STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF APOPTIN AND DESIGN OF A MONOMERIC VARIANT WITH ANTITUMOR ACTIVITY

Santiago Ruiz Martínez

Universitat de Girona

DOCTORAL THESIS

Structural and functional characterization of Apoptin and design of a monomeric variant with antitumor activity

Santiago Ruiz Martínez

2016

Universitat de Girona

DOCTORAL THESIS

Structural and functional characterization of Apoptin and design of a monomeric variant with antitumor activity

Santiago Ruiz Martínez

2016

MOLECULAR BIOLOGY, BIOMEDICINE AND HEALTH

Directed by

Dr. Antoni Benito Mundet
Associate professor in Biochemistry and Molecular Biology

Directed by

Dr. Marc Ribó i Panosa
Associate professor in Biochemistry and Molecular Biology

This thesis is submitted in fulfilment of the requirements to obtain the doctoral degree from the Universitat de Girona

This PhD thesis contains one annex

This work has been supported by a predoctoral FPU fellowship and mobility grant from the MEC (Spain) (AP2010-4803) and by grants BFU2009-06935, BIO2013-43517 and SAF-2013-49179-C2-2-R from MINECO (Spain) and MPCU2016/18 and SING12/0 from UdG (Spain).

Ha arribat el moment dels agraïments, probablement la part més llegida de la tesi. Com que és la única part amb format lliure, em permetré el luxe de fer-ho en diferents idiomes, ja que m'agradaria que tots els que hi han format part d'una manera o altra, siguin capaços d'entendre-ho.

Començaré com no, pels que m'han donat l'oportunitat d'entrar al món de la ciència, a la Maria Vilanova, en Marc Ribó i en Toni Benito. No cal dir que sense ells aquesta tesi, de la que val a dir n'estic molt content, no hagués estat possible. Moltes gràcies per la vostra constant preocupació tant a nivell professional com a nivell personal.

Words are powerless to express my gratitude to Douglas who has always shown an enormous interest in my project and has always proved to be a good scientist and a better person. Moltes gràcies també a la Marta Bruix, qui m'ha cuidat durant la meva estada a Madrid, ajudant-me a entendre el tan complicat món de la RMN. Gracias a Angélica, Miguel Mompeán y David, que han hecho mi estancia en Madrid mucho más agradable.

Ara m'agradaria passar a agrair a tots aquells amb qui no només he compartit més temps durant el doctorat, sinó que també han pogut entendre millor que ningú tot el que fer una tesi comporta. Començaré per la Vero, a qui per sort vaig conèixer al principi d'entrar al grup i amb qui he pogut compartir bons moments tant a la universitat, amb aquestes converses "frikies", com a fora. A la Txell, amb qui he descobert que no només podem tenir controlada tota la universitat, sinó que és una persona amb qui he pogut parlar sempre que ho he necessitat. A mi "pistaxin" (Marta), por todos esos buenos momentos que hemos pasado juntos, por las lecciones culinarias y de la vida en general que harán de mí una persona de provecho. No podria deixar-me a la Jess, a qui potser més he molestat fent consultes, tant de feina com no, qui tot i anar sempre justa de temps (no sé com s'ho fa) m'ha dedicat tot el temps que he necessitat. Vull també agrair a la Montse, en David, la Glòria, a tots els de la "granja"

(la Mireia, la Sara, en Pablo, la Sandra, l'Elena, la Carla i en Pau), a l'Ari, l'Àlex, en Luis i en Pedro amb qui he compartit despatx, laboratori, dinars (al mig, davant o a dins), "Pacobirres", sopars i bons moments. A la Mariona per adoptar-me quan encara era "pupilo" i amb qui he compartit i espero continuar compartint molts bons moments. A l'Anna Vert, per tots els moments de reflexió i els "tocs" improvisats. A en Pere per la fantàstica portada d'aquesta tesi. A tots vosaltres moltes gràcies!

Quiero agradecer también el apoyo recibido por dos fantásticas personas que conocí durante la semana en Palma de Mallorca gracias al Cross Border Doctorials, Silvia y Robert, sois simplemente geniales.

My time in Berlin wouldn't have been the same without Conni, Niall, Bahar, Aleksandra, Joaquim and Albert. Thank you for all the good moments that we've spent there both during the YSF and the FEBS Congress and for those moments that are yet to come.

A mis compañeros de carrera que con el tiempo acabaron formando parte de mi vida, gracias Lili, Anna, Víctor (aunque no seas biólogo), Alba y Èlia. Por supuesto a Lidia Risco ("Manola") por todas esas experiencias vividas juntos.

I would like to acknowledge Mia, who I have always looked up to. Thank you for all your wise and useful advises. Thanks to all the people I met in Sweden (Maude, Charlene, Julia, Núria, Alex and Albert) and with whom I have built a great friendship.

Fuera de la universidad, debo un enorme gracias por todo el apoyo recibido desde el primer momento a Alana, con quien espero se cumpla el viaje prometido al final de la tesis. No puedo dejarme a Estefanía ("la Paco"), con quien a pesar de habernos conocido más tarde, he compartido mucho y seguiré compartiendo.

Quiero agradecer a Cristina (“la Solano”) por esa constante preocupación. También a Rosi, un enorme punto de apoyo tanto para mí, como para mi familia, muchas gracias.

Finalmente mi mayor gracias a mi familia, quien a pesar de no tener nunca claro si continuaba estudiando o ya trabajaba me ha apoyado incondicionalmente e independientemente de la decisión que yo tomara. En especial a mi madre y mi hermano, las dos personas que han hecho que a día de hoy yo sea quien soy, estando junto a mí en cada momento. A mi hermana, que a pesar de su corta edad siempre ha sabido sacarme una sonrisa en momentos difíciles y de quien ya me siento muy orgulloso. No quisiera acabar estos agradecimientos sin mencionar a mi primo Paco, quien aprendió la palabra “ribonucleasa” en mi defensa del TFM y espero pueda continuar ampliando su vocabulario en mi defensa de tesis.

ANS	8-(Phenylamino)-1-naphthalenesulfonic acid ($C_{16}H_{13}NO_3S$)
APC/C	anaphase-promoting complex/cyclosome
Bax	bcl-2-associated X protein
Bcl-2	B-cell lymphoma protein 2
BGH	Bovine Growth Hormone
BSA	Bovine serum albumin
CAV	Chicken Anemia Virus
CD	Circular Dichroism
CDF	cumulative distribution function
CDK2	Cyclin-dependent kinase 2
CH-plot	charge-hydrophobicity plot
CMV	Cytomegalovirus
CR	Congo Red; disodium 4-amino-3-[4-[4-(1-amino-4sulfonato-naphthalen-2-yl)diazenylphenyl]phenyl]diazenyl-naphthalene-1-sulfonate ($C_{32}H_{22}N_6Na_2O_6S_2$)
DEDAF	death effector domain-associated factor
DLS	Dynamic light scattering
DNTB	Ellman's reagent; 5-(3-Carboxy-4-nitrophenyl)disulfanyl-2-nitrobenzoic acid ($C_{14}H_8N_2O_8S_2$)
DSS	4,4-Dimethyl-4-silapentane-1-sulfonic acid ($C_6H_{16}O_3SSi$)
DTT	Dithiothreitol ; (2S,3S)-1,4-bis(sulfanyl)butane-2,3-diol ($C_4H_{10}O_2S_2$)
EDTA	Ethylenediaminetetraacetic acid; 2-({2-[Bis (carboxymethyl)amino]ethyl}(carboxymethyl)amino)acetic acid ($C_{10}H_{16}N_2O_8$)
FBS	Fetal bovine serum
GSH	Reduced glutathione; ((2S)-2-amino-4-{[(1R)-1[(carboxymethyl)carbamoyl]-2-sulfanylethyl] carbamoyl}butanoic acid ($C_{10}H_{17}N_3O_6S$)
GST	Glutathione S-transferase
HMM	high molecular mass
HSQC	Heteronuclear Single Quantum Coherence
IC_{50}	half maximal inhibitory concentration
IDP	Intrinsically disordered protein
LMM	Low molecular mass
MALDI-TOF	Matrix-assisted laser desorption/ionization time of flight
MBP	Maltose Binding Protein
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide ($C_{18}H_{16}BrN_5S$)
MW	molecular weight

NES	Nuclear export signal
NLS	Nuclear localization signal
NMR	Nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
Nur77	nerve growth factor IB
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PI	propidium iodide ($C_{27}H_{34}I_2N_4$)
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase
PKCβ	β isozyme of Protein Kinase C
PML	Promyelocytic leukemia
PMSF	phenylmethanesulfonyl fluoride ($C_7H_7FO_2S$)
Ppil3	Peptidyl-prolyl cis-trans isomerase-like 3
PP2A	Protein Phosphatase 2A
PRS	Proline Rich Sequence
p53	tumor suppressor protein
R_h	hydrodynamic radius
R_1	^{15}N longitudinal relaxation rate
R_2	^{15}N transverse relaxation rate
SDS	Sodium dodecyl sulfate ($NaC_{12}H_{25}SO_4$)
SEC	Size exclusion chromatography
TCEP	Tris(2-carboxyethyl)phosphine; 3,3′,3″-Phosphanetriyltripropanoic acid ($C_9H_{15}O_6P$)
TEM	transmission electron microscopy
TEV	Tobacco Etch Virus
Th T	Thioflavin T; 4-(3,6-dimethyl-1,3-benzothiazol-3-ium-2-yl)-N,N-dimethylaniline chloride ($C_{17}H_{19}ClN_2S$)
TOCSY	Total correlation spectroscopy
VP3	Virus Protein 3 (Apoptin)
XIAP	X-linked inhibitor of apoptosis protein

Ruiz-Martínez, S., Castro, J., Vilanova, M., Bruix, M., Laurents, D. V., Ribó, M., Benito, A. Delivery of antitumor Apoptin's variant is accomplished by removal of aggregation prone domains (submitted to Molecular Pharmaceutics).

Ruiz-Martínez, S., Pantoja-Uceda, D., Castro, J., Vilanova, M., Ribó, M., Bruix M., Benito, A., Laurents, D. V. Mechanistic insight into Apoptin's exquisitely selective anti-tumor cytotoxicity from atomic level characterization of its structure and dynamics (submitted to FEBS Journal).

Additional unrelated studies performed in fulfillment of the Ph.D. degree will not be presented in this thesis and are represented in the following publication:

Vert, A., Castro, J., Ruiz-Martínez, S., Tubert, P., Escribano, D., Ribó, M., Vilanova, M., Benito, A. (2012). Generation of new cytotoxic human ribonuclease variants directed to the nucleus. Mol Pharm 9(10): 2894-2902.

The present study is aimed at the characterization of the biological and biophysical properties of Apoptin, a viral protein comprising 121 amino acids that induces tumor-selective apoptosis in a wide range of human cancer cells. Apoptin has been fused to an N-terminal His-tag (called H_6-Apoptin), which has helped in its purification process. H_6-Apoptin forms multimers composed of roughly 50 subunits that are nontoxic when exogenously added to the media of NCI-H460 cancer cells, likely because they are not internalized. At its N-terminus Apoptin possesses a proline rich sequence (PRS, residues 8-28) and a leucine rich sequence (LRS, residues 33-46) that are believed to be important for protein aggregation. In order to endow this variant with cytotoxicity when externally delivered, two variants lacking the PRS (H_6-ApopΔPro) and both the PRS and LRS (H_6-ApopΔProΔLeu) have been produced. While H_6-ApopΔPro behaves like its parental protein, H_6-ApopΔProΔLeu is purified as a soluble non-aggregating protein that maintains most of the biological properties of Apoptin when transfected into cells. Additionally, it gains the ability to kill cancer cells through apoptosis when added externally, while leaving non-tumor cells unharmed.

The small size of this variant has greatly facilitated its study at atomic resolution using the NMR spectroscopy. The conformational characterization and dynamics show that this variant is an intrinsically disordered protein, which lacks a preferred conformation, and that is neither affected by the redox state of its Cys residues nor by Thr phosphorylation.

H_6-ApopΔProΔLeu is an ideal model to study the structural properties of Apoptin and the relationship with its remarkably selective anticancer activity providing a framework for the future design of improved Apoptin variants.

En el present estudi es caracteritzen les propietats biològiques i biofísiques de l'Apoptina, una proteïna vírica de 121 aminoàcids que indueix apoptosi selectiva sobre una àmplia varietat de cèl·lules canceroses humanes. S'ha fusionat un His-tag a l'extrem N-terminal de l'Apoptina (anomenada H_6-Apoptina), que ha ajudat al seu procés de purificació. H_6-Apoptina forma multímers d'unes 50 subunitats que no són tòxics per les cèl·lules canceroses NCI-H460 quan s'afegeix de forma exògena al medi de cultiu, probablement perquè no són internalitzats. L'Apotina conté al seu extrem N-terminal una seqüència rica en prolines (PRS, residus 8-28) i una seqüència rica en leucines (LRS, residus 33-46), que es creu que són importants per a l'agregació de la proteïna. Amb la finalitat de dotar aquesta variant de citotoxicitat quan s'administra externament, s'han produït dues variants: una sense el PRS (H_6-ApopΔPro) i l'altra sense el PRS ni el LRS (H_6-ApopΔProΔLeu). Mentre que H_6-ApopΔPro es comporta com la proteïna parental, H_6-ApopΔProΔLeu es purifica com una proteïna soluble sense agregar i manté la majoria de les propietats biològiques de l'Apoptina quan es transfecta en cèl·lules. A més, adquireix la capacitat de matar cèl·lules canceroses per apoptosi quan s'afegeix externament, deixant les cèl·lules no tumorals il·leses.

La petita mida d'aquesta variant ha facilitat en gran part el seu estudi a nivell atòmic utilitzant l'espectroscòpia de RMN. La caracterització conformacional i dinàmica mostren que aquesta variant és una proteïna intrínsecament desordenada, que no té una conformació preferent, i que no es veu afectada per l'estat redox del seus residus de Cys ni per la fosforilació de la Thr.

H_6-ApopΔProΔLeu és un model ideal per estudiar les propietats estructurals de l'Apoptina i la relació amb la seva selectiva activitat anticancerígena i proporciona un marc per al futur disseny de variants millorades d'Apoptina.

En el presente estudio se caracterizan las propiedades biológicas y biofísicas de la Apoptina, una proteína vírica de 121 aminoácidos que induce apoptosis selectiva sobre una amplia variedad de células cancerosas humanas. Se ha fusionado un His-tag al extremo N-terminal de la Apoptina (llamada H_6-Apoptina), lo que ha ayudado en su proceso de purificación. H_6-Apoptina forma multímeros de unas 50 subunidades que no son tóxicos para las células cancerosas NCI-H460 cuando se añade de forma exógena al medio de cultivo, probablemente porque no son internalizados. En su extremo N-terminal, la Apoptina posee una secuencia rica en prolinas (PRS, residuos 8-28) y una secuencia rica en leucinas (LRS, residuos 33-46) las cuales se cree que son importantes para la agregación de la proteína. Con el fin de dotar a esta variante de citotoxicidad cuando se administra externamente, se han producido dos variantes que carecen una del PRS (H_6-ApopΔPro) y la otra tanto del PRS como del LRS (H_6-ApopΔProΔLeu). Mientras que H_6-ApopΔPro se comporta como su proteína parental, H_6-ApopΔProΔLeu se purifica como una proteína soluble sin agregar que mantiene la mayoría de las propiedades biológicas de la Apoptina cuando se transfecta en células. Además, adquiere la capacidad de matar células cancerosas por apoptosis cuando se añade externamente, dejando las células no tumorales ilesas.

El pequeño tamaño de esta variante ha facilitado en gran medida su estudio a nivel atómico utilizando la espectroscopia de RMN. La caracterización conformacional y la dinámica muestran que esta variante es una proteína intrínsecamente desordenada, que carece de una conformación preferente, y que no se ve afectada por el estado redox de sus residuos de Cys ni por la fosforilación de la Thr.

H_6-ApopΔProΔLeu es un modelo ideal para estudiar las propiedades estructurales de la Apoptina y la relación con su selectiva actividad

anticancerígena y proporciona un marco para el futuro diseño de variantes mejoradas de Apoptina.

La présente étude caractérise les propriétés biologiques et biophysiques de l'Apoptine, une protéine virale composée de 121 acides aminés qui induit une apoptose sélective dans une large gamme de cellules cancéreuses humaines. L'Apoptine a été fusionée à un His-tag à l'extrémité N-teminale (appelée H_6-Apoptine), ce qui a aidé à son processus de purification. H_6-Apoptine forme des multimères composés d'environ 50 sous-unités qui ne sont pas toxiques lorsqu'ils sont additionnés de manière exogène au milieu des cellules cancéreuses NCI-H460, probablement parce qu'ils ne sont pas internalisés. À son extrémité N-terminale, l'Apoptine possède une séquence riche en proline (PRS, résidus 8-28) et une séquence riche en leucine (LRS, résidus 33-46) qui sont considérées importantes pour l'agrégation de cette protéine. Afin de doter cette variante de cytotoxicité lors de la livraison à l'extérieur, deux variantes dépourvues du PRS (H_6-ApopΔPro) et à la fois du PRS et du LRS (H_6-ApopΔProΔLeu) ont été produites. Alors que H_6-ApopΔPro se comporte comme la protéine parentale, H_6-ApopΔProΔLeu est purifiée en tant qu'une protéine soluble sans agrégation qui maintient la plupart des propriétés biologiques de l'Apoptine lors d'une transfection dans des cellules. En outre, elle gagne la capacité de tuer les cellules cancéreuses par apoptose lorsqu'elle est ajoutée à l'extérieur, tout en laissant les cellules non tumorales indemnes.

La petite taille de cette variante a grandement facilité son étude à résolution atomique en utilisant la spectroscopie RMN. La caractérisation conformationnelle et la dynamique montrent que cette variante est une protéine intrinsèquement désordonnée, qui ne possède pas une conformation préférable, et qui n'est pas affectée par l'état redox de ses résidus Cys ni par la phosphorylation de la Thr.

H_6-ApopΔProΔLeu est un modèle idéal pour étudier les propriétés structurelles de l'Apoptine et sa relation avec son activité anticancéreuse

remarquablement sélective et fournit un cadre pour la conception future de variantes d'Apoptine améliorées.

Introduction

Cancer, standard and targeted therapies

Cancer is the general name for a group of more than one hundred diseases that can affect any part of the body. Human cells divide regularly in order to replace aging and dead cells, and thereby maintain the integrity and proper functioning of organs. Cell division is a well-regulated process, however when it breaks down and a normal cell grows out of control, a tumor may be formed. When this cellular growth is uncontrollable and it spreads in the body the tumor develops into cancer.

Standard cancer therapy depends on the type of cancer and stage of disease progression. Surgery removal followed by radiotherapy or chemotherapy in order to eradicate remaining cancer cells is the primary therapy when the tumor is localized. Surgery is also used in the diagnosis through a biopsy, determining also the status (how slowly or quickly it may grow). In radiation, high energy waves are used to damage or kill cancer cells preventing its growth. For these cases in which cancer cells have invaded other parts of the body, chemotherapy is the most widely used alternative. In chemotherapy, drugs that interfere with processes like DNA synthesis or cell proliferation are used to kill cancer cells. Although normal cells might be more resistant to the action of these drugs, those cells replicating fast such as bone marrow cells or the ones forming the intestinal wall are also likely affected, resulting in undesirable side-effects, i.e. gastrointestinal symptoms and myelosuppression. In the last twenty years, however, targeted therapies have been a component of treatment for many types of cancers. In this regard, one of the most successful strategies for fighting cancer in the past decade has been the antibody-based therapy. Antibodies can function by modulating the immune system or by delivering a specific drug that is conjugated to the antibody. Among the most paradigmatic examples of antibody-based anticancer drugs it should be mentioned Trastuzumab for the treatment of ERBB2-positive breast cancer, Bevacizumab for the treatment of

metastatic colon cancer or Cetuximab for the treatment of the squamous cell carcinoma of head and neck (Afghahi and Sledge, 2015; Chari et al., 2014; Scott et al., 2012).

The term "targeted therapy" was first introduced by Paul Erlich in the late 1800s when he proposed the idea of "magic bullet", originally describing a chemical with the ability to specifically target microorganisms (Imai and Takaoka, 2006). In cancer, the term targeted therapy refers to those drugs designed to specifically bind to a molecular target that is mainly expressed in tumor cells and that is believed to have a critical role in tumor growth, progression or spread of cancer. Unlike standard therapies that act on rapidly diving cells, targeted cancer therapies act on specific molecular targets that are associated with cancer. These targets include growth factor receptors, signaling molecules, cell-cycle proteins, modulators of apoptosis and molecules involved in invasion and angiogenesis (formation of new blood vessels), which are key processes for development and homeostasis in normal tissues.

Growth-signaling pathways, and especially their receptor proteins, are interesting objectives for targeted anticancer strategies. Monoclonal antibodies, immunoglobulin structures that specifically target a certain antigen or that are joined to radio-isotopes or toxins to be delivered into cancer cells, and small molecule inhibitors that interfere with molecules required for tumor development and growth, have significantly changed the treatment of cancer over the last two decades. Examples of successful targeted therapies are the tyrosine kinase inhibitor (TKI) imatinib, a small molecule approved in 2002 for the treatment of chronic myeloid leukemia (CML) (Afghahi and Sledge, 2015; Deininger et al., 2005) and the above mentioned Trastuzumab, a humanized monoclonal antibody related to EGFR that disrupts ligand-independent intracellular signaling (Afghahi and Sledge, 2015) of the HER2 transmembrane receptor, which is

overexpressed in approximately 25 percent of patients with breast cancer (Slamon et al., 1987).

More recently, a new approach using viral genes has been investigated in order to develop novel anticancer agents (Gupta et al., 2014). Many viral proteins perturb normal cell physiology of infected cells, leading them to apoptosis by providing the necessary upstream signals. Apoptosis may be used by many viruses to kill cells at the end of the infection cycle (Roulston et al., 1999). There are some viruses that selectively infect or replicate in cancer cells, and can therefore be used with therapeutic purposes (Parato et al., 2005). These viruses are known as oncolytic viruses and can be either naturally attenuated strains or genetically modified to mediate oncolytic effects (Kirn et al., 2001; Nemunaitis, 1999).

Viruses and oncogenesis

Some viral genes and their encoded products possess a great potential to be developed as novel anticancer therapeutic agents which can specifically target and kill the cancer cells leaving the normal cells unharmed. Many viruses have the ability to modulate apoptosis by either inhibiting or promoting it.

Viruses that block apoptosis protect infected cells from early death by producing certain proteins, that can inhibit one or more key pro-apoptotic factors such as Bax, Bak, caspases, p53, Fas/TNF, PKR, etc (Roulston et al., 1999), preventing the inflammatory and apoptotic response pathways. This interference with cellular processes maximizes viral replication and helps in the evasion of the immune response. However, many other viruses possess specific genes that kill the infected cells by multiple mechanisms including apoptosis, normally with the aim to assist the virus spread in the later stages of viral infection (Roulston et al., 1999). This is an important trait since many cancer cells elude apoptosis due to the accumulation of alterations in the apoptotic machinery. Some of the

proteins involved in the viral induction of apoptosis are E1A 12S and 13S proteins, E3 and E4 of human adenovirus type 5, E2 and E7 of papilloma viruses, Apoptin (VP3) of chicken anemia virus (CAV), tat of HIV-1, parvovirus NS1, SV40 large T antigen and HN protein of Newcastle disease (for a review see (Gupta et al., 2014)). These proteins are interesting in terms of drug development since they are effective in a wide range of cancer cells, functioning regardless of the genetic background and type of cancer (Argiris et al., 2011). In this work we have focused on the apoptosis-inducing VP3 protein produced by the Chicken Anemia Virus (CAV).

Chicken anemia virus (CAV)

CAV was first described in Japan in 1974 (Yuasa et al., 1979), although it is found worldwide. CAV is one of the smallest avian non-enveloped viruses (23-25 nm), has a circular negative sense single-stranded DNA of about 2.3 kbp (Noteborn et al., 1991) and is one of the four members of the genus *Gyrovirus* within the family *Circoviridae* (Chu et al., 2012; Pringle, 2014). The virus, that affects chickens, can be transmitted both vertically and horizontally and the disease is coursed with high mortality (up to 60%), weight loss, abnormal feathers, leg paralysis, intramuscular hemorrhages, anemia, growth retardation, aplasia of bone marrow, and atrophy of thymus, owing to apoptosis of the cells in the thymus, bone marrow and spleen (Noteborn and Koch, 1995; Yuasa et al., 1976).

The major transcript from the CAV genome is a single poly-adenylated poly-cistronic mRNA of about 2,100 nucleotides that comprises three partially overlapping open reading frames (ORFs) encoding three proteins (Figure 1). Viral protein 1 (VP1) is a 51.6 kDa protein that forms the virion icosahedral capsid. The 24.0 kDa viral protein 2 (VP2) is a non-structural protein which possesses a dual-specificity protein phosphatase (DSP) activity (Peters et al., 2002) and may also act as scaffold for capsid assembly (Noteborn and van der Eb, 1998). It is suggested that VP2 protein

phosphatase activity is not essential, but is required for efficient virus replication (Peters et al., 2005). CAV VP2 weakly induces apoptosis in tumor cells as well as in normal cells (Noteborn, 2004). The third viral protein, VP3, also named Apoptin, is a 13.6 kDa protein, necessary for CAV replication. Apoptin is a strong inducer of apoptosis in chicken thymocytes and chicken lymphoblastoid cell lines (Jeurissen et al., 1992) among others, although its natural function in the viral cycle has not yet been elucidated.

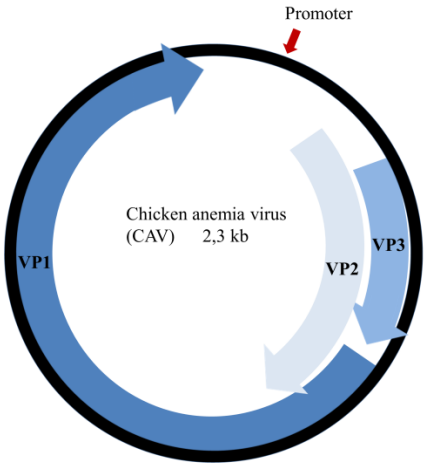


Figure 1. Schematic representation of the CAV genome. The CAV genome contains three partially overlapping genes that encode for three distinct proteins, VP1: 54.6 kDa, VP2: 24.0 kDa and VP3: 13.6 kDa.

Apoptin (VP3): an antitumor protein

CAV VP3 is able to induce apoptosis of multiple mammalian transformed and malignant cell lines while leaving primary and non-transformed cells unaffected (Maddika et al., 2006; Poon et al., 2005). Moreover, long term expression of Apoptin in normal human fibroblasts has revealed that it has no toxic or transforming activity in these cells and that Apoptin does not interfere with normal cell proliferation (Backendorf et al., 2008). Apoptin-induced apoptosis is independent of p53, an important characteristic since more than 50 percent of the human tumors contain

mutated p53 (Bieging and Attardi, 2012) and many standard therapies require functional p53 for inducing apoptosis in tumor cells. Apoptin also induces apoptosis independently of death receptors (Maddika et al., 2005; Zhuang et al., 1995) and is mediated by cytochrome c release and procaspase-9 activation (Han et al., 2010; Maddika et al., 2005). Apoptin's cytotoxicity is also independent of the antiapoptotic proteins Survivin, $FLIP_S$, XIAP and cIAP-1. Likewise the loss of function of the proapoptotic protein Bax does not influence its activity (Liu et al., 2006). The role of the antiapoptotic Bcl-2 proteins is controversial since in some tumor cell lines Apoptin seems to act independently of the status of Bcl-2 or it can be stimulated by this protein (Danen-Van Oorschot et al., 1999a, 1999b; Liu et al., 2006) while in some other cell lines both Bcl-2 and BclxL inhibit apoptosis triggered by Apoptin (Burek et al., 2006). Bcl-2 depends on the involvement of Nur77, a member of the steroid receptor family that is differentially expressed depending on the cell type, which makes that the observed differences might ultimately be due to the tumor cell assayed.

When Apoptin is expressed in living cells, it forms large aggregated complexes with a diameter of about 200 nm as stated by immune-electron microscopy (Leliveld et al., 2003a). Apoptin aggregates produced *in vitro* are toxic when they are microinjected to Saos-2 cancer cells but not to VH10 normal cells (Leliveld et al., 2003a). These aggregates bind to naked, double-stranded and single-stranded DNA with no sequence specificity (Leliveld et al., 2003b). It has been postulated that Apoptin binding to DNA is linked to its apoptotic activity (Leliveld et al., 2004). In this regard, Apoptin's localization has also been related to the DNA damage response (DDR) signaling. Relocation of Apoptin to the nucleus was observed upon the induction of DNA damage in primary cells. In addition, the inhibition of DDR signaling in transformed cells resulted in cytoplasmic localization of Apoptin (Kucharski et al., 2011). Apoptin gene has also shown efficacy in

treated xenografted tumors in mice using different delivery strategies (Li et al., 2010).

Apoptin's primary structure and functional domains

Apoptin is a relatively small protein composed of 121 amino acids that contains two cytotoxic domains (Danen-Van Oorschot et al., 2003) (Figure 2). The C-terminal region contains a bipartite nuclear localization sequence (NLS1, residues 82-88; NLS2, residues 111-121) and a nuclear export sequence (NES, residues 97-105) that is sandwiched between NLS1 and NLS2 (see Figure 2). Adjacent to the NES, lies Thr108 whose phosphorylation prevents NES recognition by the nuclear export protein 1 (XPO1) or also called chromosomal maintenance 1 (CRM1) (Figure 3) (Poon et al., 2005; Rohn et al., 2002). This region also contains a Proline-rich sequence (PRS, residues 81-86) that is located within the NLS1 (Jangamreddy et al.; Panigrahi et al., 2012). At the N-terminal domain, Apoptin possesses a Leucine-rich sequence (LRS) spanning residues 33 to 46 that promotes aggregation (Leliveld et al., 2003a) and a second PRS (residues 8-28) (Noteborn et al., 1991, 1994). When residues 1 to 65, which include the LRS, are removed from Apoptin the resulting variant seems to equilibrate between monomers, dimers and trimers, however the cytotoxicity of this new variant was not tested (Leliveld et al., 2003a). The LRS was first believed to be a CRM1-mediated NES (Wang et al., 2004), nevertheless, it facilitates nuclear accumulation rather than a cytoplasmic retention and it is strongly dependent on the bipartite nuclear targeting signal (NLS1 and NLS2) (Poon et al., 2005b).

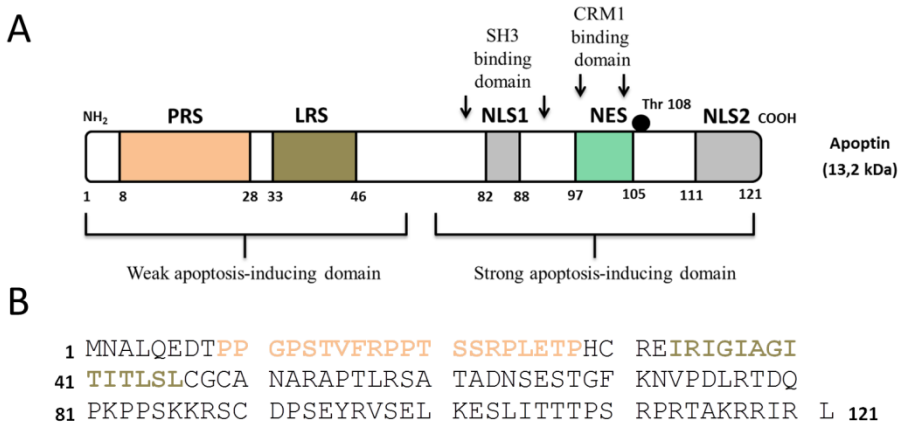


Figure 2. The primary structure of Apoptin. **A)** The key domains and important sequences are indicated. **B)** Apoptin's amino acid sequence. In orange are indicated the residues of the proline rich sequence (PRS) and in brown the residues of the leucine rich sequence (LRS).

Apoptin is believed to contain little secondary structure within the aggregate complexes that forms as stated by circular dichroism (CD) (Leliveld et al., 2003a). According to this study, Apoptin lacks both α-helical and β-sheet structure, although its conformation might not be fully random as determined by the large negative ellipticity spanning from 190 to 200 nm which could imply that some Pro residues from the PRS within the multimerization domain may adopt a helix-like conformation. Additionally, the solvent protection of the Tyr95 found by fluorescence experiments suggests that the multimeric complex contains certain order, that can be either intra- or intermolecular (Leliveld et al., 2003a). Furthermore, another study has predicted the Apoptin's 3D structure based on computational protein modeling using proteins with known 3D structure and based on sequence similarity (Panigrahi et al., 2012). In this study a well ordered globular structure for Apoptin is predicted. Although the Ramachandran plot shows that almost 82 percent of Apoptin's residues are in the most favored regions, the outcome of a computational modeling might be strongly influenced by the template proteins used. It is worth mentioning that in this study the sequence identity between Apoptin and

template was 31%, which is fairly close to the target-template sequence identity percentage where it becomes hardly reliable (Fiser, 2010).

Recently, a protein homologue of CAV's Apoptin has been isolated from Human Gyrovirus (HGyV) (Sauvage et al., 2011). HGyV encodes a 125 amino acid protein that contains the nuclear localization and export signals as well as the phosphorylation site, although the overall identity is low (Sauvage et al., 2011). It has been shown that this protein induces apoptosis in tumor cells where it accumulates in the nucleus while it does not affect 1BR3 normal cells (Bullenkamp et al., 2012).

Mechanism of antitumor action of Apoptin

The mechanism by which Apoptin specifically induces apoptosis of tumor cells is not fully understood and different non-mutually exclusive hypotheses have been proposed.

It is known that Apoptin does not act as a direct transcriptional repressor or activator but might still indirectly affect the activity of other cellular transcriptional regulators (Danen-Van Oorschot et al., 2003).

The selective cytotoxicity of Apoptin has been attributed in part to the differential subcellular localization of the protein in tumor and normal cells. Phosphorylation of Thr108 is catalyzed by tumor specific kinases and therefore, in tumor cells, Apoptin mainly accumulates in the nucleus whereas in normal cells it accumulates in the cytoplasm, where it is subsequently degraded. Apoptin's aggregates in the cytoplasm of normal cells become epitope shielded and are eventually eliminated (Figure 3) (Zhang et al., 2003). In this regard, Thr108 replacement by Ala in Apoptin partially impairs its apoptotic activity in tumor cells whereas its replacement by Glu, which is rather good mimic of phosphothreonine, gains the ability to also kill non-tumor cells (Poon et al., 2005b; Rohn et al., 2002). It has been reported that, at least *in vitro,* the preceding threonine

(Thr107) might also be phosphorylated and be able to play the role of Thr108 in the context of the Thr108Ala mutation (Lanz et al., 2012). Indeed, both residues, Thr107 and Thr108 have to be mutated to completely abolish Apoptin's cytotoxic activity (Rohn et al., 2005).

