variant is Intrinsically Disordered as Characterized by NMR Spectroscopy. Submitted

Ruiz-Martínez, S., Castro, J., Vilanova, M., Bruix, M., Laurents, D., Ribó, M., & Benito, A. (2016b). A truncated form of Apoptin is selectively cytotoxic for cancer cells when exogenously added. Submitted.

Seow, Y., & Wood, M. (2009). Biological gene delivery vehicles: beyond viral vectors. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, *17* (5), 767–777.

Tavassoli, M., Guelen, L., Luxon, B., & Gäken, J. (2005). Apoptin: specific killer of tumor cells? *Apoptosis*, *10* (4), 717–24.

Uversky, V., Oldfield, C., & Dunker, A. (2008). Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annual Review of Biophysics*, *37*, 215–246.

Ward, J., Sodhi, J., Mc Guffin, L., Buxton, B., & Jones, D. (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *Journal of Molecular Biology*, *337* (3), 635–645.

Zhang, Y., Kooistra, K., Pietersen, A., Rohn, J., & Noteborn, M. (2004). Activation of the tumor-specific death effector apoptin and its kinase by an N-terminal determinant of simian virus 40 large T antigen. *Journal of Virology*, *78* (18), 9965–9976.

Zhang, X., Xie, Q., Ji, J., Chang, S., Liu, J., Chen, F. Jingyun, M., & Bee, Y. (2012). Complete Genome Sequence Analysis of a Recent Chicken Anemia Virus Isolate and Comparison with a Chicken Anemia Virus Isolate from Human Fecal Samples in China. *Journal of Virology*, *86* (19), 10896–10897.