Even though other specific kinases could phosphorylate Apoptin, it has been shown that in some cases Thr108 may be phosphorylated by cyclin A/CDK2 (Maddika et al., 2009) which, in turn, is activated by the PI3K-Akt pathway (Maddika et al., 2008a) and by the β isozyme of protein kinase C (PKCβ) (Jiang et al., 2010). In the presence of Apoptin, Akt (or protein kinase B) translocates to the nucleus where it activates CDK2, which subsequently phosphorylates Apoptin at Thr108. Activation of CDK2 by Akt is carried out by both direct phosphorylation of CDK2 and phosphorylation-induced proteasomal degradation of the CDK2 inhibitor $p27^{Kip1}$ (Maddika et al., 2008b). It is interesting that the C-terminal PRS of Apoptin, has a sequence (PKPPSK) that matches the consensus motif recognized by SRC homology 3 (SH3) domains (PxxPxR/K), interacts with the SH3 domain of the p85 regulatory subunit of phosphoinositide 3-kinase (PI3-K) (Panigrahi et al., 2012) and that this interaction is essential for both the cytotoxic activity of Apoptin and its nuclear localization (Maddika et al., 2008a). PI3-K however, does not directly phosphorylate Apoptin (Maddika et al., 2008a).

The PKCβ has been shown to phosphorylate Apoptin in multiple myeloma cell lines (Jiang et al., 2010). A recent study has described a tumor-specific interaction between Apoptin and PKCβ1, which results in PKC-dependent phosphorylation of Apoptin as well as in nuclear translocation and activation of PKC (Bullenkamp et al., 2015).

Although the Apoptin's compartmentalization is believed to be important for its selectivity, this cannot be the sole mechanism by which this protein induces tumor-specific apoptosis. This is supported by the fact that

Apoptin does not induce cell death of normal cells when its nuclear localization is enforced by means of a heterologous NLS (Danen-Van Oorschot et al., 2003; Los et al., 2009). Also, some studies have indicated that non-phosphorylatable Apoptin mutants still induce apoptosis of cancer cells (Heckl et al., 2008; Heilman et al., 2006; Kucharski et al., 2011; Lee et al., 2007). This has been attributed to the existence of a second apoptosis motif at the N-terminus, whose action is phosphorylation independent (Danen-Van Oorschot et al., 2003). Both, N- and C-terminal regions (residues 1 to 69 and 70 to 121, respectively) have inherent cell-killing activity although they are less effective than the full-length protein. Interestingly, both regions are also able to bind dsDNA (Leliveld et al., 2004).

Apoptin interacts with a number of cellular proteins that affect its localization as well as its cytotoxic effects. These interacting partners may either be involved in the direct activation of Apoptin or in pathways that influence the induction of apoptosis. The above described molecules and other Apoptin partners are summarized in Table 1.

Table 1. Apoptin interacting partners (extracted from (Bullenkamp and Tavassoli, 2014)).

Molecule	Biological function	References
Protein Kinase C β (PKCβ)	Binding of Apoptin to PKCβ results in PKCβ-dependent phosphorylation and activation of Apoptin and nuclear translocation of PKCβ	(Jiang et al., 2010)
Fas-associated protein with death domain (FADD)	Overexpression causes co-localization of FADD and Apoptin in cytoplasmic death effector filaments, potentially interfering with death receptor signaling	(Guelen et al., 2004)
Bcl-10	Apoptin and the NF-κβ regulator Bcl-10 co-localize in cytoplasmic filament with a yet unknown consequences	(Guelen et al., 2004)

Continuation of **Table 1**

Molecule	Biological function	References
Protein kinase G (PKG-I)	High levels of PKG-I in normal cells correlate with lower Apoptin activation, however the precise role for PKG-I during Apoptin-induced cell death remains unclear	(Bullenkamp and Tavassoli, 2014)
PI3-Kinase (PI3K) and Akt	Apoptin binds to the p85 subunit of PI3K, activating PI3K which results in sustained activation and nuclear translocation of Akt. This triggers activation of CDK2 which in turn phosphorylates Apoptin	(Maddika et al., 2007, 2008a, 2009)
Heat shock cognate protein 70 (Hsc70)	Binding of Apoptin triggers nuclear translocation of Hsc70 which might be required for Apoptin-induced Akt activation and downregulation of p65	(Chen et al., 2011a, 2011b)
Promyelocytic leukemia protein (PML)	Apoptin is sumoylated and targeted to nuclear PML bodies. However, the interaction with PML is not required for the cytotoxic function of Apoptin	(Janssen et al., 2007)
Anaphase promoting complex I (APCI)	Apoptin binds to APCI, disrupting the APC/C and resulting in G2/M arrest and apoptosis	(Heilman et al., 2006; Teodoro et al., 2004)
Hip-1 protein interactor (Hippi)	In normal cells but not in tumor cells Apoptin and Hippi interact in the cytoplasm	(Cheng et al., 2003)
Death effector domain-associated factor (DEDAF)	Co-expression of Apoptin and DEDAF results in enhanced apoptosis compared to expression of either protein alone	(Zheng et al., 2001)
Peptidyl-prolyl isomerase like 3 (Ppil3)	Overexpression of Ppil3 can promote cytoplasmic localization of Apoptin	(Huo et al., 2008)
N-myc interacting protein (Nmi)	Nmi was identified as an Apoptin binding partner but no functional studies are available yet	(Sun et al., 2002)
Breast cancer associated gene 3 (Bca3)	Bca3 enhances Apoptin phosphorylation and cytotoxic function by a still unknown mechanism	(Zimmerman et al., 2012)

The C-terminal region of Apoptin is reported to interact with peptidyl-prolyl isomerase like 3 (Ppil3) (Huo et al., 2008), a member of the cyclophilin family that has a role in mitochondrial maintenance, cell cycle progression, and apoptosis (Yao et al., 2005). Thr108 is essential for its

interaction with Ppil3. Interestingly, Apoptin bound to Ppil3 is kept in the cytoplasm of cancer cells (Huo et al., 2008). This is dependent on the Pro109 residue and it has been suggested that binding of Ppil3 to this region would modify the activity of the nearby NES (Huo et al., 2008). Apoptin associates also through its C-terminus region with human Hippi (protein interactor of the huntingtin-interacting protein 1), resulting in caspase-8 recruitment (Figure 3). In normal cells, where Apoptin remains in the cytoplasm associated with Hippi, Apoptin's activity is abolished (Cheng et al., 2003). Thus it has been postulated that the interaction with Ppil3 or Hippi may prevent apoptosis in normal cells by sequestering Apoptin in the cytoplasm.

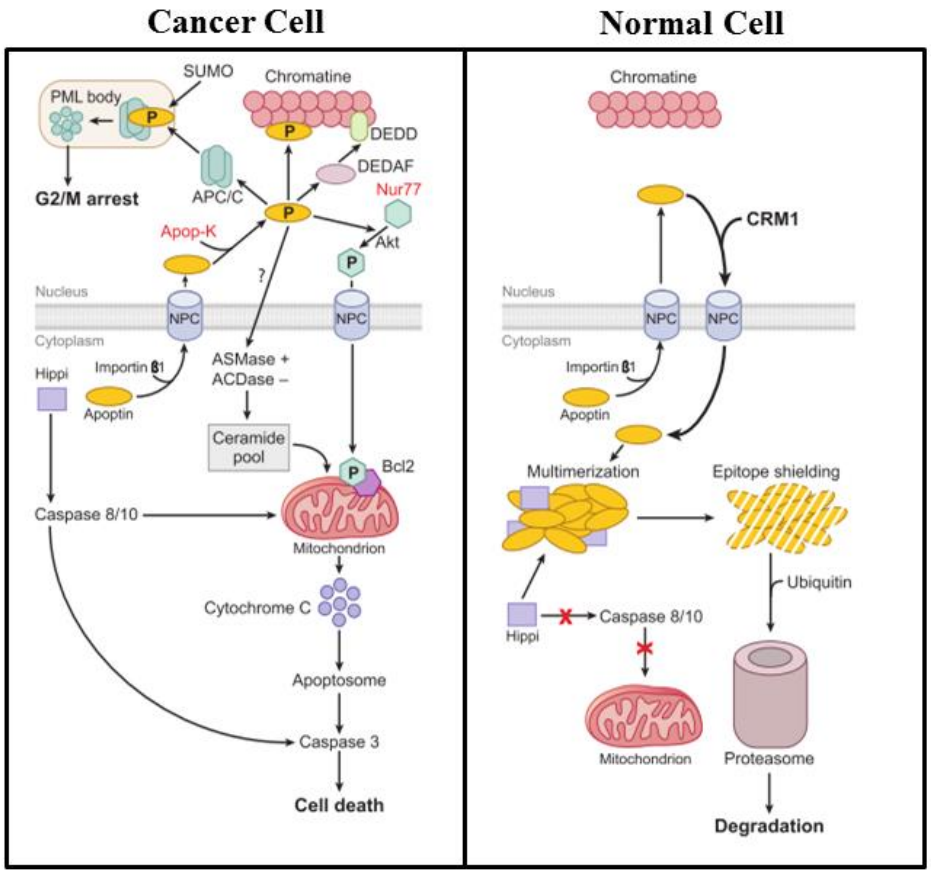


Figure 3. Schematic representation of different interacting Apoptin partners in normal (right) and cancer (left) cells (extracted from (Backendorf et al., 2008)).

Another mechanism of action proposed for Apoptin is through its interaction with the nerve growth factor IB Nur77 (also known as TR3 or NGFI-B) (Figure 3). Nur77 is a nuclear orphan receptor member of the steroid/thyroid receptor family that acts as an anti-apoptotic factor when it stays in the nucleus. However, in tumor cells, Apoptin triggers the cytoplasmic translocation of Nur77 where it acts as a pro-apoptotic factor, promoting cell death by causing the loss of mitochondrial membrane potential, the release of cytochrome c and the apoptosis inducing factor (AIF) (Maddika et al., 2005). AIF is a mitochondrial flavoprotein that enters the nucleus where it causes chromatin condensation and the large-scale DNA fragmentation (Susin et al., 1999).

Apoptin seems to interact with protein phosphatase 2A (PP2A). It has also been described that the interference with the normal function of PP2A is sufficient to activate Apoptin in a tumor-selective fashion (Zimmerman et al., 2012). The breast cancer associated gene 3 protein (Bca3 or AKIP1), an Apoptin binding partner (Table 1) that interacts with the catalytic subunit of protein kinase A (PKA), has been shown to enhance Apoptin's phosphorylation in tumor cells (Zimmerman et al., 2012). PKA interacts with PP2A, a key regulator of cellular protein phosphorylation which can either induce or prevent apoptosis. The inactivation of the nuclear PP2A, particularly of the B56-δ and –γ subunits is sufficient to trigger the tumor-selective apoptosis activity of Apoptin (Zimmerman et al., 2012). This would indicate that not only different kinases are implicated in the regulation of Apoptin's activity depending on the type of cancer, but also phosphatases might have an important role.

Additionally, it has been hypothesized that Apoptin interacts with the human death effector domain-associated factor (DEDAF) (Figure 3) (Danen-van Oorschot et al., 2004). DEDAF can promote apoptosis in tumor cells by binding to several DED-containing proteins (Zheng et al., 2001),

nevertheless when co-expressed with Apoptin, the cytotoxic activity is higher than that of either protein expressed alone, probably acting synergistically (Table 1).

Promyelocytic leukemia (PML) protein interacts with Apoptin (Figure 3) through its LRS (Poon et al., 2005b). PML is a critical player in oncogenesis and tumor progression (for a review, see (Zhou and Bao, 2014)), recruiting APC/C (anaphase-promoting complex/cyclosome) to these subnuclear structures by interacting with its subunit APC1. This disruption of the APC/C complex results in G2/M arrest (Figure 3) and ultimately in apoptosis in the absence of p53 in transformed cells (Teodoro et al., 2004). However, it has been reported that neither the sumoylation (a posttranslational modification that takes place in PML nuclear bodies) nor the interaction with PML are essential for Apoptin's cytotoxicity but might be essential for CAV replication (Janssen et al., 2007).

The N-Myc interaction (Nmi) protein has also been suggested to interact with the LRS of Apoptin, although its function has not been completely elucidated (Sun et al., 2002). Nmi is an interferon inducible protein highly expressed in leukemia cell lines but downregulated in normal tissues (Bao and Zervos, 1996) and thus, its interaction with Apoptin would be restricted to these cell lines and does not explain the Apoptin's wide antitumor activity.

Strategies for Apoptin delivery

The fact that Apoptin specifically induces apoptosis in a broad spectrum of tumor cell lines and its ability to switch survival pathways into proapototic pathways make this protein an interesting candidate to be used in cancer therapy, either alone or combined with other drugs. Nonetheless, effective delivery of Apoptin to tumor cells remains a challenge, especially because this protein is very prone to aggregate. The concentration of the protein seems to be an important trait to induce

apoptosis, hence different thresholds depending on the cell type must be reached (Guelen et al., 2004; Wadia et al., 2004). Efficient transduction tools such as recombinant viruses, bacterial or a recombinant plasmid within liposomes have been used to deliver the Apoptin's gene into tumor cells.

The Apoptin gene has been successfully delivered using non-replicative adenoviral vectors with the consequent virus production and adenovirus-mediated transfer of the Apoptin gene. Using this strategy, Apoptin's gene arrests cell growth of hepatoma cells (Pietersen et al., 1999) or prostate cancer cell lines and xenografts (Liu et al., 2006). Interestingly, the adenoviral vector has also been effective in different cholangiocarcinoma cell lines, which are representative of a form of cancer with poor prognosis (Pietersen et al., 2004). Another strategy that effectively suppressed primary and metastatic tumors both *in vivo* and *in vitro* is the intratumoral injection or systemic delivery of the conditional replication-competent adenovirus Ad-hTERT-E1a-Apoptin, where the cancer specific promoter of human telomerase reverse transcriptase (hTERT) regulates the expression of the E1a-Apoptin (Liu et al., 2012; Qi et al., 2014; Zhang et al., 2013). Other viral vectors such as pseudotype baculovirus (Pan et al., 2010), lentivirus-based vectors, oncolytic viruses and conditionally replicative oncolytic adenoviruses have been assayed as potential Apoptin delivery methods (for a review see (Rollano Peñaloza et al., 2014)). Apart from viruses, also attenuated intracellular bacterium *Salmonella typhimurium* has been assayed to deliver the Apoptin's gene into gastric cancer cells *in vivo* (Cao et al., 2010) and into human laryngeal carcinoma cells (Guan et al., 2013).

A non-viral vector as systemic delivery vehicle containing Asor-Apoptin (the asialoglycoprotein receptor ligand), PLL (poly-L-lysine linker), and the plasmid pcDNA-vp3 encoding Apoptin demonstrated specific delivery of Apoptin to *in situ* hepatocarcinoma, resulting in tumor regression (Peng

et al., 2007). Furthermore, Apoptin coding plasmid (pTracer-CMV2) administered via electroporation directly into murine B16(F10) tumors (Mitrus et al., 2005) was tested although low transfection levels in the tumor cells were achieved. Another non-viral Apoptin gene delivery system tested is based on the transfection of brain tumor cells or intratumoral injection into intradermal located U87MG brain tumors in nude mice by means of a polyamidoamine dendrimer with an arginine surface (PAM-RG4) (An et al., 2013).

These approaches have restrictions in terms of therapeutic application including insertional mutagenesis of transfected cells, transient expression or high immunogenicity. Other major drawbacks for gene transfer are low efficiency and poor specificity (Herrmann and Biology, 1995) and, for the treatment of humans, transgene toxicity as well as inherent viral toxicity must also be taken into account. Although during the past decade good progress has been made in establishing the most desirable method in terms of safety and effectiveness, many disappointing results have been obtained (Yi et al., 2011). These concerns have stimulated the search of alternative delivery methods based on the Apoptin protein. Fusion to cell-penetrating peptides (CPP), the trans-acting activator of transcription (TAT) peptide transduction domains (PTDs) (Guelen et al., 2004) and the protein transduction domain 4 (PTD4) (Jin et al., 2011; Sun et al., 2009) have been investigated for protein delivery (Guelen et al., 2004; Ho et al., 2001). Furthermore, the combination of Apoptin treatment with classic chemotherapeutic agents such as dacarbazine with PTD4-apoptin (Jin et al., 2011) or recombinant adenovirus vector encoding Apoptin in combination with paclitaxel or etoposide (Olijslagers et al., 2007) have proven efficacy in reducing the tumor growth. Additionally, a combined strategy using Apoptin and a downregulation of Survivin, an apoptosis inhibitor protein overexpressed

in many tumors significantly inhibited cancer cell growth more efficiently than a single treatment (Liu et al., 2008; Panigrahi et al., 2008).

Objectives

In biomedical research, cancer is likely the most elusive condition nowadays, one for which there is an unmet need for therapies. Over the years, many have been the strategies to develop new agents with the ability to kill cancer cells. Lately, advances in a novel approach using viruses and their encoded proteins with cytotoxic activity for cancer cells have been investigated. Apoptin, a protein produced by the chicken anemia virus is now one of the most promising anticancer drug candidates. However, its use is limited by its strong tendency to aggregate. In recent years, many efforts have been made in order to study the structural properties of Apoptin and their relationship with its mechanism of action, though until now the structural details of Apoptin remain unknown mostly because its aggregated state has hampered its characterization.

Taken into account these aspects, the main objectives of the present work can be summarized as follows:

1. To set up a purification protocol for Apoptin and to characterize its biophysical and biological properties.

2. To design, construct, produce and purify Apoptin variants less prone to aggregation.

3. To compare the new variants' biological properties with those of Apoptin and to investigate their effect when externally delivered to both tumor and non-tumor cells.

4. To characterize the conformation and dynamics of the new Apoptin variants by different biophysical techniques.

Material and Methods

Bacterial strains:

E. coli DH5α. This strain was used to obtain multiple DNA copies during the construction of the recombinant variants. It can be transformed with high efficiencies and, moreover its genotypic character hsdR17 (rk-, mk+) provides this strain with the ability to methylate DNA and impedes the destruction of any foreign DNA.

Genotype: *F-, ø80ΔlacZΔM15, Δ(lacZYA-argF), U169, deoR, recA1, endA1, hsdR17(rk-, mk+), phoA, supE44, λ-, thi-1, gyrA96, relA1*

***E. coli* BL21(DE3).** This strain was used for high-level protein expression. It is deficient in Lon and OmpT proteases and it is therefore suitable for expression of non-toxic genes. The BL21(DE3) contains the T7 RNA polymerase gene, under the control of the *lacUV5* promoter, integrated into the chromosome. Isopropyl-D-thiogalactoside (IPTG) is used to induce the expression of recombinant genes cloned into vectors downstream of T7 RNA promoter and transformed into this strain cells.

Genotype: F^-, $ompT^-$, $hsdS(rB^- mB^-)$, *dcm+, gal, λ(DE3),* lon^-

Prokaryotic vectors:

pET-22b(+). This vector (Novagen, USA) carries an N-terminal *pelB* signal sequence for potential periplasmic localization, plus optional C-terminal His-tag sequence. The cloning/expression region of the coding strand is transcribed by T7 RNA polymerase, following this sequence there is a multiple cloning site (MCS) with ten unique restriction sites, where the foreign DNA can be inserted. The f1 origin is oriented so that infection with helper phage will produce virions containing single-stranded DNA that corresponds to the coding strand. This vector also contains an origin of replication (ori) for double stranded replication, as well as the *lac promoter* (Studier and Moffatt, 1986) (T7lacUV5), that controls the expression of T7 RNA polymerase in strains with the DE3 bacteriophage λ

lysogen, which in turn controls the expression of the desired protein. Ampicilin gene is the antibiotic resistance gene of this plasmid.

pET-28a(+). This vector (Novagen, USA) carries an N-terminal His-tag/thrombin/T7·Tag configuration plus a C-terminal His-Tag sequence. The cloning/expression region of the coding strand is transcribed by T7 RNA polymerase. At 5' of this sequence there is a MCS with eleven unique restriction sites, where a transgene can be introduced. The f1 origin is oriented so that infection with helper phage will produce virions containing single-stranded DNA that corresponds to the coding strand. This vector also contains an ori for double stranded replication, as well as the expression of T7 RNA polymerase under the control of T7lacUV5, inducible with IPTG. The vector confers resistance to Kanamycin.

pRSET/EmGFP. This vector (Thermo Fisher Scientific, USA) contains the sequence encoding the Emerald Green Fluorescent Protein (EmGFP) derived from eGFP (excitation/emission: 487/509 nm). The vector contains unique restriction sites flanking the EmGFP gene to allow its transfer to any vector of choice. It possesses the bacteriophage T7 promoter for high-level, inducible expression of the EmGFP in *E. coli* and the pUC origin for high-copy replication and maintenance of the plasmid. When using the BL21-derived strain that contains the DE3 bacteriophage λ lysogen, the DE3 lysogen contains the T7 RNA polymerase under the control of the *lacUV5* promoter, allowing expression of T7 RNA polymerase to be induced by IPTG. The vector confers resistance to ampicillin.

pBADm: these series of vectors were obtained from the Protein Expression and Purification Core Facility (EMBL, Germany) and are designed to increase the solubilization of the protein of interest and facilitate its purification through binding to several fusion partners. They possess the replication origin of *E.*coli and the araBAD promoter, which is inducible with arabinose. These vectors also contain a codifying His-tag

sequence, the gene of a solubilizing protein, and the sequence of Tobacco Etch Virus (TEV) protease cleavage site (Allison et al., 1986) followed by an MCS where a transgene can be inserted. The recombinant proteins coded by these vectors are fused to a solubilizing protein that contains a 6xHisTag for affinity purification. The vectors used in this work are **pBADm-20(+)**, **pBADm-30(+)**, **pBADm-41(+)**, and **pBADm-60(+)**, which contain the thioredoxin (TrxA), the glutathione S-transferase (GST), the maltose binding protein (MBP) and the NusA, respectively. The solubilizing protein can be later on excised from the Apoptin protein using the TEV protease. The vector confers resistance to ampicillin.

Eukaryotic vectors:

pcDNA3.1(+). This is a 5.4 kb vector (Invitrogen, USA) derived from pcDNA3 (Goldman et al., 1996) and designed for high-level stable and transient expression in mammalian hosts. The vector contains a human cytomegalovirus immediate-early (CMV) promoter for high-level expression in a wide range of mammalian cells. A multiple cloning site is in the forward orientation to facilitate cloning. It contains the neomycin resistance gene for selection of stable cell lines. The vector confers resistance to ampicillin.

Plasmid purification

For purification of the plasmids used in all the routinely molecular biology applications (restriction enzyme digestion, sequencing, ligation and transformation) the commercial kit QIAprep® Spin Miniprep Kit (Qiagen, Netherlands) was used following the manufacturer's instructions.

For purification of the plasmids used in transfection (pcDNA3.1 constructions) the commercial kit EndoFree Maxi Kit (Qiagen, Netherlands) was used, following the supplier's instructions.

DNA Gel electrophoresis

Electrophoresis was performed to separate, identify, and purify DNA. 1% (m/v) agarose gels were prepared by dissolving agarose powder (Ecogen, Spain) in TAE buffer (Tris-Acetate 40 mM, EDTA 2 mM, pH 8.0). Agarose was melted in TAE by heating in a microwave. When the molten agarose had cooled down to 50°C, ethidium bromide (Sigma, USA) was added to a final concentration of 0.5 µg/ml. Gels were solidified and then submerged with TAE buffer in a gel apparatus (Bio-Rad, USA). Samples were mixed 5:1 with 6xloading buffer (Fermentas, USA) before being loaded into the gel. For size estimation, a DNA ladder (Fermentas, USA) was also loaded to the gel. An electrical current was applied to the gel tank at 110 volts (Bio-Rad, USA) until nucleic acids were appropriately separated. The gels were then visualized under an ultraviolet lamp.

Gel extraction and purification

The desired bands of DNA were excised from the agarose gels placed under UV lamp using a sterile blade. The fragments of agarose were transferred into new tubes and purified using the QIAquik® Gel Extraction Kit (Qiagen, Netherlands) following the manufacturer's instructions.

Sequencing

The different constructions were sequenced in both sense and antisense directions with oligonucleotides T7 Prom and T7 Term for the pET series vectors and T7 Prom and BGH reverse primer for pcDNA3.1-related constructions (Table 2).

DNA sequencing was carried out using the dideoxy nucleotide chain-termination method (Sanger et al., 1977) with dye terminator labelling of PCR products. The commercial kit ABI PRISM® dRhodamine Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, USA) was used.

The DNA samples were precipitated using ethanol/sodium acetate and pellets were washed with 70% ethanol. The purified templates were dissolved in deionized formamide (Bio-Rad, USA) and analyzed in an ABI Prism 310 Genetic Analyzer (Applied Biosystems, USA) at the Laboratori d'Ictiologia Genètica of the Universitat de Girona.

Site-directed mutagenesis

Site-directed mutagenesis was performed using the QuikChange™ kit (Agilent Technologies, USA) according to the supplier's instructions. The PCR reactions were treated with *Dpn*I for 1 h at 37°C before transforming 5 µl into *E. coli* strain DH5α competent cells.

Oligonucleotides

The oligonucleotides used for site-directed mutagenesis were designed according to the following standard. The length of each primer was ranged from 25-45 nucleotides with a >40% of guanine/cytosine (GC) content and ideally had one or more C or G bases at 3' end. Both of the mutagenic primers contained the desired mutation in the middle of the primers with 10-15 bases of correct sequence on both sides and annealed to the same sequence on opposite strands of the plasmid.

The expected melting temperature was estimated using the following formula:

$$T_m = 81.5+0.41\cdot(\%GC)-675/N-\%MM$$

Where T_m is the melting temperature, %GC is the percentage of G and C bases, N is the primer length in bases and %MM is the percentage of mismatch.

Table 2. Oligonucleotides used in this work.

Name	Sequence (5'-3')*	Purpose	Vector	Tm (°C)
SalIApop1	CGTATAAGACTGTAA**GTCGAC**GAATTCCC	Site directed mutagenesis to replace *Eco*RI restriction site at 3' of the Apoptin gene for *Sal*I	pcDNA3.1(+)	**62.8**
SalIApop2	GGGAATTC**GTCGAC**TTACAGTCTTATACG			
NdeIApop1	GAATTCGCC**CATATG**AACGCTCTCCAA	Site directed mutagenesis to replace *Eco*RI restriction site at 5' of the Apoptin gene for *Nde*I	pcDNA3.1(+)	**72.5**
NdeIApop2	TTGGAGAGCGTT**CATATG**GGCGAATTC			
Nco1_1	GGAGATATAC**CCATGG**ACGCTCTCC	Site directed mutagenesis to replace *Nde*I restriction site at 5' of the Apoptin gene for *Nco*I	pET-22b(+)	**69.5**
Nco1_2	GGAGAGCGT**CCATGG**GTATATCTCC			
NdeI_Pro1	CGGCCGTTGGAA**CATATG**CACTGCAGAGAG	Site directed mutagenesis to introduce a second *Nde*I restriction site after the PRS to create the variant ApopΔPro	pET-28a(+)	**62.2**
NdeI_Pro2	CTCTCTGCAGTG**CATATG**TTCCAACGGCCG			
NdeI_Leu1	GGAATTACAATC**CATATG**TCGCTGTGTGG	Site directed mutagenesis to introduce a second *Nde*I restriction site after the LRS to create the variant ApopΔProΔLeu	pET-28a(+)	**62.8**
NdeI_Leu2	CCACACAGCGA**CATATG**GATTGTAATTCC			
NdeI_1	GGACGAGCTGTACAAG**CATATG**GAGAAGCTTGCGGCCG	Site directed mutagenesis to introduce the *Nde*I restriction site after the gene of EmGFP	pRSET/EmGFP	**76.9**
NdeI_2	CGGCCGCAAGCTTCTC**CATATG**CTTGTACAGCTCGTCC			
EcoRI_1	CGGTTTCCCTCTA**GAATTC**ATTTTGTTTAACTTTAAG	Site directed mutagenesis to introduce an *Eco*RI restriction site at 5' of the genes of ApopΔpro and ApopΔProΔLeu	pET-22b(+)	**71**
EcoRI_2	CTTAAAGTTAAACAAAAT**GAATTC**TAGAGGGAAACCG			
NcoI_Leu1	GGAATTACAATCA**CCATGG**CGCTGTGTGGCTGCG	Site directed mutagenesis to replace the *Nde*I restriction site at 5' of the ApopΔProΔLeu gene for *Nco*I	pET-22b(+)	**72**
NcoI_Leu2	CGCAGCCACACAGCG**CCATGG**GTGATTGTAATTCC			
T7 Prom	TAATACGACTCACTATAGGG	Sequencing	pET series/pcDNA3.1(+)	-
T7 Term	CTAGTTATTGCTCAGCGGTG	Sequencing	pET series	-
BGH reverse	TAGAAGGCACAGTCGAGGC	Sequencing	pcDNA3.1(+)	-

*The sequence of the restriction enzyme site is indicated in bold.

$CaCl_2$ transformation

The corresponding *E. coli* strain (BL21(DE3) or DH5α) stored at -80°C in 20% glycerol (w/v) was grown overnight at 37°C in LB medium (Bertani, 1951). The culture was inoculated 1:100 in 10 ml LB and allowed to grow at 37°C with agitation until an $O.D_{550}$ of 0.6 was reached. The rapidly growing bacteria were centrifuged at 4,000 rpm and 4°C and gently resuspended in 10 ml of cold 0.1 M $CaCl_2$. After standing in ice for at least 30 min, the cells were centrifuged again and resuspended in 1 ml of cold 0.1 M $CaCl_2$. When necessary, the suspension was stored at -80°C with 20% glycerol without loss of competence. 5 µl of DNA from site-directed mutagenesis reaction or 1 µl from purified plasmid DNA was added to 0.1 ml of competent cells. After 60 min incubation on ice, the suspensions were quickly heated to 42°C for 90 s and then chilled on ice again for 10 min. They were then diluted 5 times into LB medium and incubated at 37°C for 1 h. The transformed cells were spread onto LB agar plates containing the adequate antibiotic, ampicillin or kanamycin, depending on the vector used.

Plasmid constructions

Construction of pET-22_Apop: Apoptin's gene cloned into plasmid pcDNA3.1 (Guelen et al., 2004) was a generous gift from Prof. Malvash Tavassoli (King's College, UK). Full-length Apoptin gene was amplified by PCR using two oligonucleotides that introduced *Nde*I and *Sal*I restriction sites at 5' and 3' (Table 2), respectively. The amplified DNA segment was digested with *Nde*I and *Sal*I restriction enzymes and cloned in pET-22b, which had been cleaved with the same restriction enzymes, resulting in plasmid pET-22b_Apop (Figure 4).

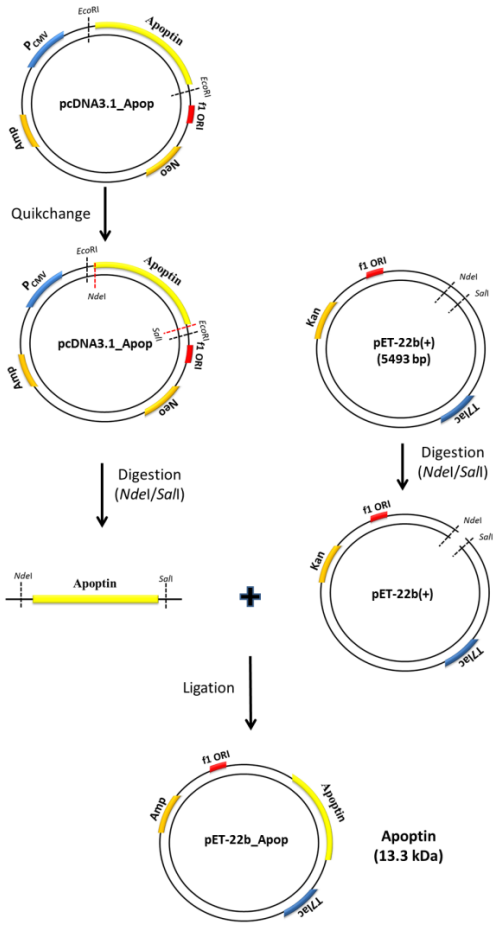


Figure 4. Schematic representation of the construction of pET-22_Apoptin.

Construction of pET-28_Apop: The DNA segment containing the Apoptin gene digested with *Nde*I and *Sal*I restriction enzymes and removed from plasmid pcDNA3.1 was cloned in pET-28a, which had been cleaved with the same restriction enzymes, resulting in plasmid pET-28_Apop. This construction contains full-length Apoptin gene fused to a sequence coding for an N-terminal hexa-histidine tag and a thrombin cleavage site and is named H_6-Apoptin (Figure 5).

Construction of pET-28_ApopΔPro and pET-28_ApopΔProΔLeu: H_6-ApopΔPro, which lacks residues 1 to 27 and H_6-ApopΔProΔLeu that lacks residues 1 to 43 were constructed from pET-28_Apop by introducing on the Apoptin gene a second *Nde*I restriction site at codons 27 and 43 (Table 2), respectively, using the QuikChange site-directed mutagenesis kit (Stratagene, USA). The resulting mutated plasmids were digested with *Nde*I to remove the desired Apoptin fragments and further ligated to create pET-28_ApopΔPro and pET-28_ApopΔProΔLeu (Figure 5).

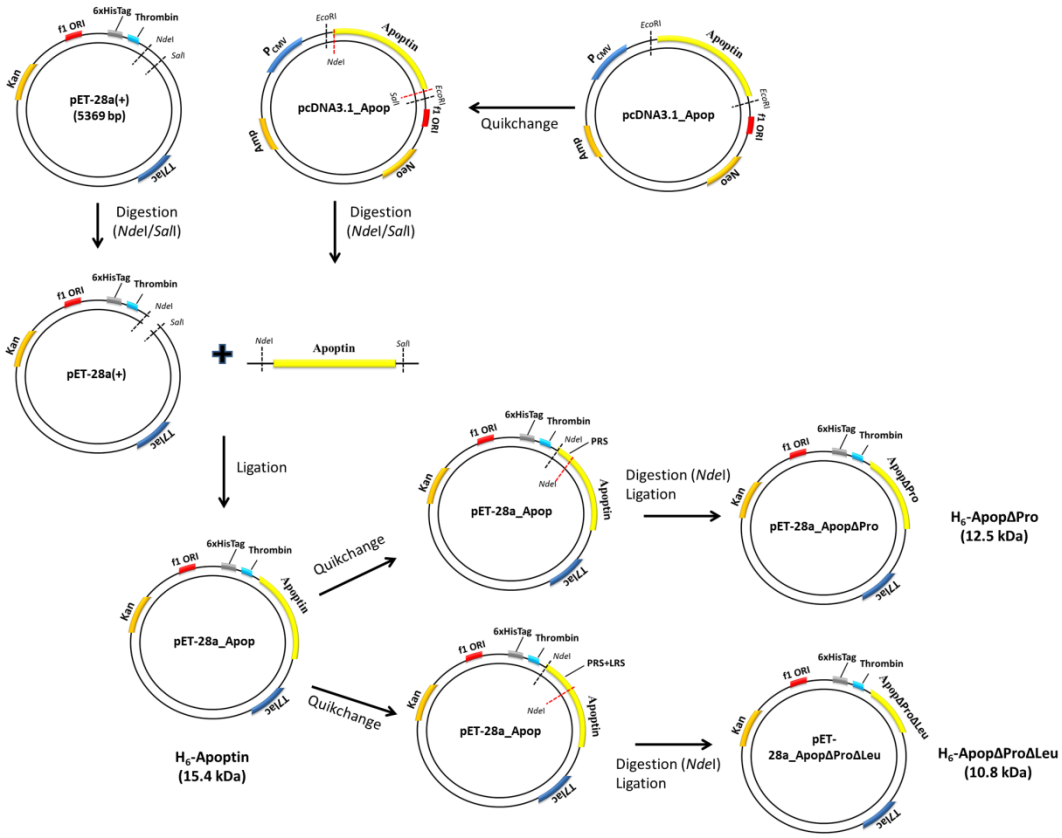


Figure 5. Schematic representation of the construction of pET-28_Apoptin, pET-28_ApopΔPro and pET-28_ApopΔProΔLeu. The recombinant proteins coded by these vectors are H_6-Apoptin, H_6-ApopΔPro and H_6-ApopΔProΔLeu, respectively (Table 3).

Construction of pET-22_ApopΔProΔLeu: The ApopΔProΔLeu gene cloned in pET-28a vector was digested with *Nde*I and *Sal*I restriction enzymes and ligated with plasmid pET-22b digested with the same restriction enzymes, resulting in pET-22_ApopΔProΔLeu.

Construction of pET-22_TEV-Apop and pET-22_TEV-ApopΔProΔLeu: The Apoptin variants' genes encoded in pET-22_Apop and pET-22_ApopΔProΔLeu were amplified by PCR using oligonucleotides that switched the *Nde*I restriction site for that of *Nco*I at 5' (Table 2). The amplified fragments were digested with *Nco*I and *Xho*I restriction enzymes and cloned into a pET-22b(+)-derived vector in which a 6 His-tag sequence followed by the cleavage site (GENLYFQGA) for the TEV protease had previously been introduced between *Nde*I and *Nco*I restriction sites (Gordo V.; unpublished PhD thesis). The resulting mutated variants contained an N-terminal 6His-tag that could be excised using the TEV protease. As a consequence of the introduction of the new restriction enzyme site (*Nco*I) Ser2 (just before the hexa-histidine tag) is replaced by Ala2.

Construction of pBADm20_Apop, pBADm30_Apop, pBADm41_Apop and pBADm60_Apop: The pBADm series of vectors incorporating the Apoptin variants were constructed digesting each vector with *Nco*I and *Xho*I restriction enzymes and ligating the Apoptin fragment already digested with the same enzymes in pET-22b plasmid. The resulting variants contained a His-tag and a fusion protein at the N-terminus of Apoptin which can be further removed by digestion of the fusion protein with TEV protease. The coded fusion proteins were TrxA-Apoptin, GST-Apoptin, MBP-Apoptin and NusA-Apoptin for pBADm20_Apop, pBADm30_Apop, pBADm41_Apop and pBADm60_Apop, respectively (Table 3).

GFP-Apoptin fusion variants: GFP variants used in the transfection experiments were constructed as follows. GFP gene from plasmid pRSET/EmGFP (Invitrogen, USA) was amplified by PCR using two

oligonucleotides that introduced an *Eco*RI and *Nde*I restriction sites at 5' and 3' (Table 2), respectively. The amplified fragment was digested with both restriction enzymes and cloned into pET-22b to create pET-22_GFP. The GFP gene was excised from pET-22_GFP using restriction enzymes *Eco*RI and *Xho*I and cloned into pcDNA3.1 digested with the same restriction enzymes to create pcDNA3.1_GFP control plasmid (figure 6). To create the GFP-Apoptin variants, the Apoptin genes from pET-28_Apop, pET-28_ApopΔPro or pET-28_ApopΔProΔLeu were excised by digesting with restriction enzymes *Nde*I and *Xho*I. The GFP gene was also excised from pET-22_GFP by cleavage with restriction enzymes *Nde*I and *Eco*RI. The GFP gene was cloned together with the excised genes of each Apoptin variants into pcDNA3.1 previously digested with *Eco*RI and *Xho*I to create plasmids pcDNA3.1_Apop, pcDNA3.1_ApopΔPro or pcDNA_3.1_ApopΔProΔLeu (figure 6).

All constructions were confirmed by DNA sequencing of both strands.

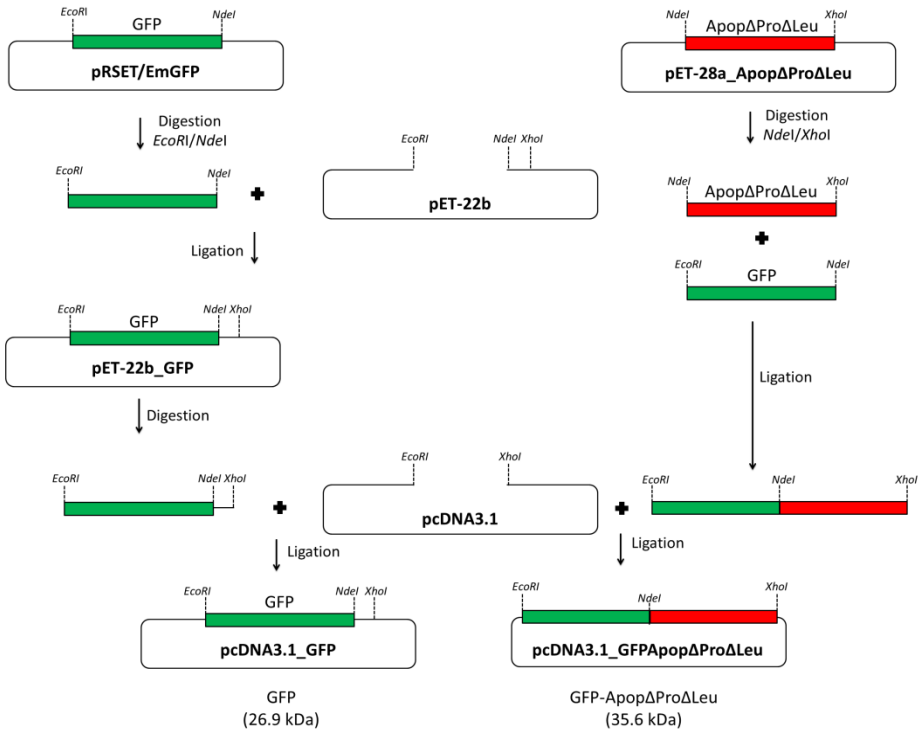


Figure 6. Schematic representation of the construction strategy of pcDNA3.1_GFP (control) and pcDNA3.1_GFP-ApopΔProΔLeu by restriction enzyme digestion and ligation.

Protein expression and purification

BL21 (DE3) cells transformed with the desired vectors were grown at 37°C in LB medium containing 100 µg/mL kanamycin until an $O.D_{550}$ of 0.8 was reached. Protein expression was induced by adding IPTG to a final concentration of 1 mM and the culture was incubated for 5 h. Cells were harvested by centrifugation and stored at -20°C until further processing.

Frozen induced cell pellets were resuspended in 30 mL of buffer containing 50 mM TRIS base, 100 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF) and the pH was adjusted to 7.4 using HCl for H_6-Apoptin and H_6-ApopΔPro variant or in 30 mL of a more acidic buffer containing 50 mM acetic acid/sodium acetate, 1 mM PMSF pH 5.0 for H_6-ApopΔProΔLeu variant. Resuspended pellets were lysed using a French Press (SLM-

Aminco, Spectronic Instruments, UK) and inclusion bodies were harvested by centrifugation. These aggregates were solubilized in 10 mL of 50 mM TRIS base, 500 mM NaCl, 7 M Urea pH 7.4 (using HCl). Reduced glutathione (GSH) 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 h under N_2 atmosphere to completely solubilize Apoptin and its variants. The remaining insoluble material was removed by centrifugation at 10,000 rpm at 4°C for 30 min. The cleared supernatant was incubated overnight at room temperature and gentle agitation with 15 ml (7.5 ml bed volume) Ni-NTA agarose (Qiagen, Netherlands) previously equilibrated with solubilization buffer.

After overnight incubation, the resin was loaded onto a column and washed with 90 ml of 50 mM TRIS base, 500 mM NaCl, 7 M Urea, 20mM Imidazole, pH 7.4 (adjusted with HCl) to remove unbound proteins. The protein was eluted from the column with 90 ml of 50 mM TRIS base, 500 mM NaCl, 7 M Urea, 500 mM Imidazole pH 7.4 (adjusted with HCl) and 9 ml fractions were collected. All fractions were then analyzed by SDS-PAGE and those containing the desired protein were gathered. The sample was diluted dropwise tenfold into 50 mM TRIS base, 200 mM NaCl, 100 mM L-Arg, 100 mM L-Glu, 1 mM oxidized glutathione (GSSG), 1 mM GSH, 1 mM EDTA, pH 8.1 (adjusted with HCl), and further incubated for 24 h at 4°C to allow protein refolding. The sample pH was adjusted to 5.0 with HCl, concentrated to 150 mL by tangential ultrafiltration (Millipore, USA), dialyzed at 4°C against ddH_2O and lyophilized.

Apoptin sample still presented some impurities that had to be removed by size exclusion chromatography (SEC). Next, the lyophilized sample was resuspended in 50 mM TRIS base, 150 mM NaCl, 1 mM EDTA, pH adjusted to 7.4 by adding HCl and further purified on a Superdex-75 10/300GL (GE

Healthcare, UK) column. The collected sample was dialyzed at 4°C against ddH_2O water and lyophilized.

The molecular mass of each variant was confirmed by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry using a Bruker-Biflex spectrometer (Bruker, USA) at the Research Technical Services of the Universitat de Girona.

For multiple dimensional heteronuclear NMR spectroscopy, the variants' expression was induced in minimal media containing ^{13}C-glucose and ^{15}N ammonium chloride (Cambridge Isotope Laboratories, USA) as the only source of carbon and nitrogen atoms, respectively, following previously published protocols (Marley et al., 2001). The degree of ^{13}C and ^{15}N incorporation was judged by comparing the theoretical mass of the completely isotopically labeled protein to the mass determined experimentally using MALDI-TOF mass spectroscopy at the Instituto de Química Física "Rocasolano" (CSIC) of Madrid.

H_6-TEV protease production and purification

One liter of *E. coli* BL21(DE3) cells previously transformed with the plasmid containing H_6-TEV protease gene (kindly provided by G. Stier, EMBL, University of Heidelberg, Germany) was cultured until an $O.D_{550}$ of 0.5 was reached, and then induced using IPTG to a final concentration of 1 mM at 18°C overnight. Cells were harvested by centrifugation at 10,000 rpm, at 4°C for 7 min. The cell pellet expressing H_6-TEV protease was resuspended in 30 ml of lysis buffer (25 mM TRIS base, 150 mM NaCl, 1 mM PMSF pH 8.0, adjusted with HCl) containing 5 mM Imidazole and lysed using a French Press. The soluble fraction was collected by centrifugation at 10,000 rpm at 4°C for 45 min and incubated for 2 h at room temperature and gentle agitation with 10 ml (5 ml bed volume) of Ni-NTA agarose previously equilibrated with lysis buffer and loaded to a column. The column was washed with two volumes of lysis buffer, and then with two extra volumes

of lysis buffer containing 10 mM imidazole. The sample was eluted using four volumes of lysis buffer with 250 mM imidazole and 10% glycerol. The purified protein was then dialyzed at 4°C against 25 mM TRIS base, 150 mM NaCl, 10% glycerol, pH 8.0 (adjusted with HCl). Finally, H_6-TEV protease was stored at -80°C in aliquots after adjusting the protein concentration at 2 mg/ml and increasing the glycerol concentration to 50%.

TEV digestion

Digestion with TEV protease was carried out in order to remove the fusion protein used to solubilize Apoptin expressed in any of the pBADm series of vectors and to excise the His-tag of ApopΔProΔLeu expressed using the pET-22b-derived plasmid (Figure 7). The digestion was performed at a ratio of 1:50 (mg TEV protease : mg protein) in a buffer containing 50 mM TRIS base, 100 mM NaCl pH 7.4 (adjusted with HCl), overnight at 20°C or alternatively at 34°C during 4 h, withdrawing samples at 0, 0.5, 1, 2, 3, and 4 h to analyze them by SDS-PAGE and MALDI-TOF.

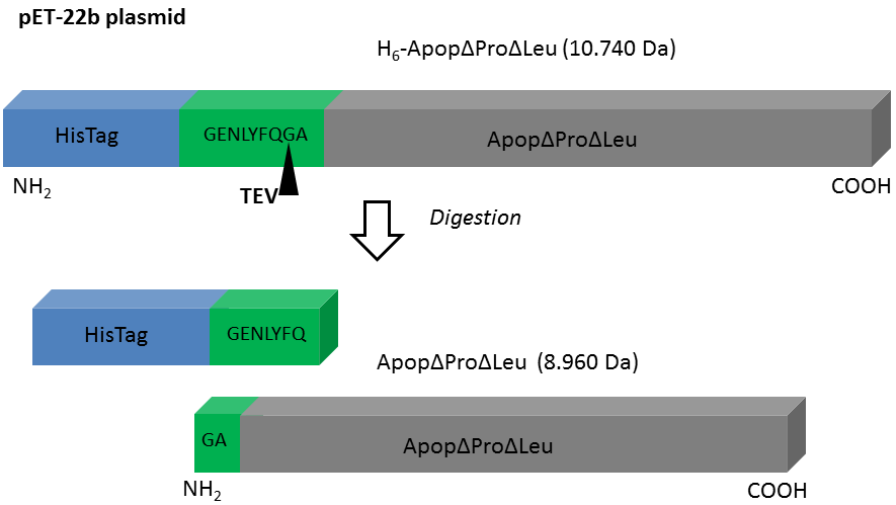


Figure 7. Schematic representation of the N-terminus His-tag excision of H_6-ApopΔProΔLeu expressed in pET-22b-derived plasmid using TEV protease.

Thrombin digestion

Digestion with thrombin (Sigma, USA) was performed to cleave the His-tag of H_6-ApopΔProΔLeu expressed using the pET-28a plasmid (Figure 8). The digestion was carried out at 25°C at a ratio of 2 NIH (Gaffney and Edgell, 1995) of thrombin per 1 mg of protein in a buffer containing 50 mM TRIS base, 150 mM NaCl, 2.2 mM $CaCl_2$, 0.1% β-mercaptoethanol pH 8.0 (adjusted with HCl). The reaction was terminated at 0, 10', 15', 20', 25', 30', 40', 50', and 60' by the addition of PMSF at a final concentration of 1 mM. The cleavage was followed by SDS-PAGE and MALDI-TOF.

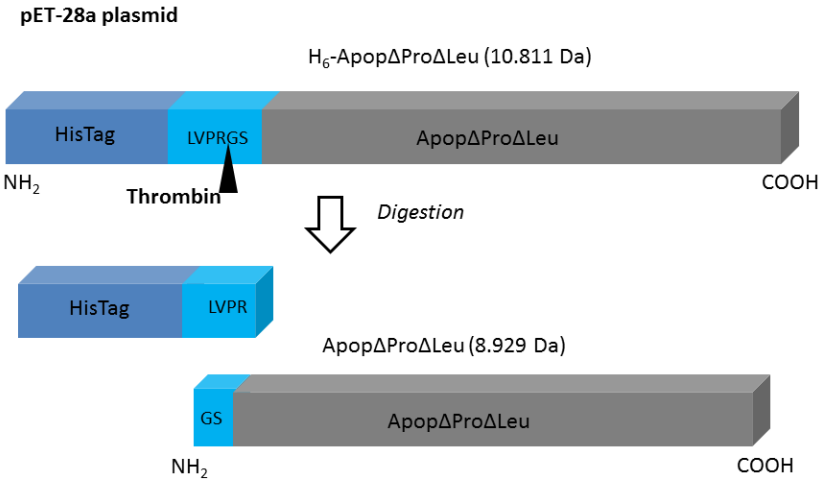


Figure 8. Schematic representation of the N-terminus His-tag excision of H_6-ApopΔProΔLeu expressed in pET-28a plasmid using thrombin

Table 3. Apoptin constructs.

Plasmid Name	Encoded protein	Mw/pI	Tag excision	Mw/pI after excision
pET-22_Apop	Apoptin (1-121)	13.3/9.95	-	-
pET-22_ApopΔProΔLeu	Apoptin (45-121)	8.6/10.14	-	-
pET-28_Apop	H_6- LVPRGSH-Apoptin (1-121)	15.4/10.14	Thrombin	13.5/9.95
pET-28_ApopΔPro	H_6- LVPRGSH-Apoptin (29-121)	12.5/10.36	Thrombin	10.7/10.15
pET-28_ApopΔProΔLeu	H_6- LVPRGSH-Apoptin (45-121)	10.8/10.36	Thrombin	8.9/10.14
pET-22-TEV_Apop	H_6-GENLYFQGA-Apoptin (1-121)	15.4/9.69	TEV	13.4/9.95
pET-22_TEV_ApopΔProΔLeu	H_6-GENLYFQGA-Apoptin (45-121)	10.7/9.85	TEV	8.7/10.14
pBADm-20_Apoptin	H_6-TrxA-GENLYFQGA-Apoptin (1-121)	27.6/7.67	TEV	13.4/9.75
pBADm-30_Apoptin	H_6-GST-GENLYFQGA-Apoptin (1-121)	43/8.47	TEV	13.4/9.75
pBADm-41_Apoptin	H_6-MBP-GENLYFQGA-Apoptin (1-121)	61.8/9.56	TEV	13.4/9.75
pBADm-60_Apoptin	H_6-NusA-GENLYFQGA-Apoptin (1-121)	70.5/4.90	TEV	13.4/9.75
pCDNA3.1_GFP	GFP	26.9/5.84	-	-
pcDNA3.1_GFPApop	GFP-Apoptin (1-121)	40.3/8.10	-	-
pcDNA3.1_GFPApopΔPro	GFP-Apoptin (29-121)	37.4/8.38	-	-
pcDNA3.1_GFPApopΔProΔLeu	GFP-Apoptin (45-121)	35.6/8.14	-	-

Bradford assay

The Bradford assay was used to determine the total protein concentration of a sample (Bradford, 1976). This method is based on the shift of the maximum absorption from 465 nm to 595 nm of the dye Coomassie Brilliant Blue G-250 (Bio-Rad, USA) in acidic solution when bounded to protein. The standard curve was prepared using BSA (Sigma, USA) concentrations ranging between 2.5 and 20 µg/ml.

Determination of extinction coefficient of Apoptin variants

The extinction coefficient of H_6-Apoptin and its variants was determined experimentally. A linear regression equation was calculated that related different Apotin variant protein concentrations calculated upon Bradford assay (see above) and its corresponding absorbance at 280 nm. The extinction coefficient of H_6-Apoptin was 1,989 M^{-1} cm^{-1} and for its variants H_6-ApopΔPro and H_6-ApopΔProΔLeu was 1,856 M^{-1} cm^{-1} and 3,775 M^{-1} cm^{-1}, respectively. The extinction coefficients were subsequently used to quantify protein concentration.

SDS-Polyacrylamide gel electrophoresis

The analysis of protein by SDS-polyacrylamide gel electrophoresis (PAGE) was performed using the buffer system of Laemmli (Laemmli, 1970) in 15% polyacrylamide gels (0.75 mm or 1.5 mm thickness). Protein samples were prepared in 2% SDS, 10% glycerol, 5% 2-mercaptoethanol, 0.005% bromophenol blue, 50 mM TRIS base/HCl, pH 6.8 (loading buffer) followed by boiling at 95°C for 10 min prior loading them into the gel.

Polypeptide bands and molecular mass standards (Thermo Fisher Scientific, USA) were stained with 0.02% (w/v) Coomassie Brilliant Blue R-250 in 10% acetic acid.

Native Gel electrophoresis

Native PAGE for proteins was performed in continuous buffer according to McLellan (McLellan, 1982). The buffer of choice was 43 mM Imidazole (basic component), 35 mM HEPES (acidic component), pH 7.4. A variant of the Human Pancreatic Ribonuclease (HP-RNase) named NLSPE5 (Vert et al., 2012) was used as a control (isoelectric point: 9.40). Protein samples were prepared in 10% (w/v) glycerol, 0.01% methyl green.

Polypeptide bands were stained with 0.02% (w/v) Coomassie Brilliant Blue R-250 in 10% acetic acid.

Thermal Stability Studies

To determine the conformational stability of the H_6-ApopΔProΔLeu variant, its UV absorbance at 280 nm using a Perkin-Elmer LS-50B spectrophotometer (Perkin-Elmer, USA) was monitored during heating (from 18°C to 70°C) as previously described (Font et al., 2006). Several experiments were performed on 1 mg/mL H_6-ApopΔProΔLeu solutions using a sealed cuvette and the following buffers: 50 mM glycine pH 3.2, 50 mM sodium bicarbonate pH 4.5, 50 mM sodium acetate pH 5.0, 50 mM MES

pH 6.0, 50 mM MOPS pH 7.0, 50 mM TRIS base adjusted to pH 8 with HCl and 50 mM glycine pH 9.2. For each buffer, the UV signal versus temperature of a blank was subtracted.

Thioflavin T (Th T) Fluorescence Experiments

The fluorescence emission and excitation spectra of Th T incubated with Apoptin protein were recorded using a Perkin-Elmer LS-50B fluorimeter (Perkin-Elmer, USA). The experiment was performed as previously described (Nilsson, 2004) with minor modifications. Th T was assayed in PBS at a final concentration of 15 µM and Apoptin was at a final concentration of 65 µM. Blank was made by registering free Th T solution.

Congo Red (CR) staining

CR was prepared in 50 mM TRIS base, 100 mM NaCl pH 7.4 (adjusted with HCl) and in 50 mM glycine pH 9.2 buffer and filtered through 0.22 µm syringe filters. The assay was performed as previously described (Nilsson, 2004) with minor modifications. CR was prepared at a final concentration of 100 µM, and it was incubated with the Apoptin variant (65 µM) for 30 min at room temperature. The absorbance spectra were registered from 400 to 650 nm wavelength on a Perkin Elmer Lambda Bio 20 spectrophotometer (Perkin-Elmer, USA).

1-anilinonaphthalene-8-sulfonic acid (ANS) binding

The fluorescence emission spectra of ANS with and without protein were recorded over a range from 400-600 nm (slit width 10 nm) with an excitation wavelength of 350 nm (slit width 5 nm) on a LS-50B fluorimeter. Protein samples were prepared at a final concentration of 1 mg/ml in 50 mM TRIS base adjusted with HCl to pH 8.0. ANS was filtered through 0.22 µm syringe filters, added to the protein solution at a final concentration of 50 µM and incubated for 20 min at room temperature before the measurements.

Transmission Electron Microscopy (TEM)

H_6-Apoptin was prepared at a concentration of 1 mg/ml in 50 mM TRIS base adjusted to pH 7.0 with HCl and incubated for 72 h at 37°C. Negative stained samples were prepared on carbon-coated grids with 2% (w/v) uranyl acetate solution (Gras et al., 2011). TEM images were obtained using a Zeiss EM910 microscope (Zeiss, Germany) at 60 kV at the Research Technical Services from the Universitat de Girona.

Dynamic Light Scattering (DLS)

DLS measurements were recorded at 532 nm on a Malvern ZetaSizer Nano ZS (Malvern Instruments, UK) at 25°C in PBS and a final protein concentration of 1 mg/ml at the Scientific and Technical Services of the Universitat Autònoma de Barcelona. The hydrodynamic radius of the particle (R_H) was determined by averaging the results of at least three independent experiments. An upper estimate of the protein molecular mass was obtained from R_H value using the empirical relationship for typical globular proteins MM (kDa) = $(1.68xR_H)^{2.34}$ using the supplied software (Friedeberg et al., 2006).

Resistance to Proteinase K

All proteins were prepared to a final concentration of 10 µM in 50 mM TRIS base/HCl, 150 mM NaCl pH 8.0. The digestion was carried out using proteinase K at a final concentration of 0.35 µM during 1 h at room temperature. PMSF (1.5 mM) was used to stop the digestion, and the samples were then analyzed by SDS-PAGE (15% polyacrylamide).

Size Exclusion Chromatography (SEC)

SEC was performed at room temperature on a Superdex 75 10/300 LG (GE Healthcare, USA) connected to AKTA Purifier UPC-10 system (GE Healthcare, USA) and run at 0.5 ml/min. The injection volume was 500 µl

of protein samples concentrated up to 2 mg/ml. The column was calibrated with proteins of known molecular weight (ICN Biomedicals, USA): cytochrome c (13 kDa), myoglobin (18 kDa), chymotripsinogen A (24 kDa), ovalbumin (45 kDa) and BSA (67 kDa). In order to establish the quantitative relationship between elution volume and molecular weight a linear regression equation was calculated (square regression coefficient (r^2) of 0.981).

Far-Ultraviolet Circular Dichroism (far-UV CD)

A JASCO J810 instrument (JASCO, Japan) equipped with a Peltier temperature control module was used to record CD spectra at 5°C or 25°C in the far UV. Twelve accumulations spanning a wavelength range of 190-260 nm were recorded at 0.5 nm steps for each spectrum. The bandwidth was set to 1.2 nm and the scan speed was 50 nm per min. All spectra were recorded in a 0.1 cm path length cuvette in 3 mM potassium hydrogen phosphate buffer. The concentration of protein was 15 µM. Percentage of secondary structure elements was estimated using the Selcon3 algorithm (Sreerama et al., 1999).

NMR Spectroscopy

The protein sample was normally prepared in 10 mM deuterated sodium acetate/deuterated acetic acid pH 5.2 buffer containing NaN_3 at a concentration of 1 mM. All samples contained 10% D_2O and 50 µM of sodium 2,2-dimethyl-2-silapentane-5-sulphonate (DSS) as an internal proton chemical shift reference. An identical sample, but in fully deuterated solvent (100% D_2O) was also prepared for the hydrogen-deuterium exchange experiment. Additional samples at pH 6.7 in 10 mM potassium hydrogen phosphate/potassium di-hydrogen phosphate, where His residues are mostly neutral, were also prepared.

All NMR spectra were recorded at 5°C on a Bruker AV-800 spectrometer (Bruker BioSpin, USA) operating at 18.8 T (800 MHz of ^{1}H Larmor frequency), equipped with a triple resonance cryoprobe and Z-gradients. Spectra were processed with TopSpin NMR (Bruker, USA) and analyzed using Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). ^{13}C and ^{15}N resonances were referenced indirectly to the ^{1}H signal of DSS using nuclei's gyromagnetic ratios (Markley et al., 1998). For labeled ^{13}C-^{15}N H_6-ApopΔProΔLeu, assignment of ^{1}H, $^{13}C'$, $^{13}C\alpha$, $^{13}C\beta$ and ^{15}N resonances was achieved using a standard suite of heteronuclear 2D and 3D spectra; namely: ^{1}H-^{15}N-heteronuclear single quantum coherence (HSQC) and triple resonance three-dimensional spectra HN(CO)CA, HNCA, CBCA(CO)NH, CBCANH, HNCO, (CBCACO)NH, HNCO, and HNCOi. Side-chain amide groups were assigned using the 2D ^{1}H-^{13}C-HSQC and 3D HBHA(CO)NH, HCC(CO)NH, CCC(CO)HN spectra. These assignments were corroborated by using ^{1}H-^{15}N HSQC spectra which are ^{13}C-filtered into several classes based on the HN's side chain or the side chain of the preceding residue (Pantoja-Uceda and Santoro, 2008, 2012). Residues preceded or followed by Pro (which lacks an HN group) were verified by specialized NMR experiments (Schubert et al., 2000). As an additional validations, and to completely assign the resonances arising from the His-tag, a (4, 2) D $^{1}H^{15}N/^{1}H^{15}N$ NMR experiment was recorded and analyzed.

The backbone and $^{13}C\beta$ conformational chemical shift values were calculated by subtracting the corresponding values measured in random coil conformers from the experimentally measured values (Kjaergaard et al., 2011) or alternatively, by subtracting the corresponding values measured in a set of intrinsically disorder proteins from the experimental measurements (Tamiola et al., 2010).

All the assignations were deposited in the Biological Magnetic Resonance Bank (BMRB, www.bmrb.wisc.edu) (Ulrich et al., 2008) under accession code BMRB #26736 and are annexed to this work.

Hydrogen/Deuterium (H/D) Exchange Measurements

Approximately 3 mg of lyophilized ^{13}C, ^{15}N H_6-ApopΔProΔLeu were dissolved in 0.20 mL of D_2O with 10 mM deuterated sodium acetate/deuterated acetic acid which had been precooled to 5°C. After briefly shimming and tuning the magnet, H/D exchange was followed by recording a series of ^{1}H-^{15}N HSQC spectra at 5°C. Although the first spectrum was recorded within just 15 min after dissolving the protein, most of the HN groups had already exchanged to DN, indicating a general absence of stable secondary structures. The exchange rates of 17 signals could be determined and these were compared to the intrinsic exchange rates expected for a structureless short peptide under the same conditions of pH* (5.6) and temperature (5°C) to determine the protection factors (Bai et al., 1993).

^{15}N Relaxation Measurements and Analysis

^{15}N longitudinal relaxation rate (R_1), ^{15}N transverse relaxation rate (R_2) and heteronuclear {^{1}H}-^{15}N NOE spectra (Farrow et al., 1994) for 1 mM uniformly ^{15}N-labeled H_6-ApopΔProΔLeu were acquired at 298 K on a Bruker Avance 800 MHz NMR spectrometer (Bruker, USA). Eleven delays (20, 60, 140, 240, 360, 460, 660, 860, 1200, 1600, and 2000 ms) were used for R_1 measurements, and a different set of nine delays (15.6, 31.4, 62.4, 112, 156, 218.4, 265.2, 405.6, and 514.8 ms) used to measure the R_2 values. {^{1}H}-^{15}N NOE experiments were carried out with an overall recycling delay of 10 s to ensure the maximal development of NOEs before acquisition and to allow solvent relaxation, thus avoiding transfer of saturation to the most exposed amide protons of the protein between scans (Renner et al., 2002).

The experiment was performed twice, using a conventional and cryogenic probe, and consistent results were obtained.

Relaxation times were calculated via least-squares fitting of peak intensities to a two-parameter exponential function, using the rate analysis routine of the Java version of NMRView (Johnson, 2004). Heteronuclear NOEs were calculated from the ratio of cross-peak intensities in spectra collected with and without amide proton saturation during the recycle delay. Uncertainties in peak heights were determined from the standard deviation of the distribution of intensities in the region of the HSQC spectra where no signal and only noise was observed.

Apoptin phosphorylation

1 mg of H_6-ApopΔProΔLeu was phosphorylated by incubating the protein with 0.5 µg of the Serine/Threonine CDK2/Cyclin A2 kinase (Sigma, USA) in 500 µl of 5 mM MOPS pH 7.2, 1 mM EGTA, 0.4 mM EDTA, 5 mM $MgCl_2$, 1 mM 3,3',3''-phophanetriyltriproponic acid (TCEP) and 250 µM ATP. Since ^{15}N-^{13}C-labeled samples give relatively broad mass spectroscopy peaks, unlabeled protein was used to facilitate the monitoring of the reaction by MALDI-TOF mass spectrometry. The phosphorylation reaction was carried out at 25°C and was monitored by non-denaturing PAGE and by MALDI-TOF mass spectrometry analyses of aliquots withdrawn at various times. After 90 min the reaction was complete and the phosphorylated protein was transferred to 10 mM sodium acetate/acetic acid buffer pH 5.2 via ultrafiltration for characterization using natural abundance ^{1}H-^{13}C HSQC NMR spectroscopy.

Eukaryotic cell lines and culture conditions

NCI-H460 human large cell lung cancer cell line was obtained from the Frederick DCTD tumor cell line repository (National Cancer Institute, USA). CCD-18Co human colon fibroblast cell line and Jurkat human T-cell

lymphocyte cell line were acquired from Eucellbank (Universitat de Barcelona, Spain). 1BR.3.G human skin fibroblast cells were obtained from European Collection of Cell Cultures (ECACC, UK). K-562 cell line (human chronic myelogenous leukemia) was a generous gift of Dr. Bruno Beaumelle (CPBS, France).

NCI-H460, K-562 and Jurkat cells were routinely grown in RPMI supplemented with 10% fetal bovine serum (FBS), 50 U/ml penicillin, 50 µg/ml streptomycin and 2 mM L-glutamine while CCD-18Co and 1BR.3.G cells were grown in DMEM supplemented with 10% FBS, 50 U/ml penicillin, 50 µg/ml streptomycin and 2 mM L-glutamine. All media and reagents used to culture the different cell lines were from Lonza, Switzerland. Cells were maintained at 37°C in a humidified atmosphere with 5% CO_2, propagated following established protocols and remained free of mycoplasma throughout the experiments.

Cell Transfection

1BR.3.G and NCI-H460 cells were transfected with 0.5 µg DNA on 24 well plates using Lipofectamine® 3000 (Invitrogen, USA) according to the specifications of the manufacturer. Briefly, 24 h prior to transfection cells were seeded at the appropriate density: 7 x 10^4 cells per well for 1BR.3.G and 3.5 x 10^4 cells per well for NCI-H460 cells. 1.5 µl of Lipofectamine® 3000 Reagent were diluted in 25 µl Opti-MEM Medium (Life Technologies, USA). The DNA was diluted in 25 µl Opti-MEM Medium, and then 1 µl of P3000 Reagent was added. The diluted DNA was added to each tube of diluted Lipofectamine 3000 Reagent (1:1 ratio), incubated for 5 min at room temperature. Finally the DNA-lipid complex was added to the cells.

CCD-18Co cells were seeded into 6-well plates 24 h prior to transfection at a density of 1.5 x 10^5 cells per well and transfected following the protocol described by M. Zhang and coworkers (Zhang et al., 2007) with minor modifications. Briefly, 3.75 µl of Lipofectamine® 3000 Reagent were diluted

in 125 µl of Opti-MEM Medium. In parallel 2.5 µg of the corresponding DNA were diluted in 125 µl of Opti-MEM Medium, and then 5 µl of P3000 reagent were added. The diluted DNA was added to each tube of diluted Lipofectamine 3000 Reagent (1:1 ratio) and incubated for 5 min at room temperature. On the day of transfection, cells were trypsinized and cell pellets collected by centrifugation at 1,600 rpm for 5 min at room temperature. Then, each pellet (from one well of cells) was resuspended in 250 µl of the transfection media and the suspension incubated at room temperature for 15 min. At the end of incubation, 2 ml of pre-warmed fresh cell growth medium and the suspension were plated onto a new well of a 6-well plate.

HeLa cells were seeded into 6-well plates 24 h prior to transfection at a density of 1.5 x 10^5 cells per well and transfected using the calcium phosphate method (Kingston et al., 2003). The medium was replaced by fresh medium 3 h prior to transfection. For each well a mix of 1.5 µg DNA and 12.2 µl $CaCl_2$ was prepared. This solution was then brought to 100 µl with ddH_2O and diluted dropwise 1:2 with 100 µl 2xHBS buffer (274 mM NaCl, 10 mM KCl, 1.4 mM Na_2HPO_4, 15 mM D-glucose, 42 mM HEPES). The solution was incubated for 30 min at room temperature and then added dropwise to the cells.

Cell proliferation assays.

To investigate the effect of Apoptin variants upon DNA transfection, NCI-H460 cells were seeded into 24-well plates at 3.5 x 10^4 cells per well. After 24 h of incubation, cells were transfected as described above and further incubated for 48 h. Then attached and floating cells were harvested by centrifugation at 460 x g for 10 min at 4°C and washed twice in cold PBS and subjected to propidium iodide (PI) staining at room temperature for 15 min in the dark. Stained cells expressing GFP or GFP-fused proteins were analyzed on FACSCalibur flow cytometer (BD Biosciences, US) using

CellQuest Pro software. A minimum of 10,000 cells within the gate region were analyzed. Data are reported as the mean ± standard error (SE) of at least three independent experiments.

To investigate the effect of H_6-Apoptin and H_6-ApopΔProΔLeu protein variants when added exogenously, cells were seeded into 96-well plates at the appropriate density, i. e., 1,900 for NCI-H460, 10,000 for NCI/ADR-RES, 3,000 for 1BR.3.G, 4,000 for CCD-18Co, 6,000 for Jurkat and 4,500 for K-562 cells. After 24 h of incubation, cells were treated with various concentrations of H_6-Apoptin or H_6-ApopΔProΔLeu for 72 h. Cell viability was determined by the MTT (Sigma, USA) method as previously described (Castro et al., 2012). The IC_{50} value represents the concentration of the assayed protein required to inhibit cell proliferation by 50% compared to untreated cells and, in each case, it was calculated by linear interpolation from the obtained growth curves. All data are reported as the mean ± SE of at least three independent experiments with three replicas in each.

Caspase activation assay

Caspase-3, -8 and -9 catalytic activities were measured using the APOPCYTO Caspase-3, -8 and -9 colorimetric assay kits (MBL, Japan) following the manufacturer's instructions. The assay is based on cleavage of the chromogenic substrates, DEVD-pNA, IETD-pNA and LEHD-pNA, by caspases-3, -8 and -9, respectively. Briefly, NCI-H460 cells (3.5 x 10^5 cells/60-mm dish) were incubated with 15 µM H_6-ApopΔProΔLeu for 24, 36 and 48 h in serum-starved medium. Then, attached and floating cells were harvested at 460 x g for 10 min at 4°C and washed twice in cold PBS. The cells were then lysed, centrifuged and the pellet was discarded. The supernatant was recovered, and the protein concentration was determined using the Bradford protein assay (Bio-Rad, USA). Afterwards, 10 µl of the cell lysate corresponding to 20 µg of total protein, 10 µl of 2x reaction buffer containing 10 mM (2S,3S)-1,4-bis(sulfanyl)butane-2,3-diol

(DTT), and 1 µl of the 10 mM DEVD-*p*NA, IETD-*p*NA or LEHD-*p*NA substrates were mixed. To confirm the specific hydrolysis of substrate, independent samples were also treated with the caspase inhibitors z-DEVD-FMK, z-IETD-FMK or z-LEHD-FMK, which are specific inhibitors of caspases-3, -8 and -9, respectively. The reactions were then incubated at 37°C for 4 h. The reactions were monitored by changes in absorbance at 405 nm. The results are reported as the mean ± SE of three independent experiments.

Phosphatidylserine exposure assay

Quantitative analysis of NCI-H460 apoptotic cell death caused by treatment with 15 µM H_6-ApopΔProΔLeu was performed by flow cytometry using the Alexa Fluor 488 annexin V/PI Vybrant Apoptosis Assay Kit (Molecular Probes, USA) as described previously (Castro et al., 2011). Again, the results are reported as the mean ± SE of at least three independent experiments.

Confocal microscopy

NCI-H460, CCD-18Co, HeLa and 1BR.3.G cells were seeded on coverslips and allowed to attach overnight. These cells were then transfected as detailed above and incubated at 37°C in a 5% CO_2 atmosphere for 24 h. Cells were fixed with 4% paraformaldehyde in PBS for 15 min at 4°C. The cell nuclei were stained with 2 µg/ml Hoechst 33258 (excitation/emission: 352 nm/461 nm) (Thermo Fisher Scientific, USA), over 15 min at room temperature. In order to discard dead cells in the intracellular localization studies, samples were washed twice with cold PBS, whereas for the study of apoptotic activity the coverslips were directly mounted. In both cases a fluorescence-mounting medium (Dako, USA) was used. All samples were then examined using a Nikon inverted Eclipse Ti microscope with built-in NIS-Elements Advanced Research (AR) software (Nikon, Japan) at the Research Technical Services from the Universitat de Girona.

DNA binding assay

A fragment of 1,135 base-pairs of dsDNA corresponding to the gene of GFP fused to Apoptin at a concentration of 25 µg/ml was incubated in the presence of either 300 µg/ml of H_6-Apoptin or its two variants, H_6-ApopΔPro and H_6-ApopΔProΔLeu, at room temperature for 20 min in PBS. After incubation, the mixture was analyzed by agarose gel electrophoresis (1% w/v) in TAE buffer and stained with ethidium bromide. Binding to DNA was detected by the retardation of the DNA migration in the electrophoresis.

Results and Discussion

Production and purification of Apoptin

In order to characterize the biophysical and biological properties of Apoptin, procedures for the production and purification of this protein were optimized. Details of the construction of the different variants specified here can be found in the Material and Methods section. Regarding the production process, I first attempted to produce it as a soluble protein. The Apoptin gene was subcloned from the plasmid pcDNA3.1 to the pET-22b vector using the *Nde*I and *Sal*I restriction enzymes (Figure 4). This vector enables the export of soluble proteins to the periplasm, which is a more favorable environment for folding and disulfide bond formation (Rietsch et al., 1996). Apoptin contains four cysteines in its primary sequence that might be important for its proper folding and activity. After lysing the cells through the French Press, I observed by denaturing electrophoresis that the protein was roughly equally distributed between the soluble and insoluble fractions (Figure 9).

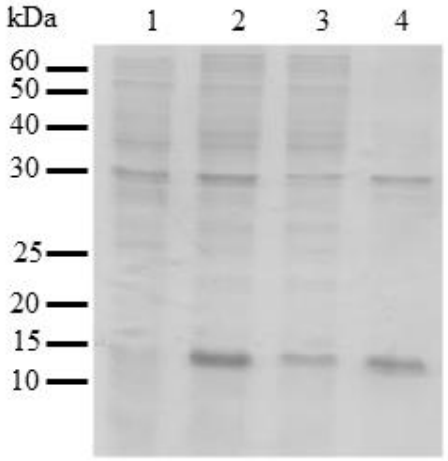


Figure 9. SDS-PAGE analysis of the production and initial purification of Apoptin: *E. coli* culture extracts before (1) and 3 h after the addition of IPTG (2); soluble (3) and insoluble (4) material from the French press. Equivalent amounts of the induced culture were loaded in each lane.

In order to increase the amount of protein in the soluble fraction, the induction was carried out at lower temperatures. One feature described for the T7 system is that many recombinant proteins often become insoluble when expressed at 37°C, but are soluble when expressed at lower temperatures presumably because lower rates of protein production allow newly synthesized recombinant proteins to fold properly (Vera et al., 2007). Additionally, this strategy reduces or impairs protein degradation due to poor activity of heat shock proteases that are usually induced during protein overproduction in *E. coli* (Chesshyre and Hipkiss, 1989). However, when the temperature was decreased to 25°C (cells induced for 6 h) or to 18°C (cells induced overnight) the amount of protein was not substantially increased in the soluble fraction (data not shown). Initial purification assays of this construction were unsatisfactory since the yield of pure protein was very low.

A widely used strategy to avoid aggregation of foreign proteins produced in *E. coli* consists in its fusion to very stable and soluble proteins that enhances the solubility of their fusion partners (for a review see (Costa et al., 2014)). In addition, the fused stretches can help in the purification process. With this aim, the Apoptin gene was fused to the genes of the TrxA, GST, MBP and NusA, using the series of vectors pBADm (see Material and Methods). Between the different genes and that of Apoptin, a sequence coding for a digestion site of TEV protease was present to allow removal of the corresponding stretch. As shown in Figure 10 when Apoptin was produced fused to TrxA and GST it remained insoluble, while both MBP-Apoptin and NusA-Apoptin were obtained in the soluble fraction. When I digested with TEV protease to release the MBP fused to Apoptin, this mostly precipitated, significantly reducing the amount of pure Apoptin that could be obtained by this strategy. Conversely, the digestion with the TEV protease to excise the NusA protein did not result in the precipitation of Apoptin. Both, TEV protease and NusA contain a

His-tag and can be separated from free Apoptin through a Ni-NTA affinity chromatography. Digested Apoptin eluted in the same fraction as NusA-Apoptin, however they ran separately in the SDS-PAGE (not shown). This indicates that even though the TEV digestion was conducted, the two proteins remained together. Likely, Apoptin and NusA could remain joined by disulfide bonds since NusA possesses three cysteines in its primary sequence and it has been seen that despite the fact that Apoptin has four cysteines, Cys90 is reactive to (5-(3-carboxy-4-nitrophenyl)disulfanyl-2-nitrobenzoic acid) (DNTB) (Leliveld et al., 2003c). Therefore, the disulfide bond formation between the two proteins could take place through the reaction of this free cysteine residue but it was cleaved after the addition of the loading buffer which contains β-mercaptoethanol.

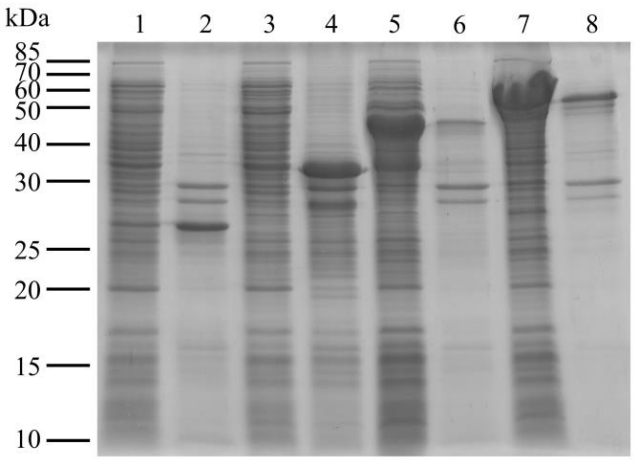


Figure 10. SDS-PAGE analysis of Apoptin fused to solubilizing partners. SDS-PAGE of soluble (odd lanes) and insoluble (even lanes) fractions of cells expressing Apoptin fused to TrxA (1, 2), GST (3, 4), MBP (5, 6) and NusA (7, 8) after lysis using a French Press.

Alternatively, I decided to purify the protein from the insoluble fraction. Since fewer proteins are found in the insoluble form, the desired protein represents a higher percentage of the sample. In order to facilitate its purification, the Apoptin gene was subcloned in pET-28a expression vector and ligated in frame with an N-terminal 6xHis-tag coding region (Figure

5). Although the 6xHis-tag can decrease the solubility of the fused protein, the purification from inclusion bodies counts with several advantages such as lower degradation of the expressed protein, resistance to proteolytic attack by cellular proteases or the presence of less contaminants among others (Singh and Panda, 2005). The inclusion bodies were solubilized using 7 M urea and the protein was purified through Ni-NTA affinity chromatography. Initially, the denatured and unfolded protein sample was dialyzed in order to gradually reduce the denaturant concentration, however at final steps the protein tended to precipitate (not shown). Different chemical chaperons such as glycerol were assayed to help the mutant protein keep it in solution. After multiple attempts, the fractions gathered during the Ni-NTA affinity chromatography were diluted dropwise in the refolding solution as detailed in the Material and Methods section. This implied the introduction of a concentration step to the purification protocol and yet a final gel filtration chromatography was necessary to remove some impurities, lastly resulting in a pure preparation according to MALDI-TOF and SDS-PAGE analysis (Figure 11). Since the purification protocol for Apoptin was finally set up using the pET-28a vector, the variant used in the following characterization was named H_6-Apoptin as it possesses a His-tag at its N-terminus. The theoretical molecular weight of H_6-Apoptin is 15,433 Da, however, the experimental molecular mass obtained by MALDI-TOF (Figure 11-B) of the purified variant was 15,305 ± 5 Da. This is in agreement with H_6-Apoptin lacking the initial N-formylmethionine, characteristic of the bacterial translation initiation. The short side-chain of Gly2 might have facilitated the N-terminal methionine removal catalyzed by *E. coli* methionine aminopeptidase (Hirel et al., 1989).

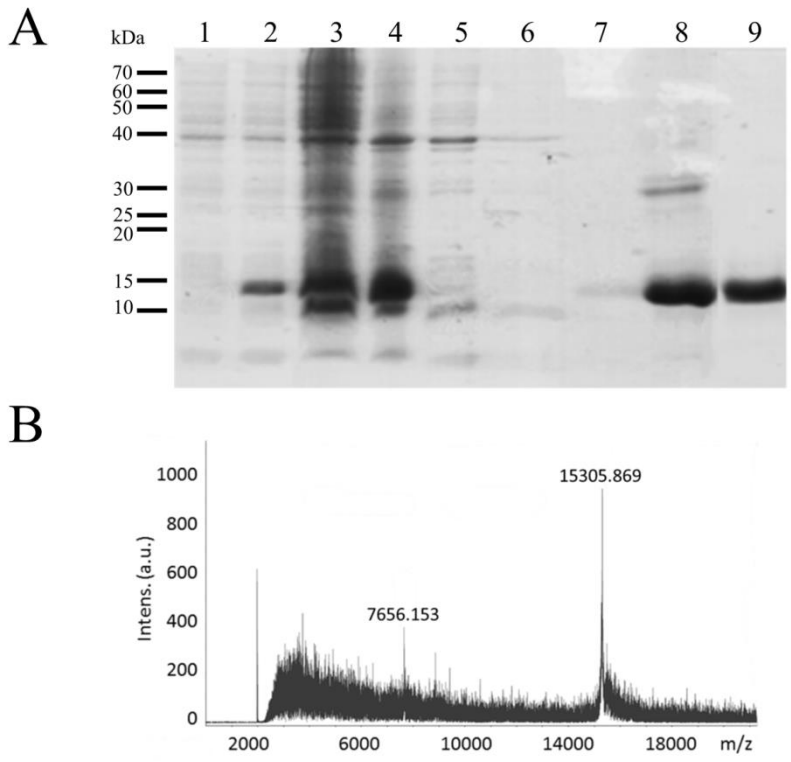


Figure 11. SDS-PAGE analysis of the production and purification steps of H_6-Apoptin. **A)** *E. coli* culture extracts before (1) and 3 h after the addition of IPTG (2); soluble (3) and insoluble (4) material from the French press; unbound (5), wash (6, 7) and imidazole-eluted fraction (8) from the Ni-NTA chromatography and H_6-Apoptin purified after SEC (9). **B)** MALDI-TOF mass spectrum of H_6-Apoptin. The molecular mass of H_6-Apoptin (15,305 Da) and the peak at half of its molecular weight (7,656 Da), corresponding to a double charged sample, are indicated.

Initial characterization of H_6-Apoptin

H_6-Apoptin forms water-soluble multimers considering that it elutes at the void volume when subjected to Superdex-75 SEC (Figure 12-A) at the last purification step. This was not surprising, since similar results have been previously reported for the wild-type protein (Noteborn et al., 1994) and for MBP-fused Apoptin (Zhang et al., 2003). I characterized the size and shape of H_6-Apoptin samples by TEM and DLS. TEM images showed that H_6-Apoptin aggregates are spherical-like structures with radii of 10-12.5 nm (Figure 12-B). Characterization using DLS confirmed that H_6-Apoptin exists

as a single solute species with an average R_H of 10.23 ± 0.16 nm (Figure 12-C). Based on this result, a value for the molecular mass of 770 kDa can be calculated which corresponds to approximately 50 subunits of H_6-Apoptin per particle. This result is very similar to the 30-40 monomers that form the multimeric globules of Apoptin fused to maltose-binding protein or fused to a C-terminal His-tag as described previously (Leliveld et al., 2003a).

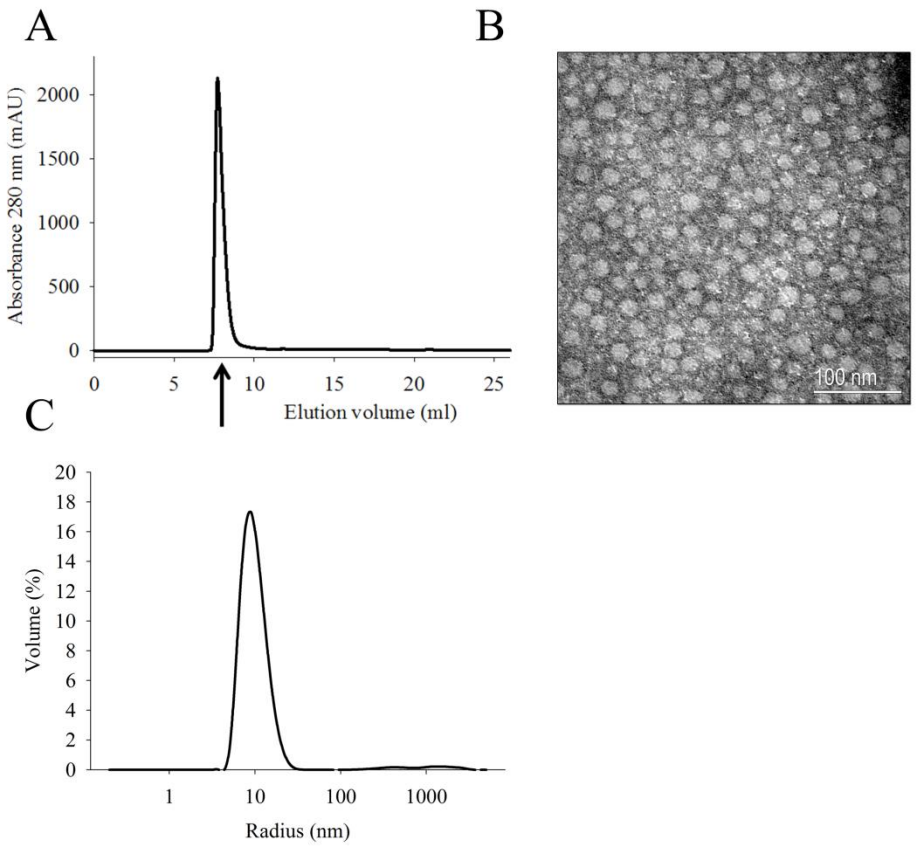


Figure 12. Purified H_6-Apoptin aggregates are water-soluble multimers. **A)** Analysis by SEC of purified H_6-Apoptin. H_6-Apoptin elutes at the void volume (arrow) of the Superdex-75 SEC column **B)** Analysis by TEM of Apoptin aggregates. Inside bar indicates the scale. **C)** Particle size distribution of H_6-Apopin according to the DLS analysis.

Apoptin and different proteins that cause proteinopathies share the feature that all of them have a high tendency to aggregate and that these

aggregates are toxic to the cells (Bucciantini et al., 2002). The mechanism of toxicity of most of these proteins has been related to the formation of amyloid aggregates (Shelkovnikova et al., 2012). However, it has been shown for the proteinopathic amyloid $\beta_{1\text{-}42}$ (Aβ42) peptide that it first aggregates to form spherical amyloid oligomers, which are later converted to protofibrils and fibrils *en route* to the deposition of amyloid plaques associated with Alzheimer's disease (Ahmed et al., 2010). These soluble proto-amyloid species are generally the most neurotoxic species forming along the amyloidogenesis pathway of several polypeptides implicated in human diseases (Ahmed et al., 2010). It is also relevant to note that a relation between protein amyloid-forming peptides and cytotoxicity has recently been rationalized (Torrent et al., 2011). Although most work on Apoptin has been focused on its behavior in the nucleus and its triggering of intracellular apoptosis programs, we thought it would be worth testing H_6-Apoptin aggregates for possible amyloid-like structure to determine whether an alternative mode of cytotoxicity is possible.

Different techniques have been broadly used over time as a diagnostic test for the presence of amyloid fibrils (for a review see (Xu et al., 2016)). I have analyzed H_6-Apoptin for amyloid-like structure formation by Congo Red (Khurana et al., 2001) and the benzothiazole dye Th T (Khurana et al., 2005).

Figure 13 shows the effects in the CR absorbance upon incubation with H_6-Apoptin aggregates. The experiment was performed at pH 7.4 and also at pH 9.2 at which the H_6-Apoptin aggregates are much less soluble (not shown). It is described that when CR binds to an amyloid structure, the spectra of bound and free CR have three characteristic features: i) there is an hyperchromicity and shift at the maximum absorbance from about 490 nm to 540 nm (Kim et al., 2003), ii) there are two isosbestic points between the spectra of bound and free CR (in which both spectra have equal molar

extinction coefficients) that fall at 406 and 477 nm (Klunk et al., 1989) and also, iii) the point of maximal spectral difference occurs at 540 nm. As it can be seen in Figure 13, when the assay was performed at pH of 9.2 both isosbestic points could be observed while at pH 7.4 there was only the one at 406 nm. In both cases it can also be seen that the maximal spectral difference takes place at 540 nm as previously described. Regardless of the buffer used, the characteristic hyperchromicity is also visible and the absorbance shift from ~490 nm to ~540 nm of the spectrum of the H_6-Apoptin bound to CR.

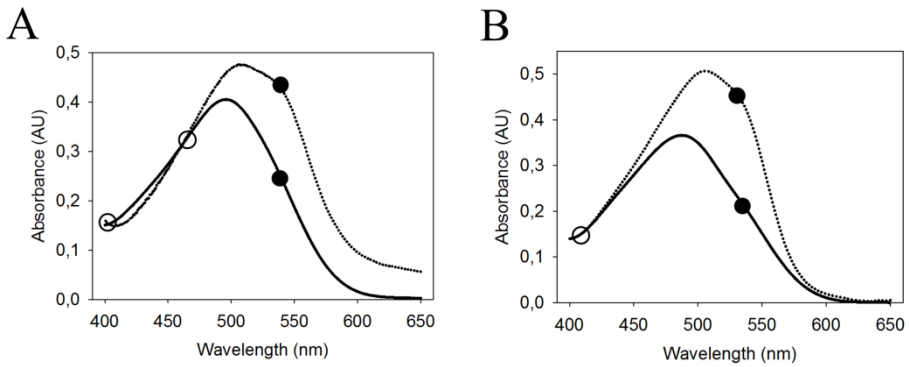


Figure 13. Congo Red assay of H_6-Apoptin. Absorbance spectra of CR in the absence (solid line) and in the presence (dotted line) of H_6-Apoptin in 50 mM Gly pH 9.2 buffer **(A)** or in TRIS base 50 mM pH 7.4 (adjusted with HCl), 100 mM NaCl buffer **(B)**. Open circles indicate isosbestic points; solid circles designate points of maximal spectral difference (540 nm).

It is known that Th T undergoes a strong increase in fluorescence emission intensity (around 480 nm, when excited at 450 nm) upon binding to amyloid fibrils (Naiki and Gejyo, 1999). When H_6-Apoptin was incubated with Th T and excited at 450 nm a 2.3-fold increase in the emission spectrum at approximately 470 nm was observed (Figure 14). It has been previously shown that, since Th T is a positively charged molecule, its binding to cationic proteins such as H_6-Apoptin can be very poor due to electrostatic repulsions. This effect can be minimized by performing the assay at a high concentration of NaCl (final concentration of 2.4 M), which

increases the fluorescence emission 10-fold (Sabaté et al., 2008). In these conditions, an increase of only 5.5-fold in the emission spectrum at ~470 nm was observed when the Th T-H_6-Apoptin was excited at the same wavelength. However, Th T binding is not restricted to complete amyloid fibrils as it has been reported to bind oligomers with typical features of molten globule state (Lindgren et al., 2005). These aggregates are spherical particles with a diameter of about 50 nm that resemble the H_6-Apoptin aggregates.

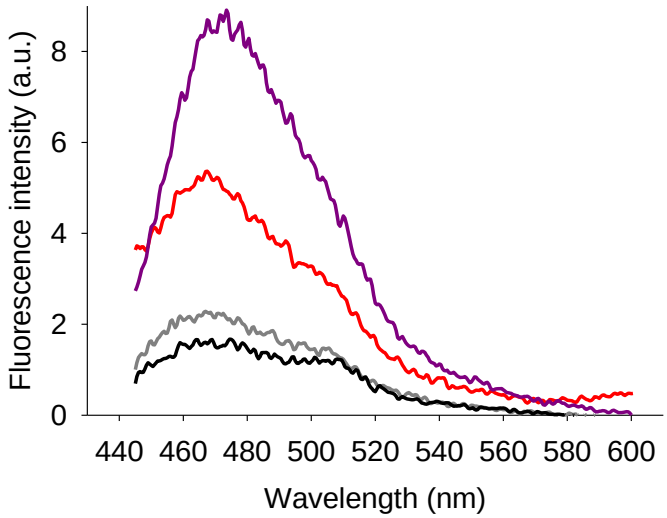


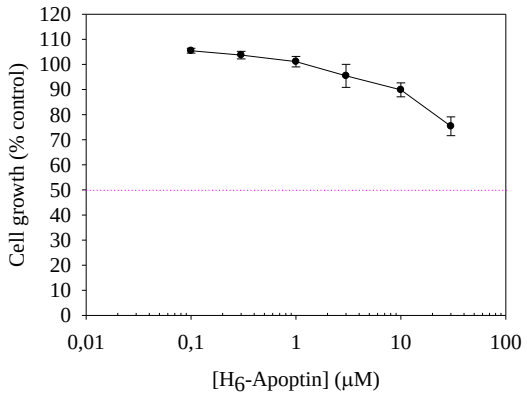
Figure 14. Fluorescence emission spectra of thioflavine T-H_6-Apoptin. Grey line: free Th T; black line: free Th T in 2.4 M NaCl; red line: H_6-Apoptin-Th T; purple line: H_6-Apoptin-Th T in 2.4 M NaCl.

Among all the techniques used to detect amyloid fibrils only the CR binding has proved positive. Conversely, since CR can bind native and partially folded conformations and amyloid fibrils of several proteins (Khurana et al., 2001), these results must be supported with other tests. Although neither Th T results nor TEM analysis indicate the presence of amyloid structures in the sample, the formation of amyloids cannot be completely discarded. Additionally, non-fibrillar but partially unfolded aggregates interact with CR and Th T in a similar amyloid manner (Komatsu et al., 2006). Thus, the results obtained for H_6-Apoptin must be

due to its probably low content in secondary structure rather than amyloid-like formation within its aggregates (see below). Alternatively, Apoptin may form or participate in liquid droplets such as those form by FUS, a prion-like protein with intrinsically disordered domains that eventually convert to an aggregated state (Patel et al., 2015). To date however, there are not evidences of this.

Apoptin is specifically cytotoxic for tumor cells. This cytotoxicity is produced in a wide range of tumor and transformed cells when the protein is microinjected or its gene is transfected (Danen-Van Oorschot et al., 1997) but not when it is exogenously added as a protein. However, when seeking for a cancer therapy, a protein-based system is preferable to a gene-delivery approach since this strategy may provoke insertional mutagenesis in transfected cells or may lack of efficacy due to transient expression. Moreover, fusion of Apoptin to different recombinant membrane-transferring penetrating peptides, like PTD4-Apoptin or TAT-Apoptin, significantly increases its cytotoxic efficiency (Guelen et al., 2004; Sun et al., 2009). As a first test, we were interested in investigating whether the H_6-Apoptin protein was cytotoxic for tumor cells. Therefore, I treated the tumor cell line NCI-H460 with different concentrations of the protein and analyzed their effect with the MTT assay. H_6-Apoptin is not significantly cytotoxic after 72 h of incubation with the cells (Figure 15-A). The observation of the cultures using an optical microscope showed that most of the protein had precipitated into the medium (Figure 15-B), and thus, it was unlikely that it would have reached the cytosol, even if it had been endocytosed by the cells. This would account for why this protein is not cytotoxic when it is added externally. The small decrease in cell growth at higher concentrations might be due to the huge amount of aggregates that hampered *per se* the normal cell division and adhesion in the flask.

A



B

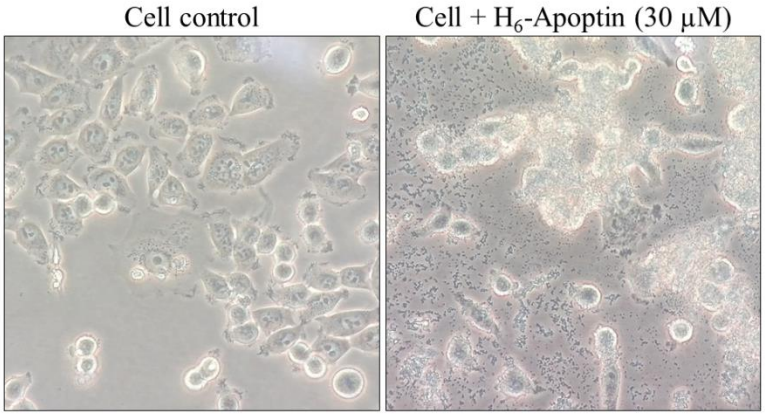


Figure 15. H_6-Apoptin is not cytotoxic and aggregates when it is added exogenously into the media of NCI-H460 cancer cells *in vitro*. **A)** Metabolic effect of different concentrations of H_6-Apoptin on NCI-H460 tumor cells. Metabolic activity was determined at 72 h post treatment by MTT assay as described in the Materials and Methods section. Cell growth is expressed as the percentage of control activity using the absorbance values. The curve in the figure is from one representative experiment using triplicates. Equivalent results were found in at least three independent experiments. **B)** Untreated cells (left) and cells treated with 30 μM of H_6-Apoptin (right) were incubated for 24 h, then examined and photographed (x400) under phase-contrast microscopy.

Production of a putative monomeric soluble form of Apoptin

We postulated that decreasing Apoptin's tendency to aggregate could be an alternative strategy to produce a non-oligomeric protein that could be cytotoxic when exogenously added. As mentioned in the Introduction section (*vide supra*), the sequence of Apoptin contains two stretches in the

N-terminal region, which are highly rich in aliphatic residues. The first is the proline rich segment (PRS) between residues 8 and 28 that hosts seven prolines and the second is the leucine/isoleucine rich segment (LRS) spanning residues 33 and 46 with seven leucines/isoleucines (Figure 2).

The LRS has been proposed to be involved in Apoptin's oligomerization (Leliveld et al., 2003a), and removal of the 65 N-terminal residues of Apoptin fused to MBP produces a soluble protein that seems to equilibrate between monomer, dimer and trimer. We aimed to produce truncated versions of Apoptin less prone to aggregation by selective removal of the PRS and LRS regions of Apoptin.

Therefore, I constructed two N-terminal truncated variants of Apoptin carrying an N-terminal His-tag (Figure 5). The first one lacks residues 1 to 27 comprising the PRS (H_6-ApopΔPro). The second variant, called H_6-ApopΔProΔLeu lacks both the PRS and the LRS (residues 1 to 43) (Figure 16).

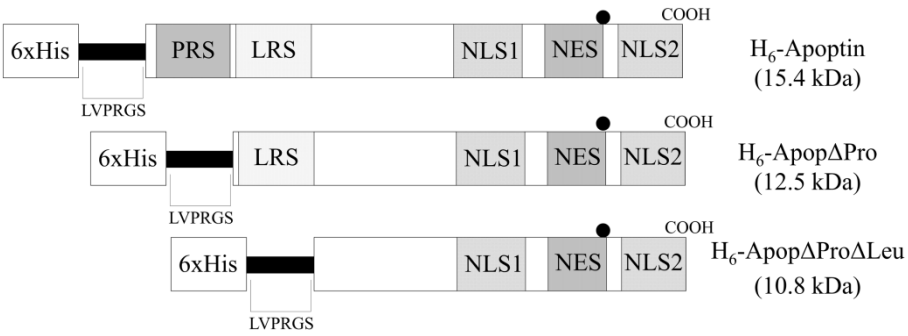


Figure 16. Schematic representation of H_6-Apoptin and its two variants H_6-ApopΔPro and H_6-ApopΔProΔLeu. There are indicated the regions involved in nuclear import/export: the putative NLSs (NLS1, residues 82-88; NLS2, residues 111-121) and NES (residues 97-105), the proline-rich sequence (residues 8-28), the leucine-rich sequence (residues 33-46) and the phosphorylation site (Thr 108). All the proteins possess an N-terminal His-tag and the thrombin cleavage site (Leu-Val-Pro-Arg-Gly-Ser).

Both Apoptin truncated variants were produced and purified using a protocol similar to that of H_6-Apoptin that required SEC as a final purification step. The identity of each variant was checked by MALDI-TOF (Figure 17). The observed masses were 12,412 ± 5 Da and 10,678 ± 5 Da, respectively, whereas the expected masses were 12,559 Da and 10,811 Da indicating that in both variants, as had occurred in H_6-Apoptin, the N-formylmethionine had been removed. The final yield for H_6-ApopΔPro was of 20 mg per liter of induced culture while for H_6-ApopΔProΔLeu the yield was of 10 mg per liter of induced culture.

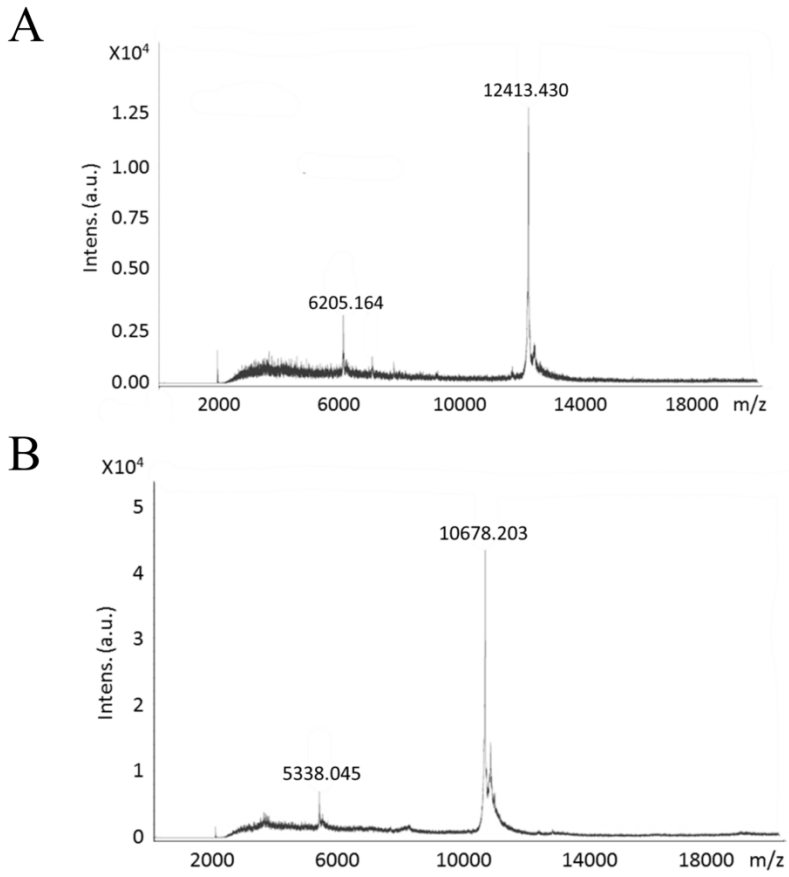


Figure 17. MALDI-TOF mass spectra of purified H_6-ApopΔPro **(A)** and H_6-ApopΔProΔLeu **(B).** The observed masses of each protein and of the doubly charged species at half the m/z value are indicated.

Resembling His-tagged wild-type Apoptin, H_6-ApopΔPro forms large water-soluble multimers as it elutes at the void volume of the Sephadex-75 SEC (Figure 18-A). The average R_H of the aggregates, measured by DLS, is of 9.34 ± 0.35 nm (Figure 18-B). The estimated molecular mass obtained based on the radius is 630 kDa, which corresponds to approximately 50 subunits of H_6-ApopΔPro per particle. Interestingly, this is the same number of subunits found for its parental protein. In contrast, H_6-ApopΔProΔLeu was mainly obtained as a protein that had an elution volume of 12 ml (87%) although a minor fraction (13%) of the purified protein had an elution volume of 10 ml (figure 18-A). The fraction of higher molecular mass was termed HMM fraction and that of lower molecular mass, eluting at 12 ml, was named LMM fraction. After calibration of the SEC column with different globular proteins, these elution volumes would correspond to molecular masses of 29 kDa for the LMM fraction and to 52 kDa for the HMM one, assuming that they are globular proteins. We analyzed both fractions of H_6-ApopΔProΔLeu by DLS (Figure 18-B). The average R_H of the two forms were 2.74 ± 0.11 nm for LMM and 3.46 ± 0.27 nm for HMM; these values corresponded, again if they were globular proteins, to estimated molecular masses of 35.9 kDa and 62.7 kDa, respectively.

The calculations of the molecular mass and the size of the oligomeric species (monomer, dimer, etc.) by gel filtration and light scattering both implicitly assume that the H_6-ApopΔProΔLeu behaves like a compact, globular protein. Additional experiments described below indicate that H_6-ApopΔProΔLeu is in fact an intrinsically disordered protein (IDP). Since IDPs are less compact than well folded, globular proteins, the calculations based on the gel filtration and light scattering results likely overestimate the size of H_6-ApopΔProΔLeu species and should be considered as upper limits.

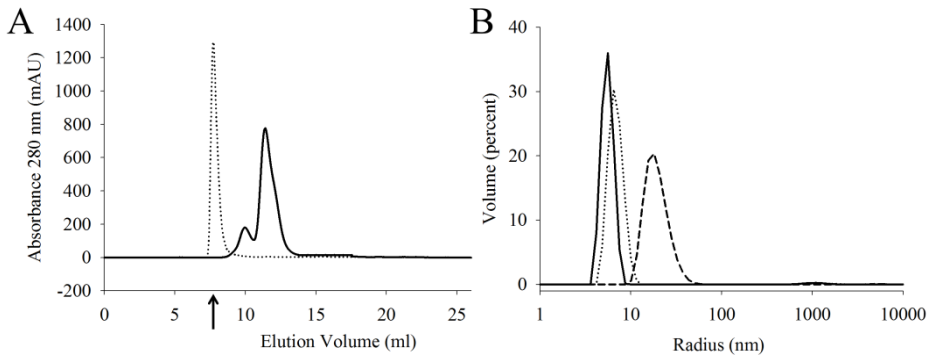


Figure 18. Purified H_6-ApopΔPro aggregates as a water-soluble multimer whereas H_6-ApopΔProΔLeu consists of a mixture of monomer and a dimer/low oligomeric forms. **A)** Analysis by SEC of purified H_6-ApopΔPro (dotted line) and H_6-ApopΔProΔLeu (continuous line). H_6-ApopΔPro elutes at the void volume of the column (arrow) whereas H_6-ApopΔProΔLeu elutes with forms compatible with 29 and 52 kDa. To estimate these molecular masses, the column was calibrated with proteins of known molecular weight (see Material and Methods). **B)** Particle size distribution (radius) of H_6-ApopΔPro (dashed line) and the two forms of H_6-ApopΔProΔLeu (major fraction (solid line) and minor fraction (dotted line)) measured by DLS. The average R_H are 9.34 ± 0.35 nm for H_6-ApopΔPro, 2.74 ± 0.11 and 3.46 ± 0.27 nm for the two forms of H_6-ApopΔProΔLeu. The estimated molecular weights are 630 kDa, 35.9 and 62. 7k Da, respectively.

I analyzed H_6-ApopΔPro and the two fractions of the H_6-ApopΔProΔLeu by SDS-PAGE in the presence of β-mercaptoethanol and by non-denaturing electrophoresis (Figure 19). Under denaturing conditions H_6-ApopΔPro presents three bands, two of them were within the resolving gel and sized 11 and 26 kDa, and the third one remained in the stacking gel. These three bands could correspond to monomeric, dimeric and oligomeric forms. In non-denaturing conditions, there is a single band, unable to enter the gel. Regarding H_6-ApopΔProΔLeu, in which both PRS and LRS were eliminated, both fractions were able to enter the native gel, which corroborates the idea that they do not form higher oligomers. The LMM fraction presented a unique band of 10-11 kDa in the SDS-PAGE, whereas the HMM fraction presented two bands of 10-11 kDa and 21 kDa. These molecular weights correspond to those expected for a monomeric and dimeric species.

In summary, these results indicate that removal of the PRS region does not reduce the tendency of Apoptin to aggregate whereas the combined elimination of both the PRS and the LRS produces a monomeric/small oligomeric soluble form of Apoptin.

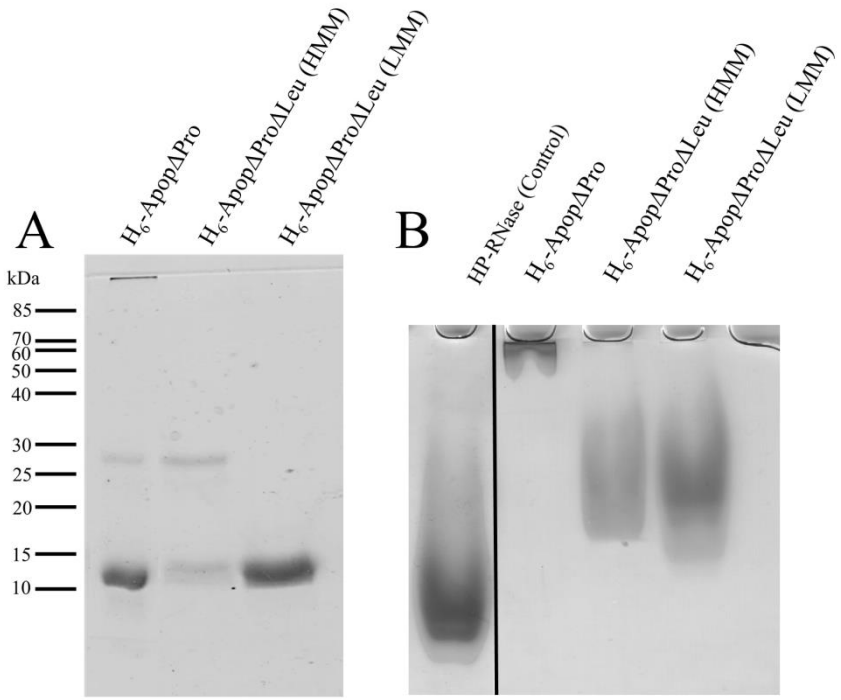


Figure 19. A) Analysis by SDS-PAGE and **B)** non-denaturing electrophoresis of the purified H_6-ApopΔPro and the HMM and LMM forms of H_6-ApopΔProΔLeu eluted from the SEC.

Biophysical characterization of the LMM and HMM fractions of H_6-ApopΔProΔLeu

In order to characterize both fractions of the H_6-ApopΔProΔLeu variant, they were reloaded onto the size exclusion chromatography once they had been separated. Most of the LMM fraction eluted at the same elution volume when it was reloaded to the Sephadex-75 column under the same conditions even after 96 h at 4°C (Figure 20).

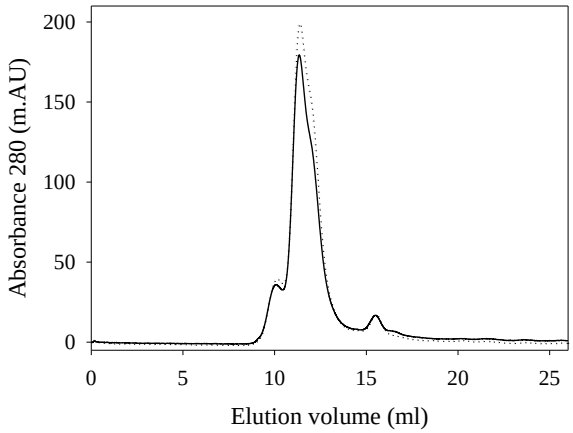


Figure 20. Analysis of the stability of the LMM form of H_6-ApopΔProΔLeu by SEC. After the purification, the protein was dialyzed against ddH_2O, lyophilized and stored at -20°C. The lyophilized protein was resuspended in a buffer containing 50 mM TRIS base/HCl, 150 mM NaCl pH 7.4 and was immediately loaded to the column (dotted line) or after 96 h incubation at 4 °C (solid line).

Regarding the HMM fraction, after incubating the sample for 2 h at 37°C, most of the molecules eluted at 10 ml in the SEC (Figure 21-A) although a minor fraction eluted as the LMM form. These results could indicate that either the equilibrium between the HMM and LMM fractions was taking place very slowly or that it was hampered by disulfide bond stabilization.

H_6-ApopΔProΔLeu maintains three of the four cysteines present in the wild type protein. We investigated if the subunits of the HMM form of H_6-ApopΔProΔLeu were stabilized through disulfide bonds by incubating it for 2 h at 37°C in the presence of 1 mM TCEP. TCEP is an irreversible reducing agent of disulfide bonds that, unlike DTT and other thiol-containing reducing agents, does not have to be removed from the protein. The treated HMM form was then loaded to the SEC (Figure 21-A) and 58.6% of the molecules eluted at 12 mL which corresponds to the elution volume of the LMM form. As an additional test, HMM was treated with and without β-mercaptoethanol (0.1%) and loaded into an SDS-PAGE (Figure 21-B). In the presence of the reducing agent we can see that the

most intense band has a molecular weight compatible with a monomeric H_6-ApopΔProΔLeu, while in the absence of β-mercaptoethanol, the most intense band is that of a dimer. These results would indicate that the HMM form of H_6-ApopΔProΔLeu might be stabilized in part by disulfide bonds, but also by other non-covalent interactions like hydrophobic interactions as it has been described in the formation of Apoptin multimers (Leliveld et al., 2003a). Also, the comparison of the results from SEC and SDS-PAGE in the presence and absence of a reducing agents seems to indicate that the LMM form may correspond to a monomer whereas the HMM form may correspond to a dimer.

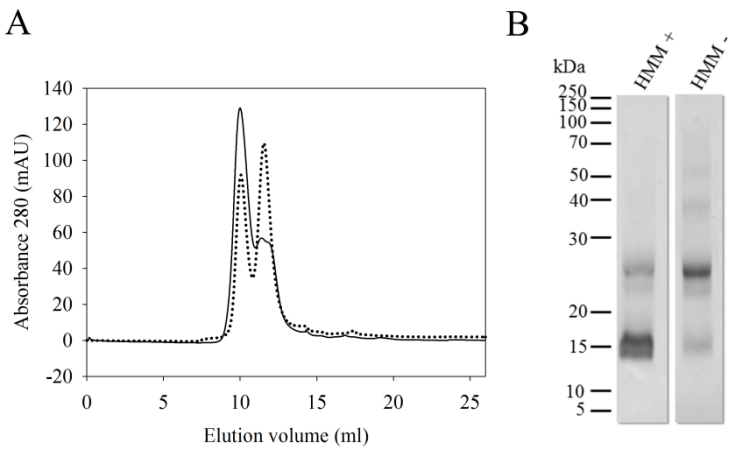


Figure 21. Analysis of the stabilization of the HMM form of H_6-ApopΔProΔLeu. **A)** The HMM form was subjected to SEC after incubation for 2 h at 37ºC in the presence (dotted line) or absence (solid line) of 1mM TCEP. **B)** SDS-PAGE analysis of the HMM form in the presence (HMM+) and absence (HMM-) of 0.1% β-mercaptoethanol.

In an attempt to better understand the quaternary structure of HMM and LMM I investigated the thermal stability of the two forms of H_6-ApopΔProΔLeu by following the changes in absorbance at 278 nm upon heating, as described in the Material and Methods section.

The thermal denaturation for the LMM form of H_6-ApopΔProΔLeu was performed at different pH values spanning from 3.2 to 9.2; however no

apparent sigmoidal unfolding transition was detected in any of these thermal denaturation experiments (Figure 22). The observed lack of transition for these experiments could be due to the fact that H_6-ApopΔProΔLeu has only two residues that absorb at UV light (Phe46 and Tyr71 in this variant; Phe70 and Tyr95 in wild type Apoptin), and that these residues could be mainly exposed on the surface. Otherwise, Apoptin could be a semi-unfolded or dynamic protein that would lack a non-polar inner core. The different behavior at pH 9.2 (Figure 22) is almost surely due to contributions of the phenolate anion as the Tyr side chain becomes negatively charged (its pKa is around 10).

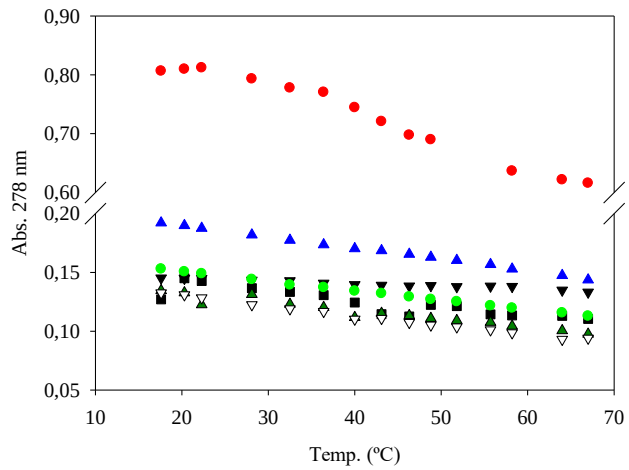


Figure 22. Changes in absorbance at 278 nm upon temperature increase of H_6-ApopΔProΔLeu at different pH: 50 mM glycine pH 3.2 (■); 50 mM sodium bicarbonate pH 4.5 (▲) 50 mM acetate pH 5.0 (▽); 50 mM MES pH 6.0 (▼); 50 mM MOPS pH 7.0 (●); 50 mM TRIS base/HCl pH 8.0 (▲) and 50 mM glycine pH 9.2 (●).

The thermal denaturation of the HMM form was also investigated in a buffer containing 50 mM TRIS base, 100 mM NaCl, 1 mM TCEP pH 7.0 (adjusted with HCl) and no unfolding transition was observed in this case (Figure 23).

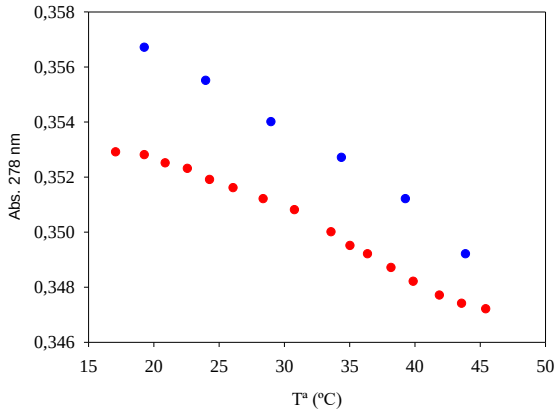


Figure 23. Changes in absorbance at 278 nm vs temperature increases (red dots) and decreases (blue dots) was followed spectrophotometrically for the HMM form of H_6-ApopΔProΔLeu.

Further investigations to evaluate the biological properties of H_6-ApopΔProΔLeu were centered on the LMM fraction i.e., that of lower molecular weight, because of its lower propensity to form high molecular mass oligomeric structures.

Characterization of the biological properties of H_6-ApopΔPro and H_6-ApopΔProΔLeu

I investigated whether Apoptin truncated variants maintained some of the biological properties of Apoptin namely, their ability to bind dsDNA, whether they were driven to the nucleus of transfected tumor cells and, in the case of the soluble non-aggregating variant, also its cytotoxic properties on tumor and normal cells.

dsDNA binding

I decided to test binding to DNA since it has been postulated that this ability may be linked to its apoptotic activity (Leliveld et al., 2004). A qualitative assay was performed in which each variant was incubated with a 1,135 base-pairs dsDNA. H_6-Apoptin and H_6-ApopΔPro associated to DNA

creating aggregates that were unable to enter the agarose gel (Figure 24). DNA incubated with H_6-ApopΔProΔLeu was observed as a smear band in the gel (the band was also weaker compared to the control), which could be indicative that this variant also associated with DNA, although to a lesser extent. The lower molecular mass of the complexes formed by DNA and H_6-ApopΔProΔLeu does not necessarily reflect a higher K_D. It has to be taken into account that the high molecular weight of Apoptin multimers may contribute significantly to delay DNA mobility upon binding and, also, that binding to DNA may be increased by an avidity (not affinity) effect. In fact, it is known that Apoptin multimers have about eight DNA binding sites of unknown number of base pairs (Leliveld et al., 2003b).

In addition, I also tested whether or not the variants' solubilities were affected by incubation with dsDNA. I incubated the Apoptin variants with this stretch of DNA in the same conditions used for the agarose-gel electrophoresis analysis and monitored protein aggregation by following the increase in turbidity (apparent absorbance at 600 nm) over 12 h of incubation. In these conditions H_6-Apoptin and H_6-ApopΔPro aggregates rapidly precipitated whereas H_6-ApopΔProΔLeu remained soluble (not shown).

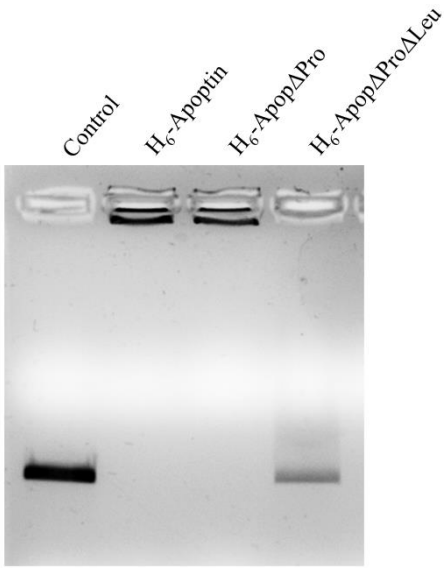
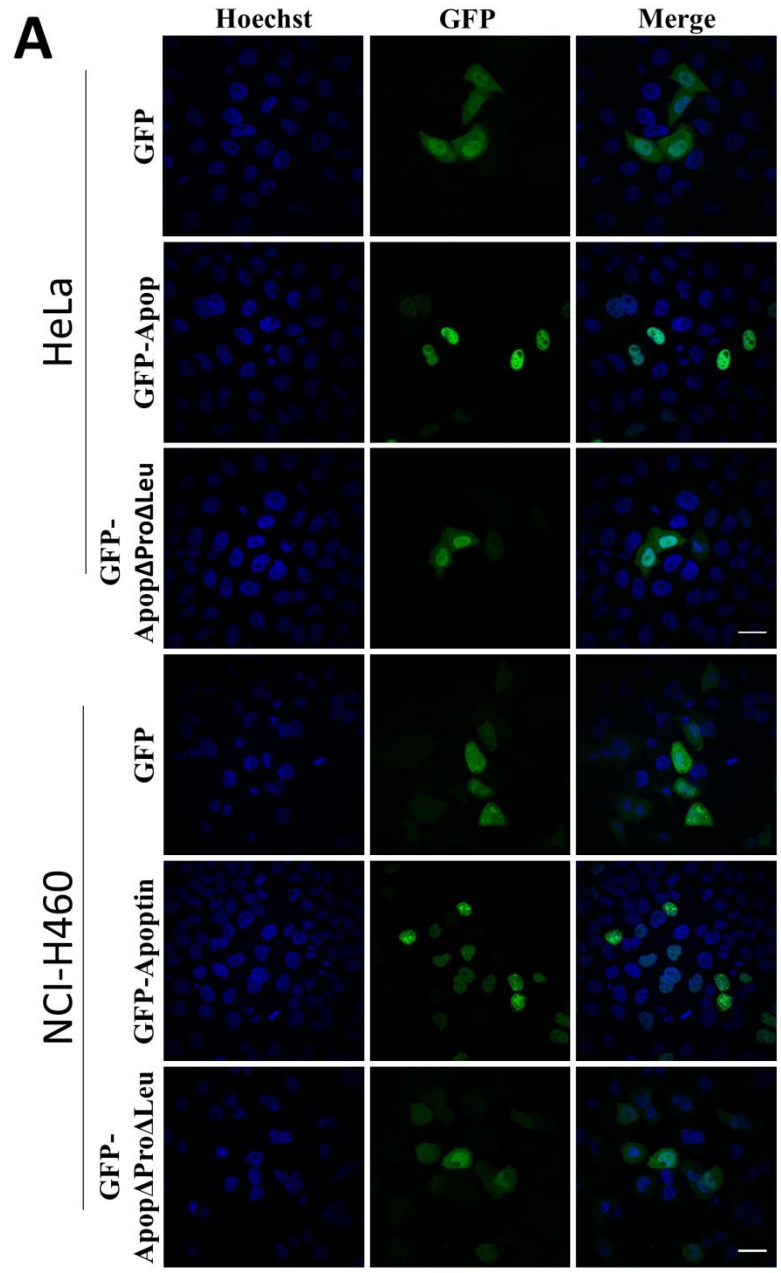


Figure 24. N-terminal truncated variants of Apoptin bind to DNA. Agarose gel electrophoresis of H_6-Apoptin, H_6-ApopΔPro and H_6-ApopΔProΔLeu incubated with a segment of 1,135 base pairs of dsDNA.

Subcellular localization

To test protein subcellular localization of Apoptin variants, their genes were fused to GFP (Figure 6), cloned under the control of a CMV promoter and transfected into HeLa and NCI-H460 tumor cells and into CCD-18Co normal cells, as described in Material and Methods. Confocal microscopy was used to investigate the subcellular localization of the GFP-chimeras in each case. As previously described, GFP-Apoptin localizes in the nucleus of HeLa cancer cells (Cheng et al., 2003) and also of NCI-H460 cells (Figure 25) but appears equally distributed between the nucleus and the cytosol of CCD-18Co normal cells. In contrast, GFP-ApopΔProΔLeu has a higher tendency to locate in the cytoplasm, especially in normal cells. Significant differences ($p<0.05$) for the values between GFP (control) and GFP-ApopΔProΔLeu within the three cell lines tested were obtained. This indicates that although ApopΔProΔLeu's overall tendency to accumulate into the nucleus is weakened, it retains part of the remarkable ability of

wild type Apoptin to selectively accumulate into the nuclei of cancer cells relative to normal cells. Besides, because H_6-ApopΔProΔLeu has a reduced tendency to accumulate in the nuclei of normal cells it could be assumed that this variant is less toxic than Apoptin to normal cells. Therefore higher doses of H_6-ApopΔProΔLeu could be used in cancer therapy increasing its effectiveness.


A
Hoechst
GFP
Merge
HeLa
GFP
GFP-Apop
GFP-ApopΔProΔLeu
NCI-H460
GFP
GFP-Apoptin
GFP-ApopΔProΔLeu

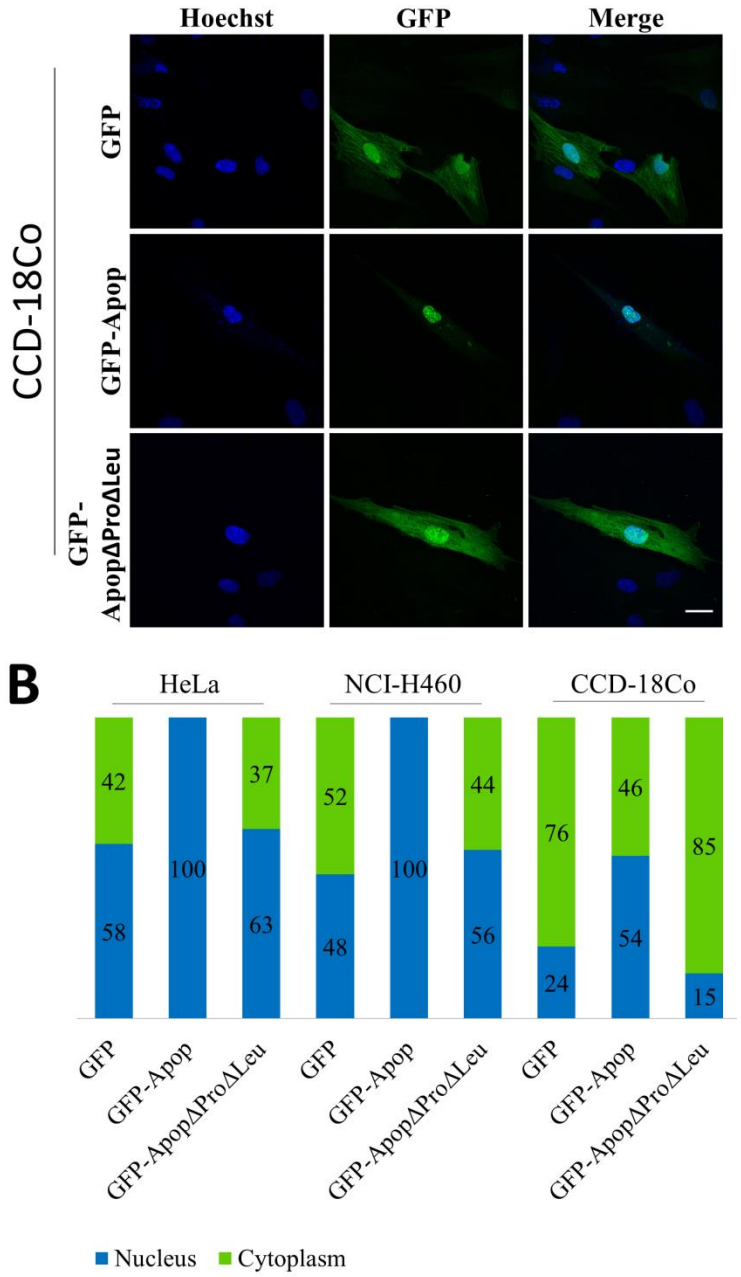


Figure 25. Subcellular localization of transfected GFP, GFP-Apoptin and GFP-ApopΔProΔLeu in HeLa, NCI-H460 and CCD-18Co cells. **A)** Confocal microscopy images of cells transfected with GFP-Apoptin or GFP-ApopΔProΔLeu before nuclear labelling with Hoechst 33258. Scale bar: 30 μm. **B)** Percentage of GFP-protein located in the nucleus is also indicated in each case.

I also investigated the nuclear localization of GFP-ApopΔPro in HeLa cells. The subcellular distribution was equal to that of GFP-Apoptin but, interestingly, GFP-ApopΔPro accumulates in structures that resemble the nucleolus of the cells (Figure 26). It was previously reported that GFP-Apoptin is present at high levels in the nucleoli of some cells (Guelen et al., 2004) and that, since Apoptin forms complexes with proteins involved in ribosome biogenesis and RNA metabolism (Leliveld et al., 2003b), it would not be surprising that this protein coordinates its apoptosis-inducing activity from inside the nucleolus of some cancer cells (Leliveld et al., 2003a).

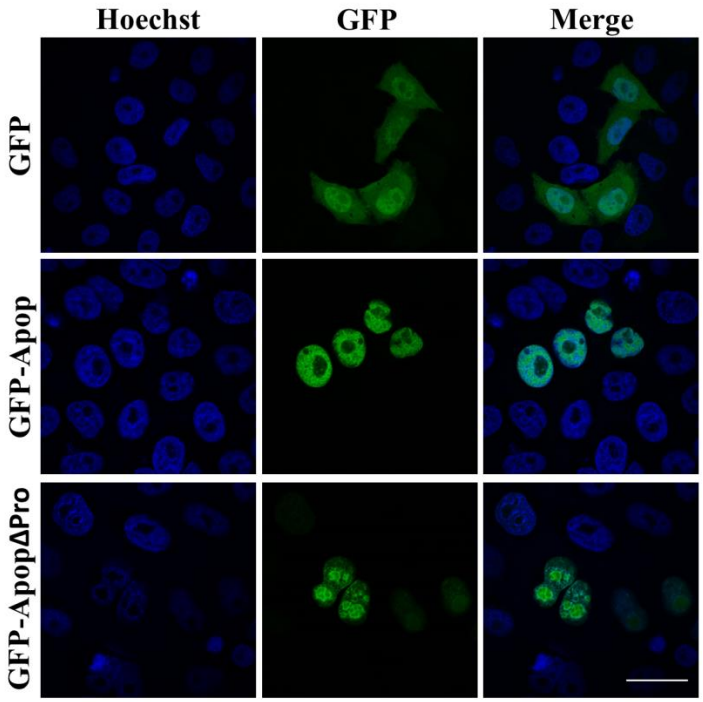


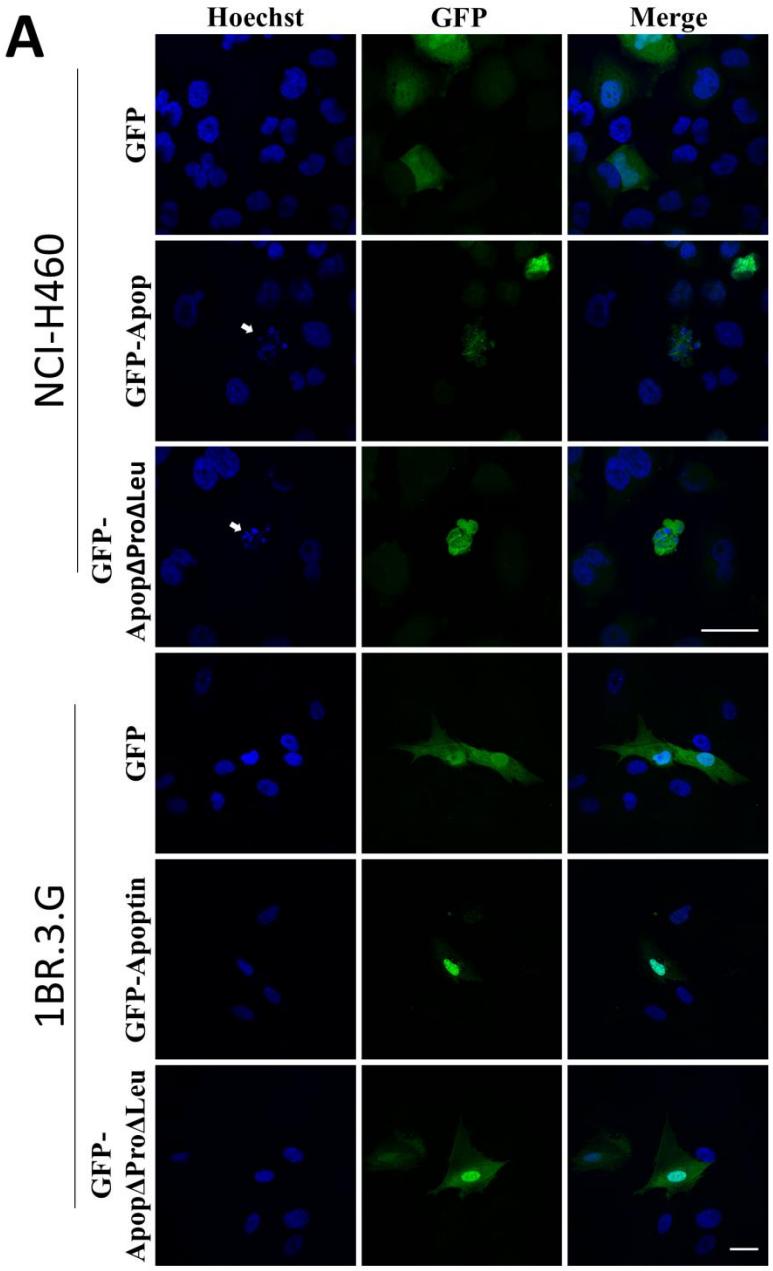
Figure 26. Nuclear localization of transfected GFP-Apoptin and its truncated variant GFP-ApopΔPro in HeLa tumor cells. Scale bar: 30 µm.

Cytotoxic properties

Next, I investigated whether ApopΔProΔLeu maintains the cytotoxic properties of wild type Apoptin by transfecting both GFP-fused genes into NCI-H460 tumor cells and CCD-18Co normal cells (Figure 27). Transfected

cells were selected by means of GFP fluorescence emission using flow cytometry and dead cells were analyzed by PI staining. The low levels of transfection of CCD-18Co precluded the cytotoxic analysis of Apoptin variant on these cells via flow cytometry. Both GFP-fused proteins killed NCI-H460 tumor cells, with the effect of the GFP-Apoptin being 1.9 times higher than that of GFP-ApopΔProΔLeu (Figure 27-B). Together, these results indicate that the truncated form of Apoptin retained the biological properties of the parental protein when transfected even though it was somewhat less active. This result might be due to the removal of the weak inducing domain at the N-terminus.

GFP-ApopΔProΔLeu and GFP-Apoptin were transfected on NCI-H460 tumor cells and CCD-18Co normal cells and they were examined using the confocal microscopy for nuclear changes associated with apoptosis after staining DNA with Hoechst 33258. In NCI-H460 transfected cells, both treatments induced nuclear morphological changes typical of apoptosis, such as the characteristic chromatin condensation and the nuclear fragmentation (Figure 27-A). Additionally, the presence of GFP-ApopΔProΔLeu in the cytoplasm of NCI-H460 cells allowed us to observe the membrane blebbing typical of apoptotic cells. In contrast, in CCD-18Co transfected cells these changes related to apoptosis were not observed.


A
Hoechst
GFP
Merge
NCI-H460
GFP
GFP-Apop
GFP-ApopΔProΔLeu
1BR.3.G
GFP
GFP-Apoptin
GFP-ApopΔProΔLeu

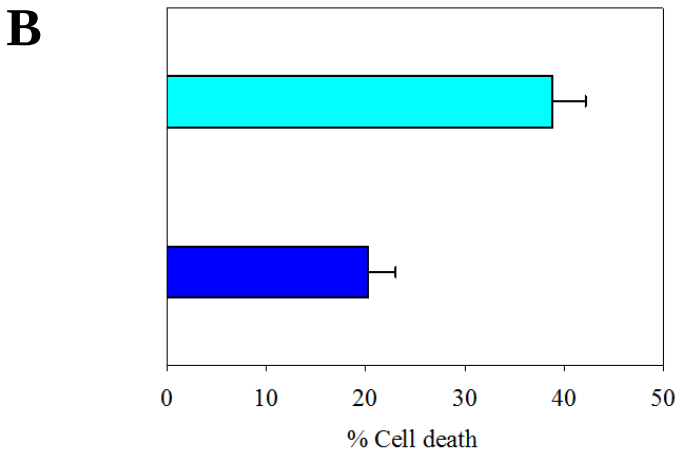


Figure 27. GFP-ApopΔProΔLeu is cytotoxic for cancer cells. **A)** H_6-ApopΔProΔLeu induces specifically signs of apoptosis in tumor cells. Hoechst 33258 stain of NCI-H460 and CCD-18Co nuclei of cells transfected with GFP-ApopΔProΔLeu, GFP-Apoptin or GFP. Hoechst 33258 staining reveals signs of induction of apoptosis in tumor cells (arrows). Scale bar: 30 μm. **B)** Cytotoxic activity of transfected GFP-Apoptin (cyan bar) and GFP-ApopΔProΔLeu (blue bar) into NCI-H460 cells. Transfected cells were incubated for 48 h and then attached and floating cells were harvested by centrifugation and subjected to PI staining. Stained cells expressing GFP or GFP-fused proteins were analyzed on FACSCalibur flow cytometer to quantify the percentage of dead transfected cells.

Remarkably, in some NCI-H460 and 1BR.3.G cells transfected with GFP-Apoptin, the protein was also located surrounding the nucleus in structures that resembled cytoplasmic filaments (Figure 28). Interestingly, Apoptin has been shown to adopt similar distribution after 24 h post transfection both in cancer cells and in normal cells (Guelen et al., 2004) and this pattern has been also observed for MBP-Apoptin (Leliveld et al., 2003a). Although these cytoplasmic filaments resembled the actin stress fibers at early stages of expression, structures that are important for maintenance of nuclear shape and integrity and cell viability (Tojkander et al., 2012), no colocalization of Apoptin with these fibers was found when MDCK dog epithelial cells expressing Apoptin-filaments were counter-stained with Texas-Red Phalloidin (Guelen et al., 2004). However, Apoptin colocalized with BCL-10 and FADD when coexpressed with either GFP-Bcl10 or FADD-GFP (Guelen et al., 2004). In our case, this distribution could

not be observed with GFP-ApopΔProΔLeu in any of the transfected cell lines. Similar results were obtained after transfecting Saos-2 and HeLa cells with MBP-Apoptin and the non-aggregating variant MBP-Apoptin(80-121), in which the aggregating protein but not the truncated variant was found surrounding the nucleus of the cells before being internalized (Leliveld et al., 2003a). It was suggested that the size of the aggregates delayed the diffusion of the fusion protein through the nuclear pores. In this regard, the small size of the non-aggregating variant H_6-ApopΔProΔLeu would allow a faster nuclear diffusion.

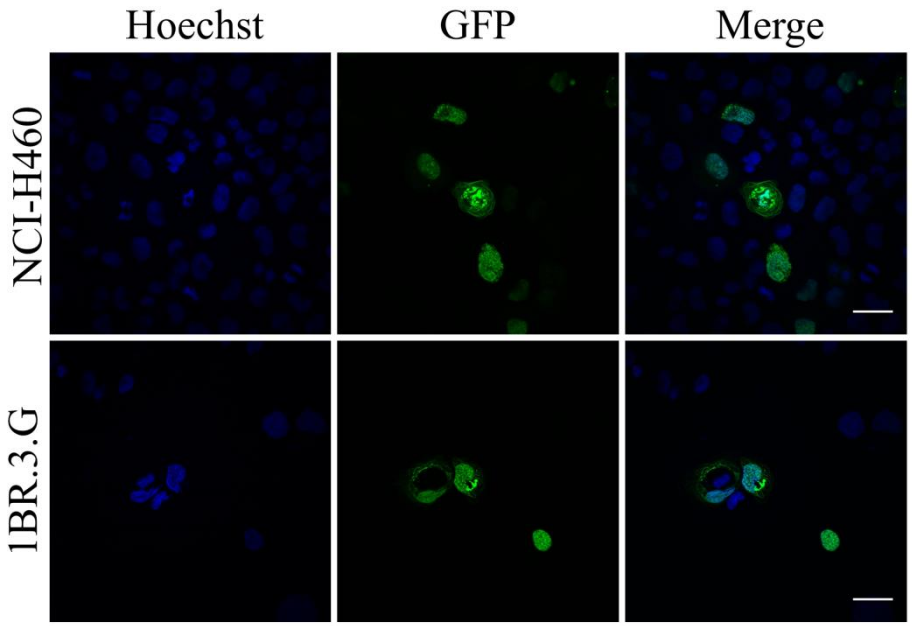


Figure 28. GFP-Apoptin localizes in cytoplasmic filaments surrounding the nucleus in NCI-H460 and 1BR.3.G cells transfected with GFP-Apoptin. Nuclei have been stained with Hoechst 33258. Scale bar: 30 μm.

H_6-ApopΔProΔLeu is selectively cytotoxic to cancer cells when added exogenously

The aim of generating the above described truncated variants was to reduce the predisposition of Apoptin to form multimers, a phenomenon that hampers the efficient internalization of Apoptin when externally added, therefore precluding its cytotoxic activity (Figure 29). Consequently, we investigated the cytotoxic properties of H_6-ApopΔProΔLeu on tumor and non-tumor cells when added exogenously. Using the MTT cell viability assay, H_6-ApopΔProΔLeu is cytotoxic for NCI-H460 (IC_{50} of 7.5 μM ± 1.3), K-562 (13 μM ± 0.6), OVCAR-8 (25 μM ± 0.01) and Jurkat tumor cell lines (14 μM ± 0.36) (Figure 29). In contrast, the truncated variant does not produce any effect on 1BR.3.G and CCD-18Co normal cells (Figure 29) indicating that H_6-ApopΔProΔLeu is selectively cytotoxic for cancer cells when it is added exogenously.

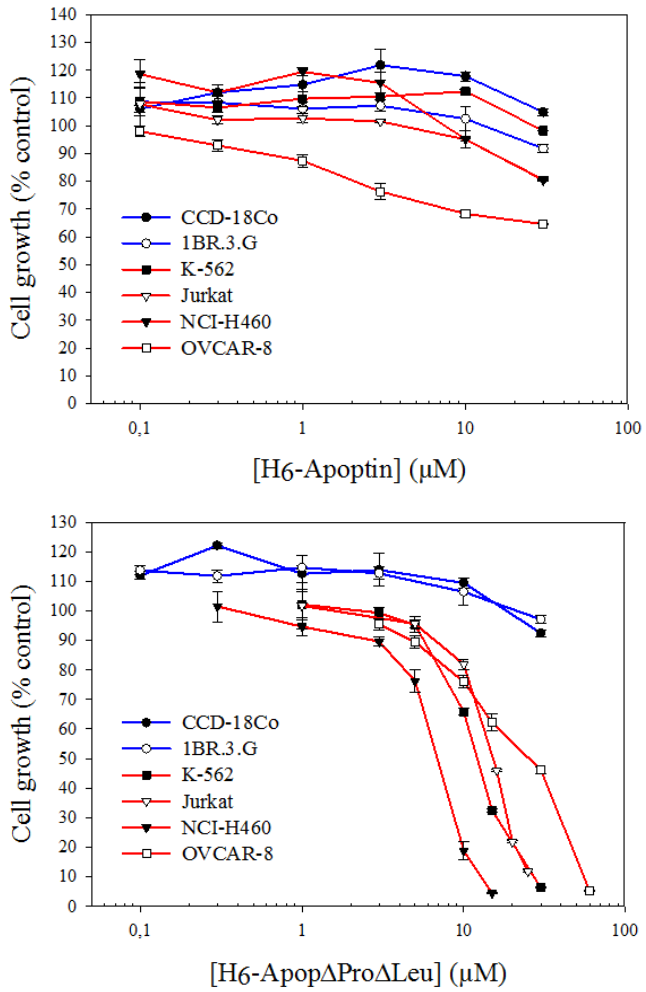


Figure 29. H_6-ApopΔProΔLeu is cytotoxic for cancer cells when added exogenously. Metabolic effect of different concentrations of H_6-Apoptin (upper figure) and H_6-ApopΔProΔLeu (lower figure) on tumor (red lines) and normal cell lines (blue lines). OVCAR-8 (□), K-562 (■), Jurkat (▽) and NCI-H460 (▼) tumor cell lines, and CCD-18Co (●) and 1BR.3.G (○) normal cells were treated with different concentrations of H_6-ApopΔProΔLeu for 72 h. Metabolic activity was determined by MTT assay as described in the Materials and Methods section. Cell growth is expressed as the percentage of control activity using the absorbance values. Curves in the figure are from one representative experiment. Equivalent results were found in at least three independent experiments.

I also investigated whether cell death induced by the treatment with H_6-ApopΔProΔLeu displayed specific features of apoptosis on NCI-H460 cells. We sought to quantify the percentage of NCI-H460 cells in early and late

apoptosis after 24, 36 and 48 h of incubation with H_6-ApopΔProΔLeu. FACS analysis of these cells, stained with Annexin V-Alexa Fluor 488 and PI (Table 4), demonstrated induction of apoptosis by the translocation of phosphatidylserine to the cell external hemi-membrane. Apoptosis was evident at 36 h of treatment and further increased at 48 h. Induction of apoptosis was confirmed by evaluating the activation of procaspases-3, -8 and -9. H_6-ApopΔProΔLeu induces the activation of these procaspases (Figure 30) reaching its maximum at 36 h. In the presence of the caspase-specific inhibitors (z-DVED-FMK, z-IETD-FMK or z-LEHD-FMK), the activity dropped to the level of the control extracts.

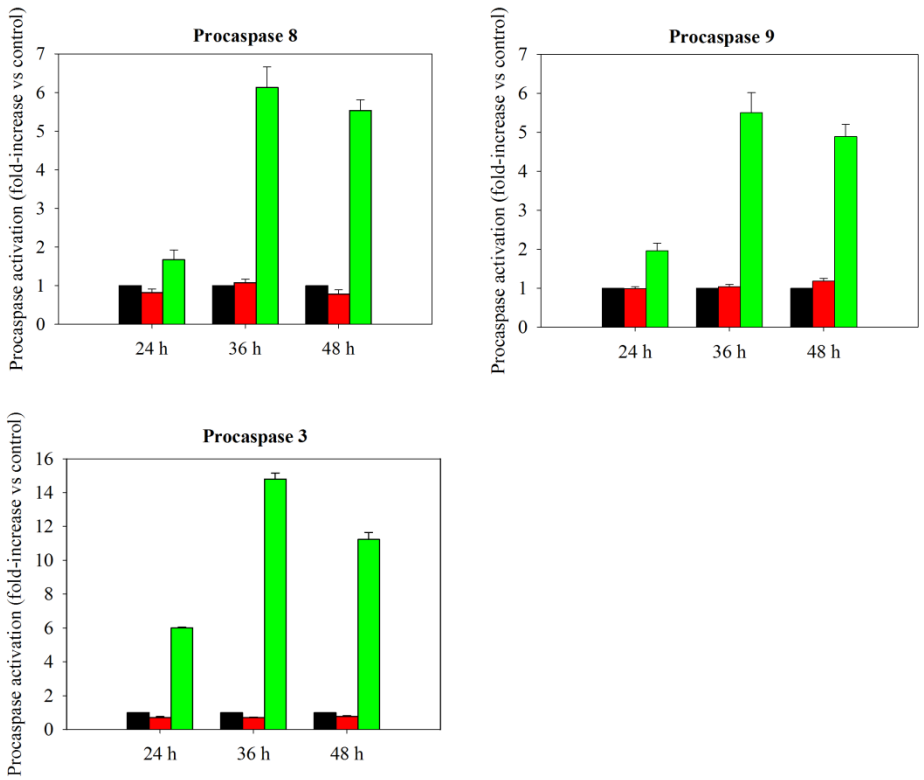


Figure 30. Procaspase-3, -8, and -9 activation in NCI-H460 cells treated with 15 μM of H_6-ApopΔProΔLeu for 24, 36 and 48 h (green). Black bars indicate untreated cells (control). Parallel samples were treated with the specific inhibitors of caspase-3 (z-DEV-FMK), caspase-8 (z-IETD-FMK), and caspase-9 (z-LEHD-FMK) (red) as a negative control. Results are expressed as the mean ± SE of three independent experiments.

Table 4. Apoptosis measured by Alexa Fluor 488 anexin V/PI staining.

	Untreated			H_6-ApopΔProΔLeu		
	24 h	36 h	48 h	24 h	36 h	48 h
Early apoptotic cells (%)	4.0±0.5	3.4±0.7	4.8±1.0	8.6±0.9	18.2±0.1	10.7±1.0
Late apoptotic cells (%)	10.0±0.5	9.1±0.5	10.4±0.6	19.1±1.2	51.7±4.1	73.0±1.1
Necrotic cells (%)	0.3±0.1	0.3±0.1	0.2±0.1	0.9±0.1	2.0±0.9	1.7±0.6
Viable cells (%)	85.7±0.1	87.2±1.1	84.7±1.6	71.5±2.0	28.0±5.0	14.6±0.6

NCI-H460 cells were treated with 15 μM H_6-ApopΔProΔLeu for 24, 36 and 48 h, stained with annexin V-Alexa 488 and PI and analyzed by flow cytometry. Cells undergoing early apoptosis were positive for annexin V and negative for PI (annexin V+/PI-), late apoptosis cells were V+/PI+ and necrotic cells were annexin V-/PI+. All data are expressed as mean ±SE of three different experiments. Values were analyzed from 10,000 total events.

Preliminary structure characterization of H_6-ApopΔProΔLeu

I have shown here that most of the tested biological properties of Apoptin are maintained by H_6-ApopΔProΔLeu and therefore, the characterization of this variant could help to understand the structural properties of Apoptin. *In silico* low resolution structural models for full length Apoptin and its complex with the Bcl-Abl oncoprotein have been advanced (Panigrahi et al., 2012) and CD spectra of Apoptin-H_6 indicated that this protein might be devoid of α-helical or β-sheet structure (Leliveld et al., 2003a). However, up until now the conformation of Apoptin at atomic resolution remains unknown, mainly because the protein tends to form large soluble aggregates that thwart characterization by high resolution techniques like X-ray crystallography or NMR. The sequence of H_6-ApopΔProΔLeu is shown in Figure 31.

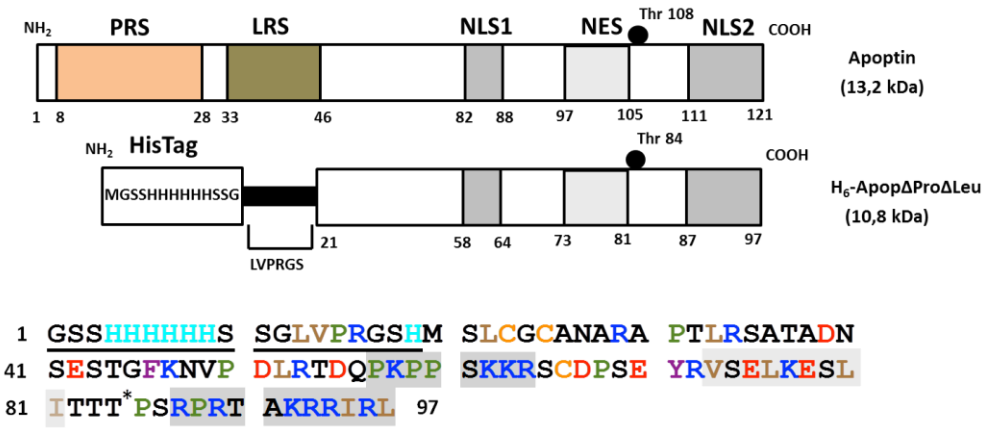


Figure 31. Primary structure of H_6-ApopΔProΔLeu. Top: Comparison of the functional regions of H_6-ApopΔProΔLeu and Apoptin. Bottom: Sequence of H_6-ApopΔProΔLeu. Nuclear localization sequences (dark grey) and the nuclear export sequence (light grey). Anionic residues (Asp & Glu) are colored red and cationic residues (Arg & Lys) are colored blue. His residues, which bear a positive charge at pH 5.5 but become neutral above pH 6.5, are colored cyan. The sequence contains an excess of positively charged residues; in particular, the nuclear localization sequences are highly cationic (Asp + Glu = 8; Arg + Lys = 17; Arg + Lys + His = 24). The construct's three Cys are colored orange and the ten aliphatic residues are colored brown. The sequence is rich in Pro (green) but is poor in aromatic residues; the lone Phe and Tyr residues are colored purple. Thr 108 (Thr 85 in this construct), whose specific phosphorylation by cancer specific kinases putatively masks the nuclear export sequence (Zhuang et al., 1995), is marked with an asterisk.

We analyzed *in silico* the probability that H_6-ApopΔProΔLeu was structurally disordered using different prediction software programs. These were PONDR® (Romero et al., 2001), IUPred (Dosztányi et al., 2005) and FoldIndex© (Prilusky et al., 2005).

PONDR® takes into account features like hydrophobocity or the fractional composition of particular residues to provide a value for every residue which will be considered disordered if greater than 0.5. It uses two binary prediction tools, the charge-hydropathy plot (CH-plot) (Uversky et al., 2000) and the cumulative distribution function (CDF) (Xue et al., 2009). The CDF predicts an order-disorder score for every residue in the protein,

based on the distribution of prediction scores, that have been stablished determining the relative characteristics of three sets of proteins (i.e. completely disordered, fully ordered and ordered proteins with disordered residues). According to CDF analysis, proteins with curves below the given boundary (i.e. high percentage of residues with high predicted disorder scores) are predicted to be disordered while predicted ordered proteins have curves above the boundary. The CH-plot compares the absolute, mean net charge (neglecting histidine) and the mean, scaled Kyte-Doolittle hydropathy (Kyte and Doolittle, 1982). Ordered and disordered proteins are plotted in CH-space and separated by a linear boundary above which, predicted disordered proteins will be located. The IUPred dataset is based in globular proteins with known structures and it estimates the capacity of polypeptides to form stabilizing contacts. It takes into account the contribution of a residue to order/disorder depending not only on its own chemical type, but also on its sequential environment, including its potential interaction partners. Finally the FoldIndex© is based on the average residue hydrophobicity and net charge of the protein sequence. It implements an algorithm to predict if a given protein sequence is intrinsically unfolded, providing a single score for the entire sequence. All positive values represent proteins (or domains) likely to be folded, while negative values represent those likely to be intrinsically unfolded. Foldindex© allows examination of the fold properties of overlapping segments, or sliding windows to identify ordered (represented in green) and disordered (red) regions within the protein.

The IUPred analysis of unstructured regions in the H_6-ApopΔProΔLeu sequence envisages that it is a disordered protein since scores of each residue in the sequence are above 0.5 (Figure 32-A). The two distinct binary classification methods of PONDR® also predict that the Apoptin variant has a disordered conformation. According to former analysis the protein appears under the linear limit of the CDF boundary (Figure 32-B).

In the CH-plot, H_6-ApopΔProΔLeu appears in the disordered region (Figure 32-C). It has been proposed that when CDF and CH predictors both predict the protein to be disordered, it will most probably be unfolded (Huang et al., 2012). Finally, the FoldIndex© algorithm predicts once again that H_6-ApopΔProΔLeu is disordered since the value obtained is -0.15 ± 0.10. In this case, if attention is paid to the whole sequence of the protein, the disorder is predicted for the segment from residue 28 to 98, but certain order is predicted at the N-terminus (Figure 32-D).

These predictions are in accordance with previously described CD results indicating that Apoptin mostly lacks secondary structure (Leliveld et al., 2003a). On the other hand, the comparison of the SDS-PAGE and size-exclusion chromatograms of denatured protein (Figures 19-21), could indicate that de LMM form of H_6-ApopΔProΔLeu is a monomer and therefore it has an anomalous elution volume in size-exclusion chromatography and a high Stoke radius in DLS experiments. Finally, they were in accordance with the fact that no apparent sigmoidal unfolding transition was detected in any of the thermal denaturation experiments performed with H_6-ApopΔProΔLeu (Figure 22). Taking all these considerations altogether, it can be hypothesized that this variant of Apoptin would be poorly structured. Taking advantage of the fact that the LMM form of H_6-ApopΔProΔLeu is a small oligomer or (more probably) a monomer, we aimed to study the conformation and dynamics of this variant.

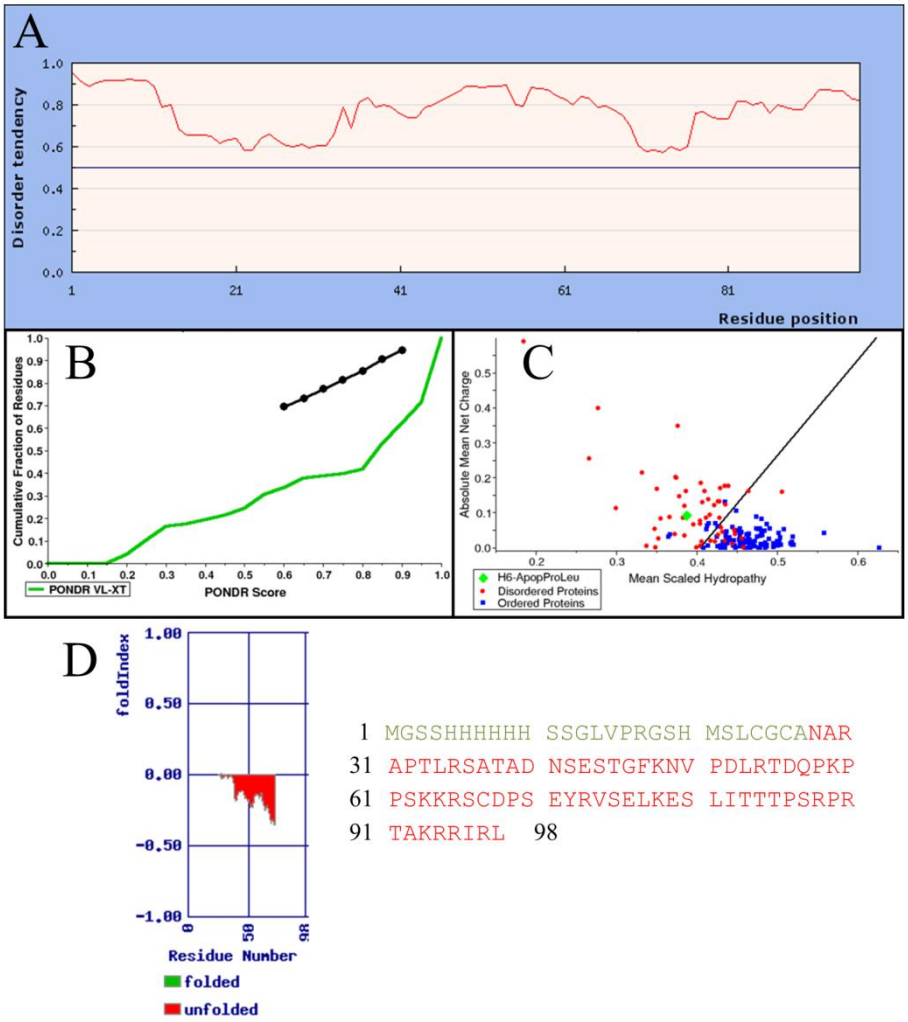


Figure 32. H_6-ApopΔProΔLeu is predicted to be an intrinsically disordered protein. In silico disorder analysis of H_6-ApopΔProΔLeu using IUPred PONDR and FoldIndex© prediction programs. **A)** IUPred prediction of unstructured regions in the H_6-ApopΔProΔLeu sequence. Scores above 0.5 indicate disorder. **B)** Cumulative distribution function (CDF) analysis of H_6-ApopΔProΔLeu. CDF summarizes the predicted order/disorder of every residue by plotting PONDR scores vs. their cumulative frequency. The appearance under the linear limit is an indicator of disorder. **C)** Charge-hydropathy (CH) plot of H_6-ApopΔProΔLeu. In the CH plot which is based on the net charge and hydropathy of a protein, its presence above the linear boundary also suggests a disorder state. **D)** The FoldIndex© (plotted with window size 51) predicts disorder within the fragment comprising residues 28 – 98 (red), while the first part of the protein shows certain order (green).

I carried out a primary characterization of the compactness and degree of structure of this variant through the characterization of ANS binding, CD and protease digestion.

ANS binding

ANS is used to detect the exposure of hydrophobic residues in the native or molten globule states of a protein (Cardamone and Puri, 1992; Semisotnov et al., 1991). This assay is based on the fact that the hydrophobic exposed clusters of the protein bind to the nonpolar aniline-naphtalene moiety of the ANS molecule. It has been previously described that Apoptin binds to ANS (Leliveld et al., 2003c). I investigated the binding of ANS to H_6-Apoptin and its two variants. Both, H_6-Apoptin and H_6-ApopΔPro induce an increase in fluorescence emission intensity of 15.3 and 40.5 fold, respectively, when excited at 385 nm and compared to free ANS. It is noteworthy that the intensity of the peak of fluorescence emission of H_6-ApopΔPro is 2.6-fold higher than that of H_6-Apoptin (Figure 33). In contrast, the emission spectrum of H_6-ApopΔProΔLeu is the same as that of free-ANS. Taken together these data suggest that ANS would be mainly binding to the LRS sequence and that H_6-ApopΔProΔLeu does not have exposed hydrophobic patches.

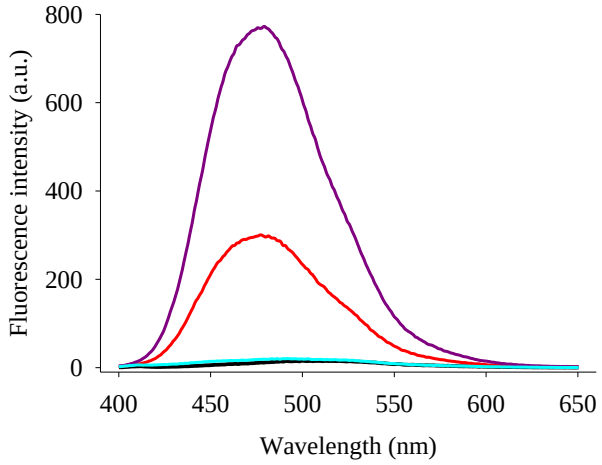


Figure 33. Fluorescence emission spectra of ANS/H_6-Apoptin (red line) and its two purified variants, H_6-ApopΔPro (purple line) and H_6-ApopΔProΔLeu (cyan line). Free ANS spectrum is represented with a black line.

CD characterization

The far UV CD spectrum of H_6-ApopΔProΔLeu is shown in Figure 34. The spectra recorded at 5 and 25°C are very similar. The minimum near 200 nm is characteristic of denatured proteins; in fact, according to the scheme proposed by Uversky and Fink (Uversky and Fink, 2004), H_6-ApopΔProΔLeu would be classified as a random coil on the basis of its $[\Theta]_{200nm}$ and $[\Theta]_{222nm}$ values. Using the Selcon3 algorithm, a more quantitative spectral analysis, the secondary structural content is estimated to be 10% helix, 9% β-strand, 14% turn and 67% coil.

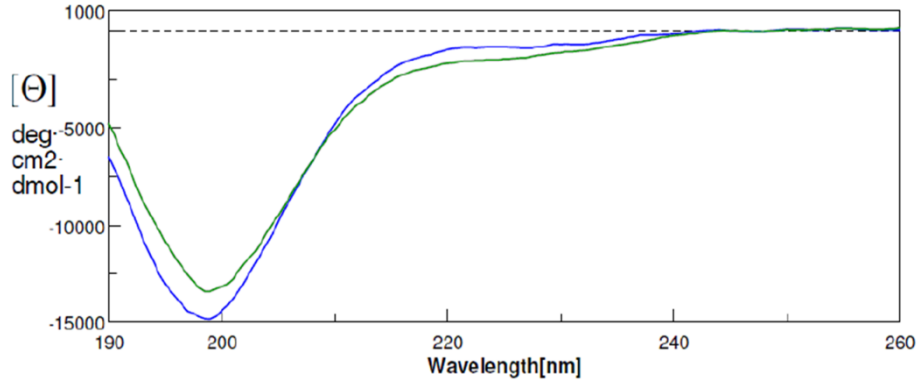


Figure 34. Far UV-CD spectra of H_6-ApopΔProΔLeu at 5ºC (blue) and 25ºC (green).

Protease digestion

I also tested the resistance of H_6-Apoptin and its variants to proteinase K digestion. Interestingly, the three proteins were completely cleaved after 1 h of digestion (Figure 35). In summary, the results of all these analysis suggest again that H_6-ApopΔProΔLeu lacks a well-defined structure. To corroborate these findings and to identify and characterize possible preferred structures that are partially populated or formed in a particular region of the protein, we turned to NMR spectroscopy.

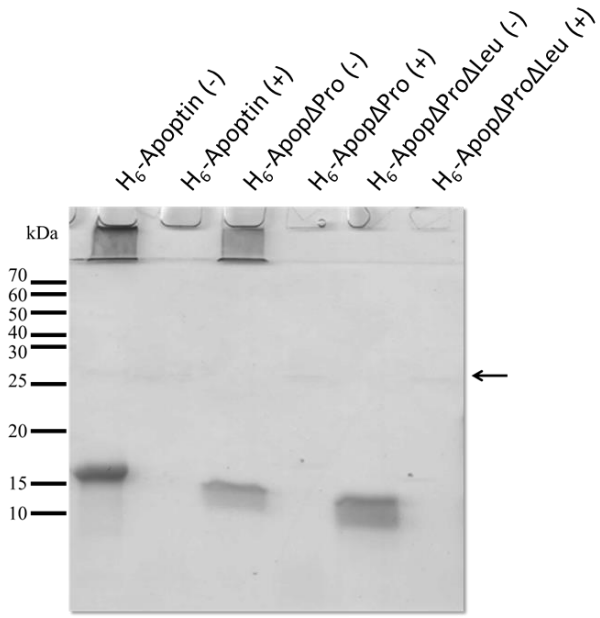


Figure 35. SDS-PAGE electrophoresis of H_6-Apoptin and its two variants, H_6-ApopΔPro and H_6-ApopΔProΔLeu after 1h incubation with proteinase K (+). In control samples (-), PMSF was added before proteinase K (indicated with an arrow).

Structural characterization of H_6-ApopΔProΔLeu by Nuclear Magnetic Resonance (NMR)

IDPs are proteins that lack fixed tridimensional structure in their putatively native states, either entirely or in part (Sickmeier et al., 2007). These proteins exist as dynamic ensembles of conformations and exert their biological function without a prerequisite stably folded structure.

These proteins are implicated in cell signaling, molecular recognition and regulation activities. IDPs are highly abundant in all species with an unequal evolutionary distribution, being more abundant in eukaryotes than prokaryotes. According to computational predictions, typically 10-35% of prokaryotic proteins contain long disordered regions of more than 30 consecutive residues, whereas in eukaryotes the fraction of such proteins reaches 15-45% (Tompa and Fersht, 2009). IDPs have also been

found in viruses; some assist in the construction of their capsid whereas others are viral regulatory and accessory proteins (Uversky and Longhi, 2011).

IDPs have a characteristic amino acid composition namely, a high content of charged residues which leads to charge repulsion, and a low content of the hydrophobic ones, which means less driving force for protein compaction.

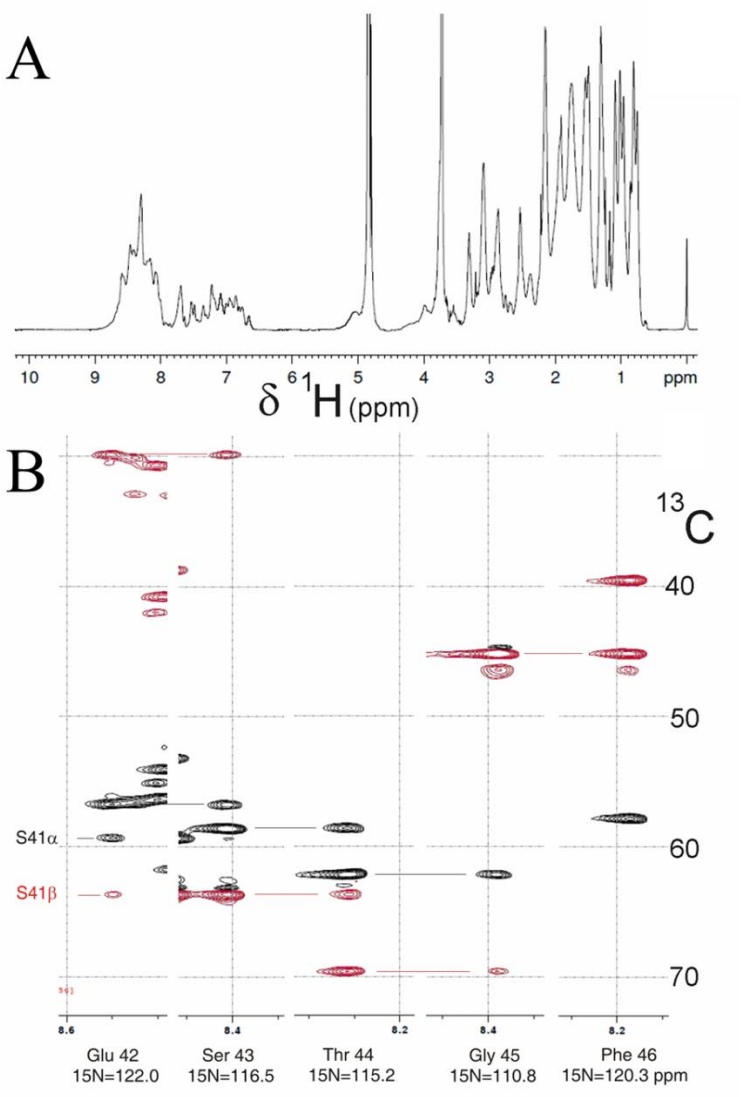
Nowadays, NMR spectroscopy and X-ray crystallography are the most powerful techniques to obtain site-specific information on the structure and dynamics of proteins. However, for characterizing disordered regions or intrinsically disordered proteins, X-ray crystallography is not suitable, since the corresponding regions are missing from electron density map due to their intrinsic flexibility. Also, IDPs typically possess highly dynamic conformations in solution with high mobility at different timescales, and therefore such proteins almost never form crystals, limiting the use of X-ray crystallography for this type of proteins. Moreover, if they formed crystals, they would be locked in a structure that might not be physiologically relevant.

H_6-ApopΔProΔLeu was expressed in minimal media containing $^{15}NH_4Cl$ alone or $^{15}NH_4Cl$ and ^{13}C-glucose for uniform isotopic labeling with ^{15}N or ^{15}N ^{13}C, respectively. The purified proteins were judged to be pure by SDS-PAGE and their characterization by MALDI-TOF showed that the constructs had the correct molecular weight and an average incorporation of ^{13}C and ^{15}N of 95%.

The 1D NMR spectrum of H_6-ApopΔProΔLeu (Figure 36-A) shows poor chemical shift dispersion, which is a characteristic of denatured proteins. In particular, the backbone amide HN signals are clustered between 8.0

and 8.75 ppm. Furthermore, the native protein signals resonating at high field (< 0.7 ppm) are absent in the spectrum.

The resonances of H_6-ApopΔProΔLeu backbone nuclei and the aliphatic side-chain carbon resonances were assigned using conventional triple resonance spectra with amide proton signal detection, following connections between adjacent Cα and Cβ nuclei (Figure 36-B). The assignments were corroborated by analysis of the ^{1}H-^{15}N HSQC spectra filtered in ^{13}C on the basis of the side chain structure and the (4, 2)D $^{1}H^{15}N$-$^{1}H^{15}N$ spectrum. The later enabled the complete assignment of the ^{1}H-^{15}N resonances of the His-tag, some of which showed split signals. The spectra which select residues preceding or following prolines were particularly useful to identify Lys58 (Lys82 in wild-type Apoptin), which is trapped between two prolines. A total of 84 out of the 88 non-proline residues (omitting fMet) were unequivocally assigned (Figure 36-C). The four remaining residues correspond to the Leu22-Cys23-Gly24-Cys25 stretch (Figure 31). The Cα and Cβ chemical shifts values of Cys25 (in this construct) can be observed in 3D experiments via correlations with the HN of Ala26 and correspond to the oxidized form of cysteine (cystine) as do the $^{13}Cα$ and $^{13}Cβ$ values of Cys66 (Table 6). H_6-ApopΔProΔLeu contains nine proline residues, eight of which could be assigned based on the connectivities with the preceding residue. All of these Pro are in the *trans* configuration, as inferred by the differences in the proline $^{13}Cγ$ and $^{13}Cβ$ chemical shifts (Table 5) (Schubert et al., 2002). Differences in the proline $^{13}Cγ$ and $^{13}Cβ$ chemical shifts of around 5 ppm indicate a *trans* conformation, while differences of 9 ppm would indicate a conformation in *cis*. The remaining proline (Pro59 in this construct, Pro83 in wild type Apoptin) could not be analyzed; its signals are missing as it is followed by another Pro residue.


A
10
9
8
7
6
5
4
3
2
1
ppm
δ ^{1}H (ppm)
B
^{13}C
40
50
60
70
S41α
S41β
8.6
8.4
8.2
8.4
8.2
Glu 42
15N=122.0
Ser 43
15N=116.5
Thr 44
15N=115.2
Gly 45
15N=110.8
Phe 46
15N=120.3 ppm

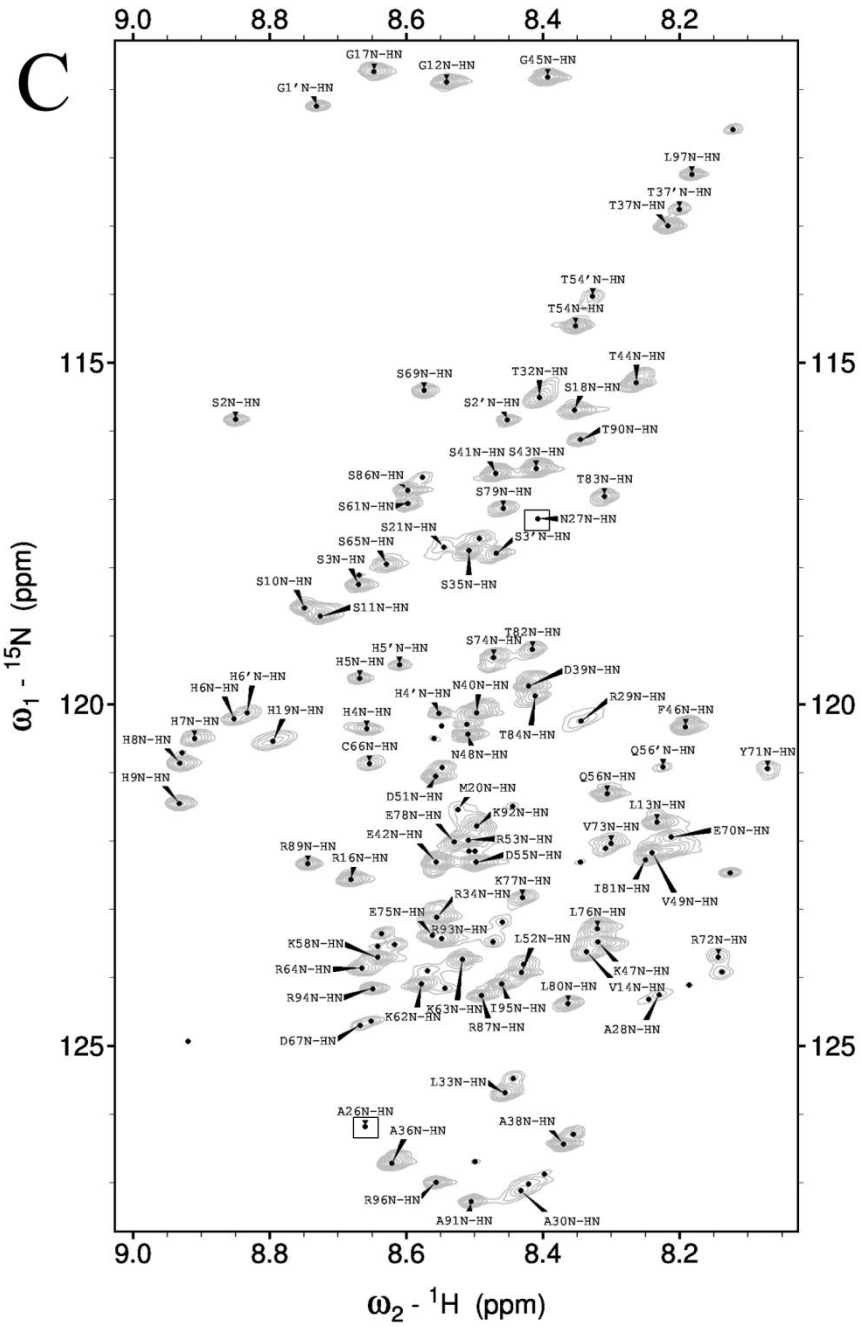


Figure 36. A) 1D ^{1}H NMR Spectrum of H_6-ApopΔProΔLeu. **B)** Strip plot showing ^{15}N planes with consecutive Cβ (red) and Cα (black) connectivities for residues 42-46. The y-axis is the ^{13}C dimension and the chemical shifts range from 70 ppm (bottom grid line) to 30 ppm (top grid line). **C)** Assigned ^{1}H-^{15}N HSQC spectrum of H_6-ApopΔProΔLeu. The boxed regions indicate the position of cross peaks whose intensity was weak at the selected signal threshold.

Table 5. Chemical shift differences of $^{13}C\beta$ and $^{13}C\gamma$ of prolines and their conformation.

Residue[a]	$^{13}C\beta$ (ppm)	$^{13}C\gamma$ (ppm)	$^{13}C\beta$-$C\gamma$ (ppm)	Conformation
(R)P15	29.186	24.526	4.660	*Trans*
(T)P31	29.317	24.389	4.928	*Trans*
(D)P50	29.411	24.375	5.036	*Trans*
(K)P58	29.644	24.234	5.410	*Trans*
(S)P60	29.311	24.355	4.956	*Trans*
(S)P68	29.159	24.177	4.982	*Trans*
(S)P85	29.311	24.355	4.956	*Trans*
(R)P88	29.186	24.526	4.984	*Trans*

Note that only eight out of the nine prolines present in the H_6-ApopΔProΔLeu sequence are analyzed since the remaining one (Pro59) is preceded by another proline.
[a] The residue preceding each proline residue is shown in brackets.

The obtained chemical shift values were compared with chemical shift values predicted for this sequence at 5°C and pH 5.2 in a random coil ensemble. The differences are small and have no coherent pattern, which indicates that there are no portions of the polypeptide with significant tendencies to adopt preferred secondary structure (Figure 37-A). Similar results were obtained upon comparison of the experimental data to chemical shift values predicted for intrinsically disordered proteins (data not shown). The experimental and predicted chemical shift values are strongly correlated (Figure 37-B and 37-C) and the RMSD values for the $^{13}C\alpha$, $^{13}C\beta$, ^{13}CO and $^{1}H\alpha$ are 0.23, 0.24, 0.31, and 0.05 ppm, respectively (Figure 37-D). These are small values that are typical of unfolded proteins (Tamiola et al., 2010).

As an additional test for the presence of stable secondary structure, we measured the H/D exchange rates for H_6-ApopΔProΔLeu. For most backbone HN groups, the exchange with deuterium was too fast to permit an accurate and precise determination of the H/D exchange kinetics.

Whereas the H/D protection factors are typically very high – hundreds of thousands to millions – for stably folded proteins, those of H_6-ApopΔProΔLeu are all less than 100 (Figure 37-E). This indicates that no stable secondary structure or intramolecular H-bonding network is formed in this Apoptin variant.

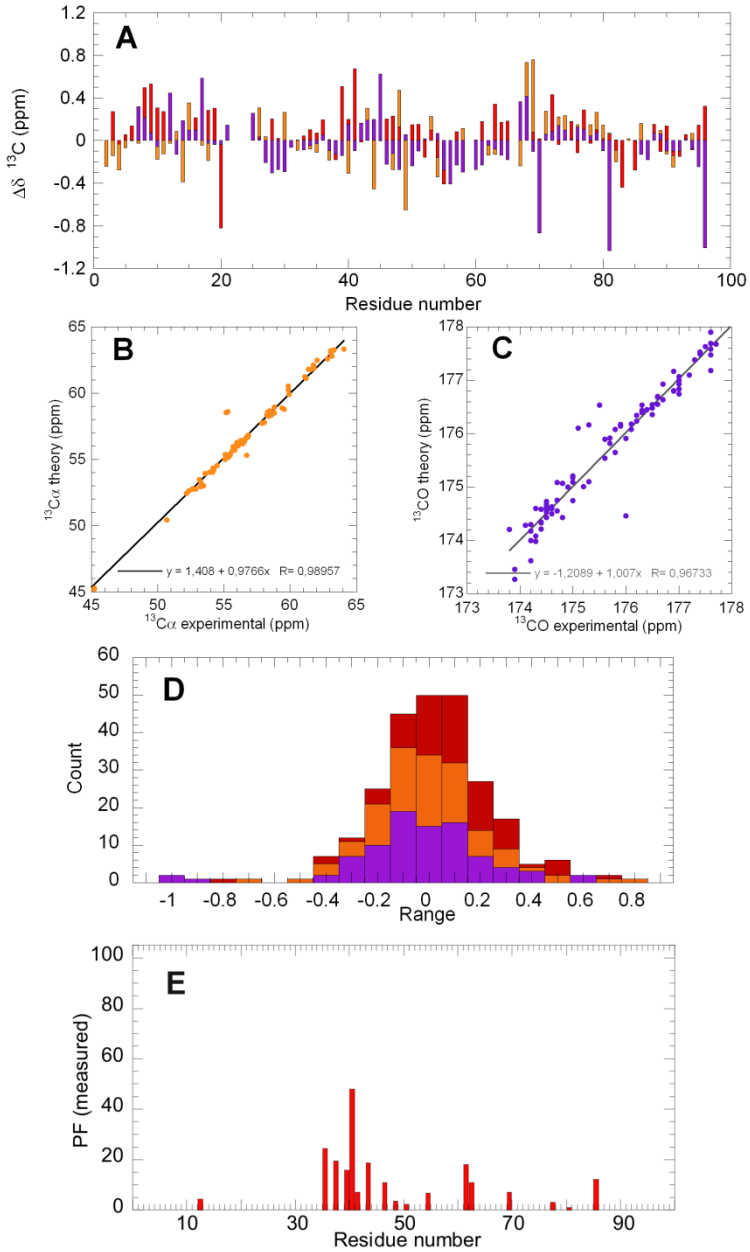


Figure 37. Structural analysis of chemical shift values. **A)** Conformational chemical shifts for ^{13}Cα (orange), ^{13}Cβ (red) and ^{13}CO (purple). **B)** Correlation between experimental and predicted ^{13}Cα and **C)** ^{13}CO chemical shift values. **D)** Distributions of the differences between experimental and predicted chemical shifts for ^{13}Cα (orange), ^{13}Cβ (red) and ^{13}CO (purple). **E)** Measured Protection Factors obtained from measurement of the H/D exchange kinetics. Only 17 HNs groups could be followed; the rest of them exchange too fast to permit their kinetics to be measured.

Apoptin backbone dynamics

The backbone dynamics of H_6-ApopΔProΔLeu were characterized by measuring the ^{15}N longitudinal relaxation rates (R_1), ^{15}N transverse relaxation rate (R_2) and heteronuclear $\{^1H\}$-^{15}N NOE (Figure 38). The heteronuclear $\{^1H\}$-^{15}N NOE, R_1 and R_2 values were measured for 68 of the 97 residues (all except the N-terminal residue, the nine prolines and 19 residues whose signals were very weak or overlapped).

The R_1 and NOE values provide information in backbone motions occurring on a fast time scale (ps to ns), while R_2 monitors conformation exchange processes on the μs to ms timescale. The $\{^1H\}$-^{15}N NOE average value is 0.4 ± 0.1, which is characteristic of a rather flexible protein chain as it is far from the value expected for a completely rigid chain (0.85). However, there are a few segments that show above average values (0.52-0.65) corresponding to the residues His6-His9, Gly17-Ser18 and Ser65-Arg72. It is interesting to note that residues 16-17 are adjacent to the Leu22-Cys23-Gly24-Cys25 stretch that could not be assigned (see above) and that Cys66 is in the third rigid segment. Therefore all three Cys residues are in or close to relatively rigid segments. It is tempting to speculate that these Cys could somehow increase the rigid behavior of these regions. Both motional anisotropy and chemical exchange contribute to the value of R_2. Normally, the motional anisotropy can be eliminated by taking the ratio between the transverse and longitudinal relaxation rates (Figure 38-D). The region spanning residues Ser65-Lys77 has higher than average values; this indicates that these residues are undergoing chemical exchange processes.

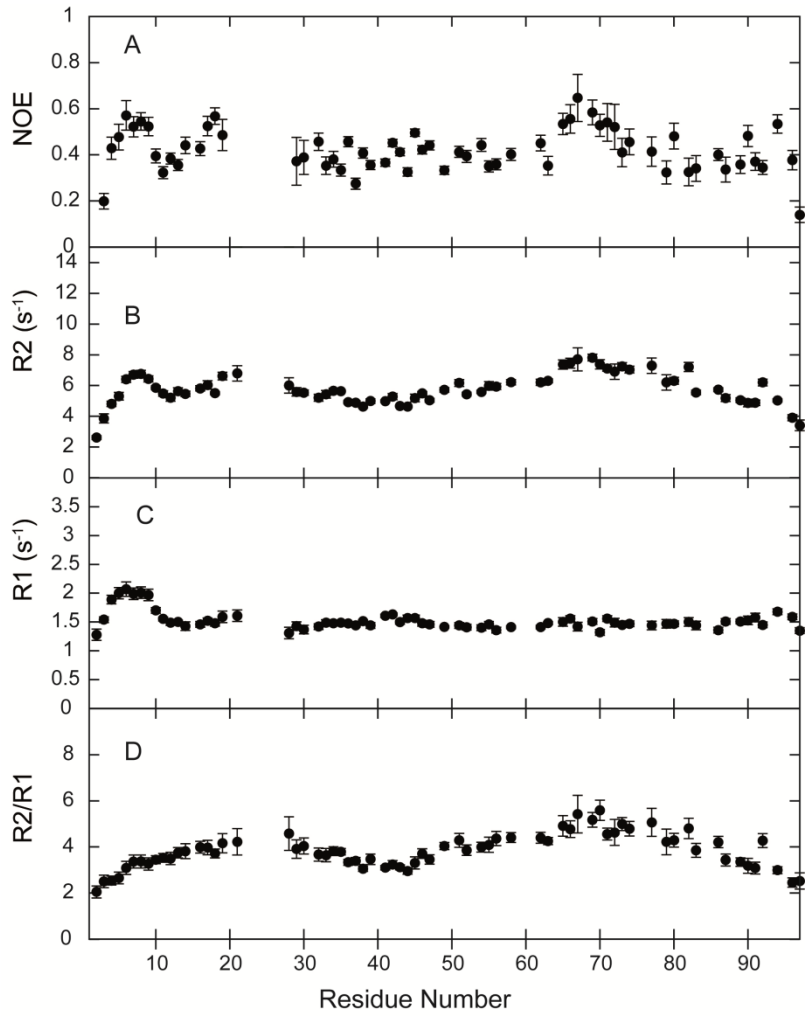


Figure 38. Backbone NMR relaxation data for H_6-ApopΔProΔLeu at pH 5.4, 5ºC. **A)** Heteronuclear {^{1}H}-^{15}N NOEs, **B)** ^{15}N transverse, R_2 and **C)** longitudinal, R_1 relaxation rates and **D)** ratio R_2 by R_1. The error bars in the R_1 and R_2 data denote curve-fitting uncertainties. The errors in the heteronuclear {^{1}H}-^{15}N NOEs are estimated from signal/noise ratio of the spectra.

H_6-ApopΔProΔLeu contains three of the four cysteines present in the wild type Apoptin. ^{13}Cβ chemical shift data indicate that Cys25 and Cys66 are oxidized (Table 6). In MBP-Apoptin aggregates, only Cys90 (Cys66 in H_6-ApopΔProΔLeu) appears to be reactive with DTNB (Leliveld et al., 2003c). To test the effect of the oxidation state of the Cys residues on H_6-ApopΔProΔLeu's conformational ensemble, we recorded 1D ^{1}H and 2D ^{1}H

TOCSY and NOESY and 2D ^{1}H-^{13}C HSQC spectra in the absence and presence of 4.7 mM TCEP. This reducing agent is active over a broad range of pH values and cannot form mixed disulfide bonds with Cys like β-mercaptoethanol or DTT can. In independent samples, we recorded 1D ^{1}H spectra of freshly dissolved TCEP and TCEP oxidized by an excess of hydrogen peroxide. These spectra revealed multiplets at 2.6 and 2.4 ppm for reduced TCEP and signals at 2.5 and 2.1 ppm for oxidized TCEP (data not shown). In 1D ^{1}H spectra of the H_6-ApopΔProΔLeu sample containing TCEP recorded before and after the acquisition of the series of 2D spectra, signals of reduced TCEP were clearly observed. Considering that TCEP is a strong and irreversible reducing agent, we can be sure that these Apoptin variant's Cys are reduced under these conditions (Table 6).

Table 6. Values obtained for the computed $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts of Cys before and after the addition of the reducing agent TCEP.

Residue	Carbon	Oxidized	Reduced
C23	α	-	58.66
	β	-	27.74
C25	α	55.19	58.46
	β	40.58	28.06
C66	α	55.29	58.06
	β	40.98	28.09

The 2D ^{1}H, ^{13}C HSQC spectra and the aliphatic, Hα and aromatic portions of the 1D ^{1}H and 2D ^{1}H TOCSY and NOESY spectra are very similar under oxidized versus reduced conditions, except that some novel peaks, which arise from TCEP and reduced Cys are seen in the latter (Figure 39). In the amide ^{1}H region of the 1D, and 2D TOSCY and NOESY spectra, the resonances generally retained their chemical shift values in the presence of TCEP but become narrower. Additionally, only under reducing conditions we could see the signal corresponding to Gly24 (Figure 40). By

adding the reducing agent, all the different disulfide bonds were broken and the reduced Cys adopted a more similar set of conformers, which allowed Gly 24 to be detected. In any case, the fact that H_6-ApopΔProΔLeu is an IDP is independent of the redox conditions.

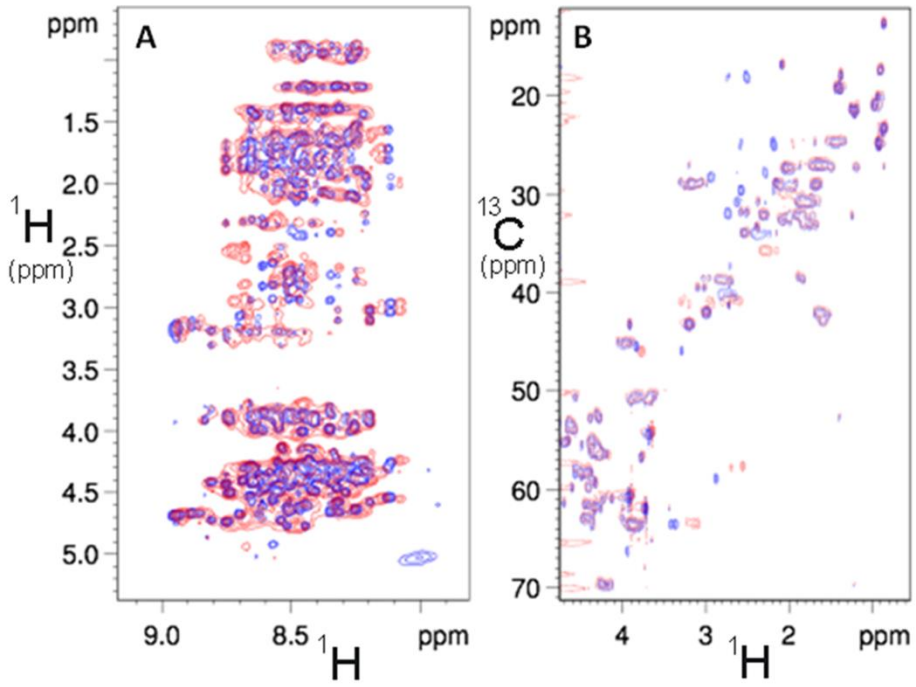


Figure 39. ^{1}H 2D NOESY **A)** (150 ms mixing time) and ^{1}H-^{13}C HSQC **B)** of H_6-ApopΔProΔLeu under oxidizing (red) and reducing (blue) conditions. The amide ^{1}H of the NOESY spectra and aliphatic regions of the HSQC spectra are shown. The reduced and oxidized spectra are plotted using the same base contour and multiplication factors between contours.

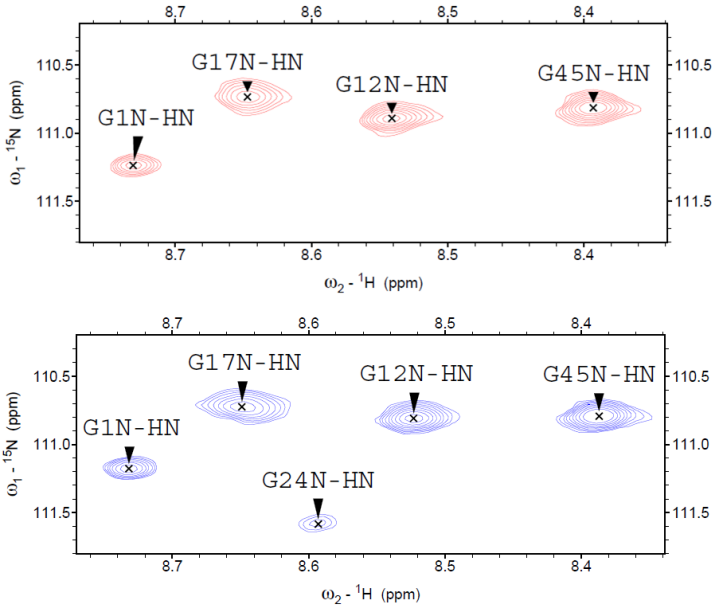


Figure 40. Portion of the ^{1}H-^{15}N HSQC spectra before (upper) and after (lower) TCEP addition. Note the appearance of the Gly24 under reducing conditions.

The six consecutive His located at the N-terminus of the variant were successfully assigned using standard 3D NMR spectra and corroborated by analysis of ^{1}H-^{15}N/^{1}H-^{15}N 4(2)D spectra. The ^{1}H-^{15}N signals of all six His residues could be unambiguously assigned and most show two distinct signals in the ^{1}H-^{15}N HSQC spectrum (Figure 41). This suggests that their conformational ensemble is not completely disordered.

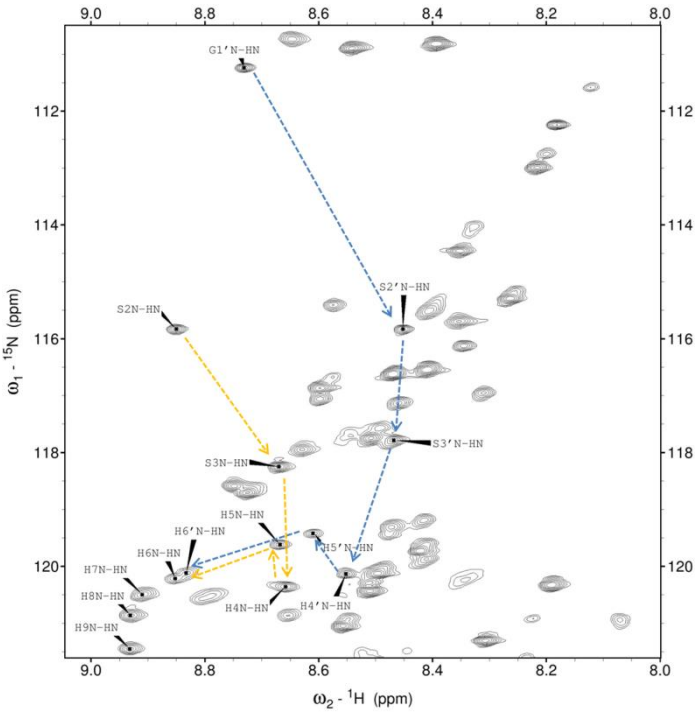


Figure 41. Portion of the ^{1}H-^{15}N HSQC spectrum showing the two alternative conformations. The consecutive signals assigned by ^{1}H-^{15}N/^{1}H-^{15}N 4(2)D experiment are indicated (blue and orange arrows).

In addition, this region appears slightly more rigid than some other portions of the construct as gauged by the relaxation experiments (*vide supra*). To test if the relative stiffness of the His-tag region might be due to metal cation coordination, we added EDTA to a final concentration of 1 mM. EDTA is a strong chelator relative to imidazole (K_a=4x10^{18} M for Ni^{+2}, for example) and should extract and bind any divalent or trivalent cations present. However, no significant changes in the dispersion of the His-tag signals were observed (Figure 42-A and 42-B). This strongly suggests that cation binding is not responsible for the presence of putative alternative conformations in the His-tag. As a further test, an excess (5 mM) of $MgCl_2$ was added to the sample and additional spectra were recorded. Except for an apparent reduction in the line width of some of the signals, no large changes in the spectrum were seen (Figure 42-C). Considering all these

results, we conclude that metal cation binding to the His-tag is not responsible for the relative rigidity of this segment.

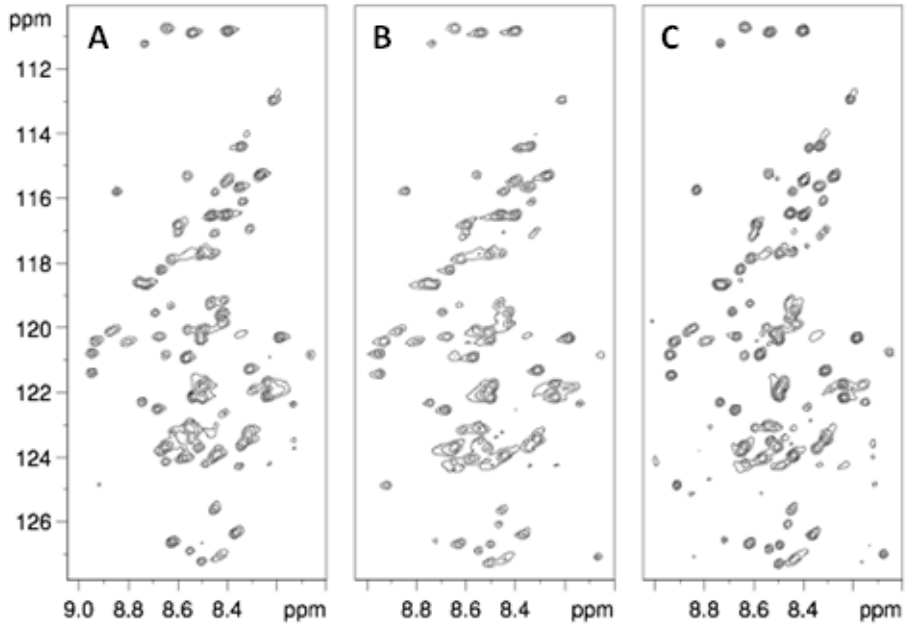
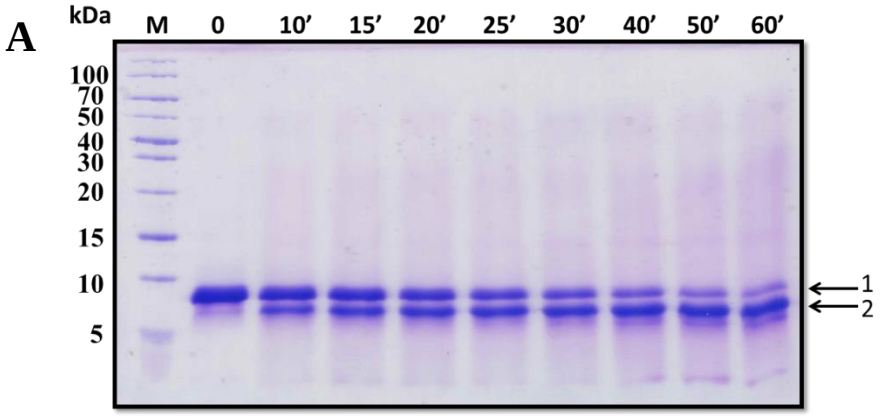


Figure 42. ^{1}H-^{15}N HSQC spectra of H_6-ApopΔProΔLeu **A)** as prepared for the 3D spectra at pH 5.4, **B)** in the presence of 1 mM EDTA **C)** and after the addition of $MgCl_2$ to a final concentration of 5 mM.

The most direct way to test the possible effects of the His-tag would be to record NMR spectra on H_6-ApopΔProΔLeu constructs that lack this motif. We endeavored to prepare such a construct, either by attempting to express a genetic construct lacking the His-tag coding sequence, or alternatively by its posttranslational proteolytic cleavage, as described in the following paragraphs.

The purification of the H_6-ApopΔProΔLeu had been achieved taking advantage of its N-terminal 6xHisTag, and a different strategy of production of truncated Apoptin had to be developed. The ApopΔProΔLeu was subcloned from the plasmid pET-28a to the pET-22b vector using the *Nde*I and *Sal*I restriction enzymes. BL21(DE3) *E. coli* cells were transformed in order to produce ApopΔProΔLeu, however, after multiple attempts the

protein was never expressed. As an alternative, the protein was produced with an N-terminal His-tag and two different digestion strategies were used to cleave it. In the first one we used the serine protease Thrombin (Sigma, USA), for which a cleavage site (Lys-Val-Pro-Arg-Gly-Ser) was already located after the His-Tag when the protein was expressed using the pET-28a plasmid (Figure 8). The digestion was performed at different incubation times as detailed in the Material and Methods section and the cleavage of the His-tag stretch at each incubation time was followed by SDS-PAGE and MALDI-TOF (Figure 43). After 25 min incubation at 25°C the majority of the protein had lost the 6xHis-Tag. However, at longer times unspecific digestion of the protein took place, generating a protein fragment of about 7.9 kDa. Further ion exchange chromatography was carried out in order to separate the cleaved protein (not shown). Unfortunately an insufficient yield was obtained using this strategy and thereby prevented further NMR studies.



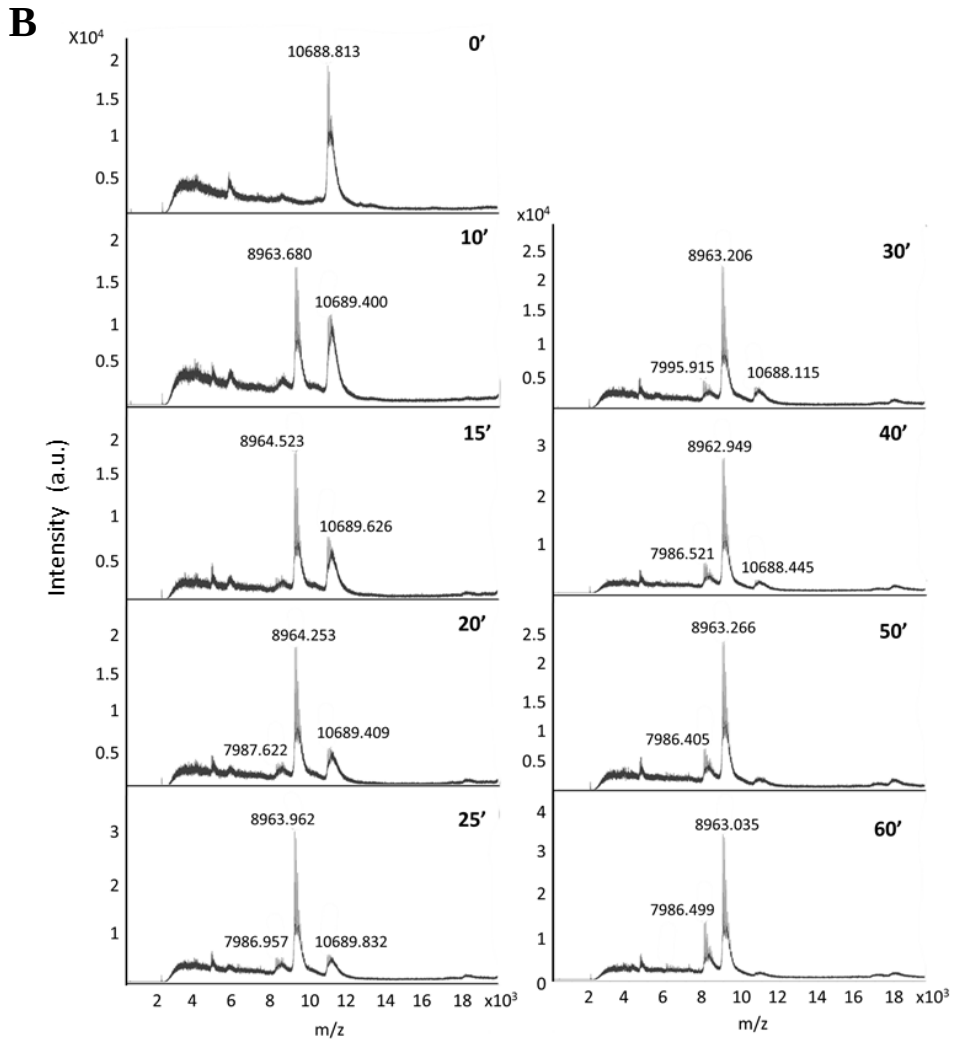


Figure 43. Analysis of the thrombin digestion. **A)** SDS-PAGE analysis of the thrombin digestion at different times. 1 indicates His-tagged protein while 2 indicates cleaved protein **B)** MALDI-TOF mass spectra of the samples withdrawn at different times of the thrombin digestion.

A second digestion strategy, involving the TEV cysteine protease was used. In this case, ApopΔProΔLeu was cloned into pET-22b-derived vector in which the sequence coding for the cleavage site for TEV protease (Gly-Glu-Asn-Leu-Tyr-Phe-Gln-Gly-Ala) had been previously introduced after a 6xHis tag coding sequence (Gordo V.; unpublished PhD thesis) (Figure 7). The protein was produced and purified following the protocol set up for

H_6-ApopΔProΔLeu expressed in pET-28a. The digestion was followed by SDS-PAGE and MALDI-TOF (Figure 44), however the different peaks observed in mass spectroscopy spectra (Figure 44-B) indicated that the protein had been non-specifically cleaved.

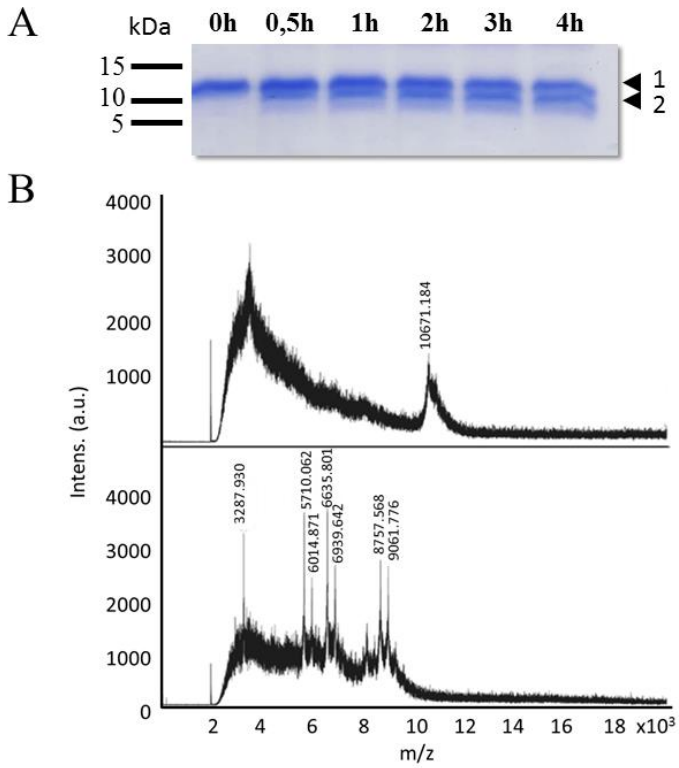
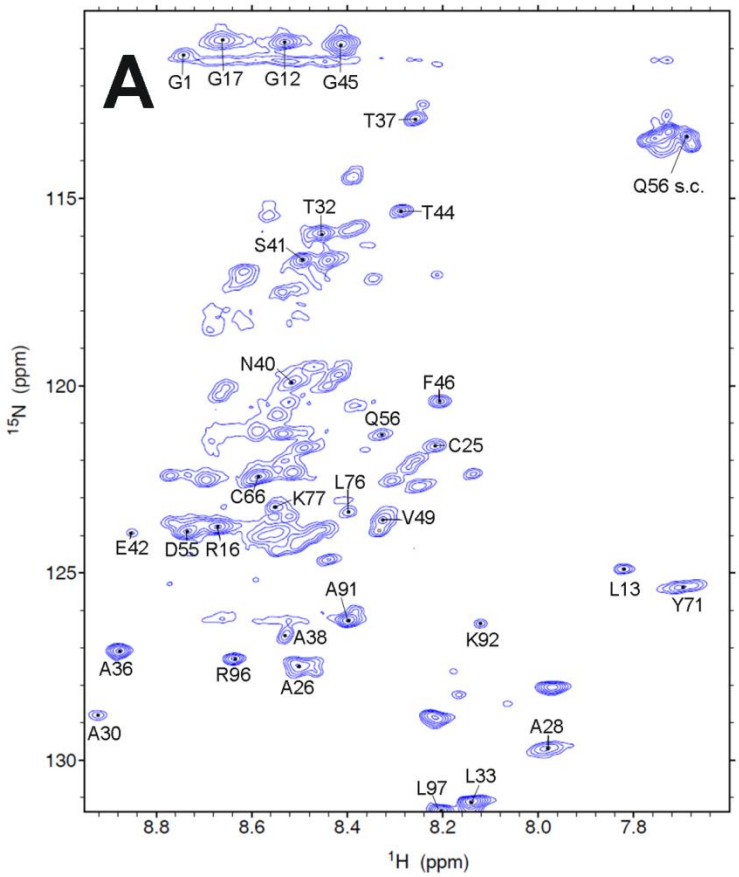


Figure 44. Analysis of the TEV digestion. **A)** SDS-PAGE analysis of the TEV digestion at different times; the arrows indicate undigested (1) and digested (2) protein **B)** MALDI-TOF mass spectra of the TEV digestion at time 0 (upper) and after 4 h (lower).

Unfortunately, despite multiple attempts, none of the envisaged methods produced the desired H_6-ApopΔProΔLeu construct without the 6xHis tag. Therefore, to further explore the conformation of the His-tag and its possible effects on other regions of H_6-ApopΔProΔLeu, we recorded a second series of 2D and 3D spectra at pH 6.7, where His residues are largely in the neutral form. Despite significantly increased line broadening and the absence of some peaks, including those of the His-tag (Figure 45-A), it

was possible to assign the backbone resonances of 31 residues. The conformational analysis of these chemical shifts (Figure 45-B) indicates that H_6-ApopΔProΔLeu is still chiefly unfolded with no apparent tendency to adopt any preferred secondary structure at pH 6.7 and 5°C. This strongly suggests that the His-tag is not responsible for the lack of preferred conformers seen in H_6-ApopΔProΔLeu.



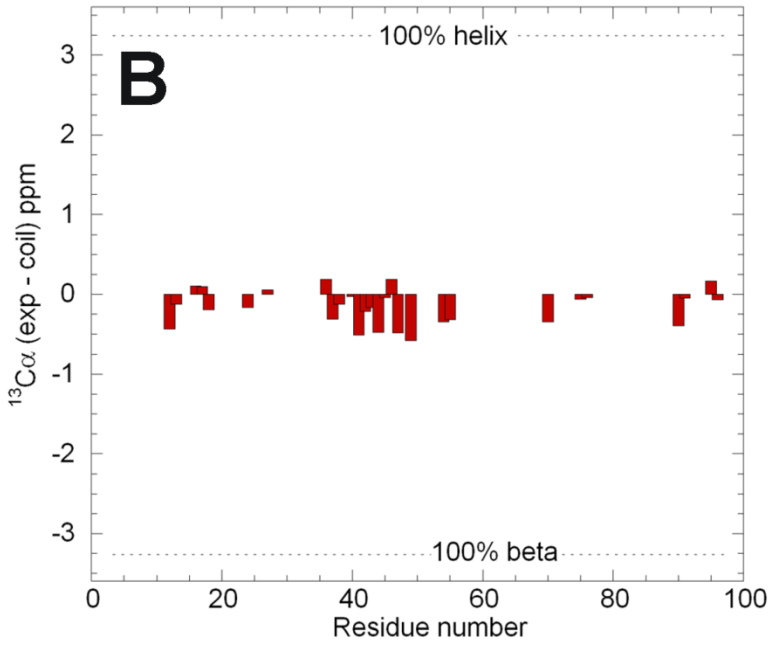


Figure 45. A) ^{1}H-^{15}N HSQC spectrum of H_6-ApopΔProΔLeu at pH 6.7, 5ºC. Note that the signal dispersion is still quite narrow, which means that the protein is still unfolded**. B)** Conformational chemical shifts of H_6-ApopΔProΔLeu at pH 6.7, 5ºC. No significant trends towards preferred structures can be detected. The dotted lines show the values expected for fully helical or sheet conformations.

Thr Phosphorylation does not significantly alter Apoptin's conformational ensemble

H_6-ApopΔProΔLeu was incubated under phosphorylation conditions as described in the Material and Methods Section. Using non-denaturing PAGE, we detected mobility changes consistent with the addition of negatively charged groups to Apoptin (Figure 46-A). MALDI-TOF mass spectrometry analysis of Apoptin after the kinase reaction revealed a mass increase of ca. 76 amu, which is in good agreement with the value expected for a single phosphate group (Figure 46-B). NMR spectroscopy was used to search for the phosphorylation site. Changes in the chemical shift of a Thr $^{13}C\gamma$ were observed that are consistent with the value expected for a phosphorylated Thr (Bienkiewicz and Lumb, 1999) (Figure 46-C). Additional NMR spectra were registered to characterize the

conformation of phosphorylated H_6-ApopΔProΔLeu (Figure 46-D). The spectra are clearly very similar to those of non-phosphorylated Apoptin and on this basis, we conclude that Thr phosphorylation does not induce H_6-ApopΔProΔLeu to adopt a well-defined tertiary structure; instead the protein continues to exist as a disordered conformational ensemble.

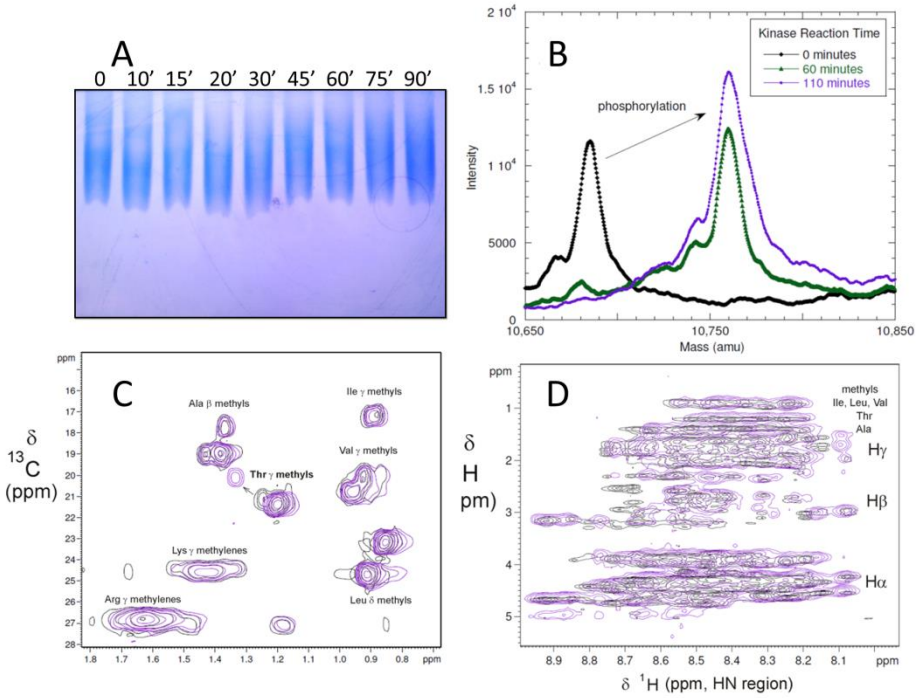


Figure 46. Phosphorylation of H_6-ApopΔProΔLeu followed by **A)** non-denaturing PAGE and **B)** MALDI-TOF mass spectrometry. **B)** Portion of the ^1H-^{13}C HSQC spectra before (black) and after (purple) phosphorylation. The shift of one Thr $^1H^{13}C\gamma$ correlation is marked by an arrow. **C)** ^1H 2D NOESY spectra (150 ms mixing time) of H_6-ApopΔProΔLeu before (black) and after (purple) phosphorylation. The HN to Hα, Hβ, *etc* correlations show little change, indicating that the protein is still unfolded in the phosphorylated state under these conditions.

General Discussion

Apoptin is a very promising antitumor drug candidate, as it possesses a highly specific cytotoxic action on tumor cells, while leaves non-transformed cells unharmed. However, due to the protein's tendency to form large aggregates, effective delivery of Apoptin to tumor cells *in vivo* remains a challenge and limits its therapeutic effectiveness. Protein aggregation may impair biological activity and also constitutes a major risk factor for immunogenicity and the consequent formation of anti-drug antibodies (Ring et al., 1977). We show that the Apoptin variants H_6-Apoptin and H_6-ApopΔPro form multimers composed of approximately 50 monomers. Similar results have been observed with other variants of Apoptin fused to MBP or to a C-terminal His-tag (Leliveld et al., 2003a).

These Apoptin aggregates are spherical-like structures as seen by TEM analysis (Figure 12-B) and resemble the aggregates that give rise to proto-amyloid fibrils (Ahmed et al., 2010). We cannot discard the presence of amyloid-like structures within Apoptin aggregates in the conditions assayed because we have found controversial evidences of this with the techniques commonly used for the detection of amyloid fibrils, i.e. Th T and CR (Figures 13-14). However, further studies using different conditions (i.e. high pH Th T experiments where Apoptin would have a much smaller positive charge), other dyes, or the conformational antibodies, A11 and OC (Kayed et al., 2007) will be necessary to corroborate whether or not the amyloid state is present.

It was previously shown that removal of the first N-terminal 65 residues of Apoptin fused to MBP generated a variant that, according to the size-exclusion elution volumes, might correspond to an equilibrium between a monomer and a dimer or a trimer (Leliveld et al., 2003a). This variant lacked the PRS and LRS consequently indicating that the determinants responsible for Apoptin's aggregation *in vivo* are chiefly located in the N-terminal part of the protein (Leliveld et al., 2003a). Here we show that

removal of residues 1 to 43 but not the smaller deletion of residues 1 to 27 abolishes Apoptin's aggregation and confirms the role of the LRS in this process.

H_6-ApopΔProΔLeu (10.8 kDa) behaves like a globular protein of between 29 kDa (G75 chromatogram, Figure 12-A) and 36 kDa (DLS results, Figure 12-C). These results seem to indicate that H_6-ApopΔProΔLeu could be a trimer. However, this conclusion should be considered with caution because our results and previous experiments based on CD and intrinsic fluorescence (Leliveld et al., 2003a) show that Apoptin is poorly structured, and it is also known that unfolded or partially folded proteins elute at anomalous volumes in the size-exclusion chromatography and have higher R_H in DLS than those expected for globular folded proteins (Gualfetti et al., 1999). Taking this into account, the comparison of the SDS-PAGE and SEC analyses of reduced H_6-ApopΔProΔLeu (Figure 21) seems to indicate that the LMM and HMM forms of H_6-ApopΔProΔLeu could correspond to monomeric and dimeric forms, respectively. Related to these results it is worth mentioning that it would have been interesting to calibrate the gel filtration with IDPs whose MW and oligomeric state are known. However the limited availability of such proteins has made so far this experiment impossible.

It was curious that H_6-ApopΔProΔLeu exists in solution as a monomer or small oligomer that maintains most of the biological properties of Apoptin. It is reasonable to think that the characterization of its conformation would shed light and deepen our understanding of the relationship between structure and function of Apoptin and this has been one of the objectives of this work.

The results obtained for H_6-ApopΔProΔLeu underline the biological importance of the central and C-terminal portions of Apoptin, since large

segments of the N-terminal region can be deleted with only a minor decrease in tumor specific cytotoxicity. In future studies, H_6-ApopΔProΔLeu may work as a control for the creation of full length Apoptin variants with point mutations in the LRS region with the aim of abolishing the tendency to aggregate while retaining the cytotoxicity of the N-terminus (Danen-Van Oorschot et al., 2003).

We have shown that H_6-ApopΔProΔLeu binds to dsDNA although probably to a minor extent relative to H_6-Apoptin or H_6-ApopΔPro (Figure 24). The capability of the multimeric C-terminal truncated form of Apoptin MBP-Apoptin(1-80)-H_6 and the non-multimeric N-terminal truncated form MBP-Apoptin(66–121) to bind dsDNA was reported previously (Leliveld et al., 2004) although the latter binds dsDNA to a minor extent. The fact that H_6-ApopΔPro variant seems to bind DNA at the same degree than H_6-Apoptin would suggest that the LRS may be involved in DNA binding. We show that the large spherical superstructures formed by Apoptin (which have been shown to strongly interact with DNA) are not necessary to specifically kill tumor cells. In accordance, it is known from transfection experiments that the C-terminal region of Apoptin (residues 80-121, lacking the LRS) induces the apoptosis of tumor cells (Danen-Van Oorschot et al., 2003).

Nuclear translocation is at least partially required for the ability of Apoptin to induce cell death (Bullenkamp and Tavassoli, 2014). GFP-ApopΔProΔLeu has a decreased capacity to accumulate into the nucleus compared to GFP-Apoptin but it maintains the differential subcellular localization of the protein in tumor and normal cells. In this regard, it is interesting to note that the LRS has been implicated in the nuclear accumulation of Apoptin since mutations in this region reduce its nuclear accumulation in both tumor and normal cells (Poon et al., 2005b). Nevertheless, despite the decreased nuclear accumulation and DNA

binding of GFP-ApopΔProΔLeu, transfection of tumor cells with the GFP-ApopΔProΔLeu gene produces a cytotoxic effect on tumor cells that is still half of that obtained with GFP-Apoptin (Figure 27-B). Besides, the observation of nuclear changes of tumor cells transfected with either GFP-Apoptin or GFP-ApopΔProΔLeu evidences that the cytotoxic effect of both transfected variants is produced through apoptosis.

Strategies for delivering Apoptin into tumor cells are being developed and most of them are based on the transfection of the Apoptin gene. However, these approaches may provoke insertional mutagenesis in transfected cells or may lack of efficacy due to transient expression. In addition, it has been reported that nuclear localization of Apoptin in cancer cells could depend on the protein intracellular levels instead of the tumorigenic status of the cell (Guelen et al., 2004; Wadia et al., 2004) weakening the selectivity of Apoptin against tumor cells. Different laboratories have described strategies to deliver Apoptin protein into tumor cells by fusing it to a protein transduction domain (PTD). Apoptin variants tagged with the HIV-TAT protein transduction domain (TAT) (Guelen et al., 2004), the C-terminus of human extracellular superoxide dismutase (hC-SOD3) (Zhao et al., 2011) or the protein transduction domain 4 (PTD4) (Sun et al., 2009) were all able to specifically induce apoptosis of tumor cells. Comparison of the *in vitro* antitumor efficiency of these TAT-Apoptin and PTD4-Apoptin with H_6-ApopΔProΔLeu is difficult since no cytotoxicity curves were presented. For Apoptin-hC-SOD3, its IC_{50} values are similar to those of H_6-ApopΔProΔLeu (IC_{50} around 4 μM on HeLa cells). Fusion of TAT-Apoptin to GST to improve the production of engineered Apoptin in *E. coli* produced a variant that at 2 μM (90 μg/ml) induced apoptosis in 12% of HL-60 treated cells (Lee et al., 2012). More efficient protein delivery methods have been recently developed. They include the use of degradable polymeric nanocapsules (Zhao et al., 2013) that have an IC_{50} around 30-100 nM or the transfection with a vector that allows Apoptin to be secreted and reach

surrounding cancer cells through a bystanding effect (Flinterman et al., 2009; Ma et al., 2012).

Here we present an alternative delivery strategy of Apoptin that uses a soluble, non-aggregated variant to generate a cytotoxic response specifically in tumor cells.

In vitro, exogenously added H_6-ApopΔProΔLeu is cytotoxic for tumor cells (Figure 29) and may be an interesting candidate to be evaluated as an antitumor drug. The use of a known concentration of protein, a parameter that is difficult to estimate when transfecting the Apoptin gene, has allowed us to quantify the cytotoxic action of H_6-ApopΔProΔLeu on different cell lines. The IC_{50} values of H_6-ApopΔProΔLeu range from 7 to 25 μM, depending on the assayed tumor cell line. At this stage, it is difficult to discern whether the mechanism of exogenously added H_6-ApopΔProΔLeu is similar to that of the transfected gene. It has been shown that in Jurkat cells, which lack functional caspase -8, Apoptin-induced apoptosis is independent of death receptors but triggers caspase-9 activation (Maddika et al., 2005) and hence, acts through the intrinsic pathway. H_6-ApopΔProΔLeu when exogenously added to NCI-H460 cells not only activates the effector procaspase-3 but also the initiators procaspase-8 and -9, acting via death receptor pathway and the apoptosome pathway as well. As discussed by Bullenkamp and Tavassoli (Bullenkamp and Tavassoli, 2014), Apoptin activation of the extrinsic apoptotic pathway has not been ruled out since Apoptin binds to different partners involved in death receptor signaling like FADD (Guelen et al., 2004) or DEDAF (Huo et al., 2008).

Apoptin possesses a net positive charge at low and neutral pH and its isoelectric point is estimated to be 10.3 on the basis of its amino acid sequence. Resembling other cationic proteins that internalize after being exogenously added to tumor cells, i.e. different pancreatic ribonucleases

(for a review see (Benito et al., 2008)) it seems plausible that Apoptin could be able to cross the lipid bilayer to exert its biological action inside the cell. H_6-ApopΔProΔLeu does not multimerise and it is soluble in all the assayed conditions. This seems to indicate that its mechanism of cytotoxicity does not require the aggregation as it has been reported for Apoptin. In fact, it has been described that even though both N- and C-terminal regions of Apoptin have inherent cell-killing activity (Danen-Van Oorschot et al., 2003), the C-terminus is stronger at inducing apoptosis and it could be hypothesized that only the cytotoxicity of the N-terminus would be dependent on protein aggregation.

The solubility and small size of H_6-ApopΔProΔLeu in solution have greatly facilitated the characterization of its conformational ensemble and dynamics by different biophysical techniques reported in the results section. We have found that the H_6-ApopΔProΔLeu is intrinsically disordered with no significant propensity to adopt preferred secondary structural elements and there are no stable intramolecular backbone hydrogen bonds. The dynamics studied by ^{1}HN relaxation experiments show that the backbone is significantly more flexible than a typical globular protein, although it is not as flexible as short, structureless peptide.

Regarding wild type Apoptin, a previous study advanced a structural model for the full length protein in which it adopts a well ordered globular structure (Panigrahi et al., 2012). The data reported here do not provide support for that model. On the basis of our data, which were mainly obtained with the H_6-ApopΔProΔLeu variant, we cannot rule out that the deleted segments might be crucial for folding full length Apoptin into a partially folded or well-ordered globular protein upon binding to other proteins. Nevertheless, a previous low resolution CD study which showed that full length Apoptin is poor in secondary structure (Leliveld et al.,

2003a) is consistent with our findings. We have also shown that ANS binding to H_6-Apoptin requires the presence of the hydrophobic LRS that in a folded protein would be expected to be buried inside the structure. Once again, this shows that Apoptin lacks globular structure and that it is indeed an IDP that does not form a stable hydrophobic core and probably is mostly solvent exposed (Dunker et al., 2001; Uversky et al., 2000). The fact that H_6-ApopΔProΔLeu maintains most of the biological properties of Apoptin seems to corroborate this conclusion. The differences between our results and the *in silico* structure of Apoptin reported by Panigrahi *et al.* highlights the danger of utilizing folded protein structures to model the conformation of an intrinsically disordered protein.

The IDP character of Apoptin aids our comprehension of its mode of action. The ability of fragments lacking half of the Apoptin's sequence to retain part of the full length protein's cytotoxicity (Danen-Van Oorschot et al., 2003) would be perplexing if Apoptin had a well-defined 3D structure but is easily understood considering that Apoptin is intrinsically disordered. It may seem surprising that its action is highly specific to tumor cells and that it interacts with several molecules to induce apoptosis (Bullenkamp and Tavassoli, 2014). However, since IDPs are uniquely well suited to bind quickly with moderate specificity to multiple partners (Berlow et al., 2015; Pontius, 1993), it is logical that Apoptin adopts an ensemble of unfolded conformations. Indeed, many regulatory proteins involved in gene expression and cell cycle control are IDPs (Pontius, 1993).

It can be envisaged that the fact that H_6-ApopΔProΔLeu is an IDP enhances its ability to cross the lipid bilayer when it is exogenously added to the cells. In this case, the IDP character could have facilitated the cytotoxic activity of this variant. In fact, one strategy to successfully transduce cells

is through denatured TAT-fusion proteins, which are believe to later refold in the cell (Nagahara et al., 1998).

Analysis of the 2D and 3D spectra of H_6-ApopΔProΔLeu reveal that Cys25 and Cys66 are oxidized, whereas Cys23 is unassigned. The appearance of Cys23 following reduction with TCEP suggests that it was oxidized prior to treatment with TCEP. It has been described that Cys90 (Cys66 in this variant) is the only one reduced in wild type Apoptin (Leliveld et al., 2003c). Although we cannot directly ascertain this variant's disulfide bond pairings, it is tempting to speculate that Cys23 and Cys25 would be linked by an intramolecular disulfide bond whereas Cys66 would form a disulfide bond with a GSH molecule, which is used along the purification process. In any case, the redox state of Cys residues does not affect the overall conformational ensemble of H_6-ApopΔProΔLeu since TOCSY and NOESY spectra are very similar under oxidized and reduced conditions (Figure 39); the resonances generally retained their chemical shift values in the presence of TCEP but become narrower. This reduced line broadening in the presence of the reducing agent suggests that the Cys oxidization may contribute to distinct interconverting conformations. Overall, however, the protein construct behaves as an intrinsically disordered chain under both reducing and oxidizing conditions. It is known that the redox status changes among the subcellular compartments into a cell and among the different metabolic states of the cells (Jorgenson et al., 2013). The formation of these disulfide bonds could be promoted by the particular redox conditions which could lead to conformational changes in Apoptin and consequently different partner recognition. The recent observations that post-transcriptional modifications facilitate regulation of IDP functions (Babu et al., 2011) is relevant to this point.

It was previously reported that *in vitro* Cyclin A-associated CDK2 specifically phosphorylates Thr108 (Maddika et al., 2009). We took

advantage of this to specifically compare the conformation of phosphorylated and non-phosphorylated Apoptin. We show that Thr phosphorylation does not induce any significant conformational change in H_6-ApopΔProΔLeu. As Apoptin is an IDP, phosphorylation could promote or hamper binding to specific partners inside cells. Rather than inducing folding, Thr phosphorylation could mask the nuclear export sequence leading to nuclear accumulation along the lines proposed for Thr108 (Thr84 in H_6-ApopΔProΔLeu) by Rohn and coworkers (Rohn et al., 2002). In fact, we show that this variant accumulates significantly in the nuclei of cancer cells (Figure 25).

The monomeric/small oligomeric nature of this variant not only allows it to be cytotoxic when added exogenously but also may facilitate its diffusion through the nuclear pores once it has been internalized into the cells. Accordingly, in the transfection experiments with GFP-ApopΔPro the oligomeric behavior of this variant may have retarded its nuclear accumulation since it was initially found in the cytoplasm appearing as a filamentous pattern surrounding the nuclei, prior internalization and apoptosis induction (Figure 28). This effect has also been noticed by other authors for wild type Apoptin (Guelen et al., 2004; Leliveld et al., 2003a). Since Apoptin interacts with Bcl10 and FADD, it has been suggested that these fibers are the cytoplasmic structures called death effector filaments (DEFs) formed in Fas-mediated apoptosis (Guelen et al., 2004). Another study showed the interaction of Apoptin with HSP-70, which associates with α-tubulin, β-tubulin and β-actin, suggesting an association with filamentous networks (Teodoro et al., 2004). None of these interactions have been found in any of the cell lines transfected with GFP-H_6-ApopΔProΔLeu.

The results obtained in this thesis provide an alternative delivery strategy for Apoptin, a clear candidate to be used as antitumoral drug, either alone

or in combination with other drugs in order to overcome intrinsic tumor cell resistance to apoptosis. Many *in vitro* and *in vivo* studies performed by world-wide researchers using different delivery systems have proven Apoptin to be safe and efficient and it seems ready to go towards clinical trials (Backendorf and Noteborn, 2014). However, despite the relative success of Apoptin in preclinical research, its mechanism of action has remained poorly understood. Our findings advance the knowledge about its conformation and mechanism of action and the chemical shift values reported here could be valuable tools for studying the interaction of Apoptin with DNA and numerous protein partners which is believed to be key for selectively triggering apoptosis in cancer cells (Bullenkamp and Tavassoli, 2014).

Conclusions

1. A purification protocol for an Apoptin variant fused to an N-terminal His-tag named H_6-Apoptin has been set up and the protein has been successfully produced and purified from *E. coli* BL21 (DE3) with a yield of approximately 20 mg per L of induced culture.

2. H_6-Apoptin forms water-soluble multimers composed of about 50 subunits. The presence of an amyloid-like structure within the aggregates in the conditions assayed cannot be absolutely discarded and other conditions and reagents must be examined to completely corroborate this.

3. Two N-terminal truncated H_6-Apoptin variants lacking the PRS, residues 1-27 (H_6-ApopΔPro) and both the PRS and LRS, residues 1-46 (H_6-ApopΔProΔLeu) have been successfully produced and purified following the protocol established for H_6-Apoptin. The final yield for H_6-ApopΔPro is approximately 20 mg per L of induced culture while for H_6-ApopΔProΔLeu is roughly 10 mg per L of induced culture.

4. H_6-ApopΔPro forms water-soluble multimers of around 50 subunits as is the case of H_6-Apoptin while H_6-ApopΔProΔLeu is purified mainly as a monomer/small oligomer (LMM form) although a minor form of twice its molecular weight (HMM form) is also produced. These results indicate that the removal of LRS stretch is necessary to reduce the tendency of Apoptin to aggregate.

5. The HMM form of H_6-ApopΔProΔLeu is stabilized by disulfide bonds although other non-covalent interactions such as hydrophobic interactions could also be involved.

6. All three proteins are able to bind dsDNA, although H_6-ApopΔProΔLeu likely binds it with less strength, suggesting that the LRS may be involved in DNA binding.

7. The large spherical superstructures formed by Apoptin, which have been shown to strongly interact with DNA, are not necessary to specifically kill tumor cells.

8. GFP-ApopΔProΔLeu has a decreased capacity to accumulate into the nucleus compared to GFP-Apoptin and GFP-ApopΔPro but it maintains the differential subcellular localization of the wild type protein in tumor and normal cells. Hence the LRS is implicated in the nuclear accumulation of Apoptin in tumor cells.

9. Transfected GFP-Apoptin and GFP-ApopΔProΔLeu kill NCI-H460 tumor cells, being the effect of GFP-Apoptin 1.9 times higher than that of GFP-ApopΔProΔLeu. NCI-H460 cells transfected with GFP-ApopΔProΔLeu show typical morphological hallmarks of apoptosis such as chromatin condensation, nuclear fragmentation and membrane blebbing, as observed for cells treated with GFP-Apoptin.

10. We have developed an alternative strategy for *in vitro* delivery of Apoptin by reducing its tendency to aggregation. The IC_{50} values of H_6-ApopΔProΔLeu, when externally added, range from 7 to 25 μM, depending on the assayed tumor cell line. In contrast, H_6-ApopΔProΔLeu does not produce any effect on any of the non-tumor cell lines assayed. This indicates that aggregation is neither crucial for Apoptin's cytotoxicity nor for its selectivity.

11. When H_6-ApopΔProΔLeu is exogenously added, it produces the translocation of phosphatidiylserine to the outer face of the cells and induces the activation of procaspases-8, -9 and -3, inducing the apoptosis of NCI-H460 tumor cells. The procaspase activation reaches its maximum at 36 h after exogenous treatment with H_6-ApopΔProΔLeu.

12. The dynamics of H_6-ApopΔProΔLeu obtained by ^{1}HN relaxation experiments show that the backbone is significantly more flexible than a typical globular protein. Although it is not as flexible as short, structureless peptide, this protein has no significant propensity to adopt preferred secondary structure and it would be classified within the Intrinsically Disordered Proteins (IDPs) group. Nevertheless, we cannot rule out that the deleted segments might be crucial for folding of full length Apoptin into partially folded or well-ordered globular protein.

13. The similarity of the experimental chemical shifts values of H_6-ApopΔProΔLeu with those predicted for its random coil ensemble and for IDPs, the very low H/D protection factors, the lack of sigmoidal unfolding transition upon temperature increase and the sensitivity to proteinase K digestion suggest that this protein does not have a preferred ensemble or an intramolecular H-bonding network and therefore, it lacks an stable secondary structure.

14. The IDP character of H_6-ApopΔProΔLeu opens new alternatives to explain the mechanism of cytotoxicity of Apoptin since this character would explain Apoptin's ability to bind multiple targets to specifically induce apoptosis in a wide range of tumor cells.

15. The redox state of Cys residues does not affect the overall conformational ensemble of H_6-ApopΔProΔLeu.

16. Thr phosphorylation does not induce any significant conformational change in H_6-ApopΔProΔLeu.

References

Afghahi, A., and Sledge, G.W. (2015). Targeted Therapy for Cancer in the Genomic Era. Cancer J. *21*, 294–298.

Ahmed, M., Davis, J., Aucoin, D., Sato, T., Ahuja, S., Aimoto, S., Elliott, J.I., Van Nostrand, W.E., and Smith, S.O. (2010). Structural conversion of neurotoxic amyloid-β1–42 oligomers to fibrils. Nat. Struct. Mol. Biol. *17*, 561–567.

Allison, R., Johnston, R.E., and Dougherty, W.G. (1986). The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: evidence for the synthesis of a single polyprotein. Virology *154*, 9–20.

An, S., Nam, K., Choi, S., Bai, C.Z., Lee, Y., and Park, J.-S. (2013). Nonviral gene therapy *in vivo* with PAM-RG4/apoptin as a potential brain tumor therapeutic. Int. J. Nanomedicine *8*, 821–834.

Argiris, K., Panethymitaki, C., and Tavassoli, M. (2011). Naturally occurring, tumor-specific, therapeutic proteins. Exp. Biol. Med. (Maywood). *236*, 524–536.

Babu, M.M., van der Lee, R., de Groot, N.S., and Gsponer, J. (2011). Intrinsically disordered proteins: regulation and disease. Curr. Opin. Struct. Biol. *21*, 432–440.

Backendorf, C., and Noteborn, M.H.M. (2014). Apoptin towards safe and efficient anticancer therapies. Adv. Exp. Med. Biol. *818*, 39–59.

Backendorf, C., Visser, A.E., de Boer, A.G., Zimmerman, R., Visser, M., Voskamp, P., Zhang, Y.-H., and Noteborn, M. (2008). Apoptin: therapeutic potential of an early sensor of carcinogenic transformation. Annu. Rev. Pharmacol. Toxicol. *48*, 143–169.

Bai, Y., Milne, J.S., Mayne, L., and Englander, S.W. (1993). Primary structure effects on peptide group hydrogen exchange. Proteins Struct. Funct. Genet. *17*, 75–86.

Bao, J., and Zervos, A.S. (1996). Isolation and characterization of Nmi, a novel partner of Myc proteins. Oncogene *12*, 2171–2176.

Benito, A., Vilanova, M., and Ribó, M. (2008). Intracellular routing of cytotoxic pancreatic-type ribonucleases. Curr. Pharm. Biotechnol. *9*, 169–179.

Berlow, R.B., Dyson, H.J., and Wright, P.E. (2015). Functional advantages of dynamic protein disorder. FEBS Lett. *589*, 2433–2440.

Bertani, G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli.* J. Bacteriol. *62*, 293–300.

Bieging, K.T., and Attardi, L.D. (2012). Deconstructing p53 transcriptional networks in tumor suppression. Trends Cell Biol. *22*, 97–106.

Bienkiewicz, E.A., and Lumb, K.J. (1999). Random-coil chemical shifts of phosphorylated amino acids. J. Biomol. NMR *15*, 203–206.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. *72*, 248–254.

Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C.M., and Stefani, M. (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. Nature *416*, 507–511.

Bullenkamp, J., and Tavassoli, M. (2014). Signalling of Apoptin. Adv. Exp. Med. Biol. *818*, 11–37.

Bullenkamp, J., Cole, D., Malik, F., Alkhatabi, H., Kulasekararaj, A., Odell, E.W., Farzaneh, F., Gäken, J., and Tavassoli, M. (2012). Human Gyrovirus Apoptin shows a similar subcellular distribution pattern and apoptosis induction as the chicken anaemia virus derived VP3/Apoptin. Cell Death Dis. *3*, e296.

Bullenkamp, J., Gäken, J., Festy, F., Chong, E.Z., Ng, T., and Tavassoli, M. (2015). Apoptin interacts with and regulates the activity of protein kinase C beta in cancer cells. Apoptosis *20*, 831–842.

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

Cao, H.-D., Yang, Y.-X., Lü, L., Liu, S.-N., Wang, P.-L., Tao, X.-H., Wang, L.-J., and Xiang, T.-X. (2010). Attenuated *Salmonella typhimurium* carrying TRAIL and VP3 genes inhibits the growth of gastric cancer cells *in vitro* and *in vivo*. Tumori *96*, 296–303.

Cardamone, M., and Puri, N.K. (1992). Spectrofluorimetric assessment of the surface hydrophobicity of proteins. Biochem. J. *282 (Pt 2*, 589–593.

Castro, J., Ribó, M., Navarro, S., Nogués, M.V., Vilanova, M., and Benito, A. (2011). A human ribonuclease induces apoptosis associated with p21WAF1/CIP1 induction and JNK inactivation. BMC Cancer *11*, 9.

Castro, J., Ribó, M., Puig, T., Colomer, R., Vilanova, M., and Benito, A. (2012). A cytotoxic ribonuclease reduces the expression level of P-glycoprotein in multidrug-resistant cell lines. Invest. New Drugs *30*, 880–888.

Chari, R.V.J., Miller, M.L., and Widdison, W.C. (2014). Antibody-drug conjugates: an emerging concept in cancer therapy. Angew. Chem. Int. Ed. Engl. *53*, 3796–3827.

Chen, K., Luo, Z., Tang, J., and Zheng, S.J. (2011a). A critical role of heat shock cognate protein 70 in Apoptin-induced phosphorylation of Akt. Biochem. Biophys. Res. Commun. *409*, 200–204.

Chen, K., Luo, Z., and Zheng, S.J. (2011b). Gallus Heat shock cognate protein 70, a novel binding partner of Apoptin. Virol. J. *8*, 324.

Cheng, C.-M., Huang, S., Chang, Y.-F., Chung, W.-Y., and Yuo, C.-Y. (2003). The viral death protein Apoptin interacts with Hippi, the protein interactor of Huntingtin-interacting protein 1. Biochem. Biophys. Res. Commun. *305*, 359–364.

Chesshyre, J.A., and Hipkiss, A.R. (1989). Low temperatures stabilize interferon α-2 against proteolysis in Methylophilus methylotrophus and *Escherichia coli*. Appl. Microbiol. Biotechnol. *31*, 158–162.

Chu, D.K.W., Poon, L.L.M., Chiu, S.S.S., Chan, K.H., Ng, E.M., Bauer, I., Cheung, T.K., Ng, I.H.Y., Guan, Y., Wang, D., et al. (2012). Characterization of a novel gyrovirus in human stool and chicken meat. J. Clin. Virol. *55*, 209–213.

Costa, S., Almeida, A., Castro, A., and Domingues, L. (2014). Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. Front. Microbiol. *5*, 63.

Danen-Van Oorschot, A. A., Fischer, D.F., Grimbergen, J.M., Klein, B., Zhuang, S., Falkenburg, J.H., Backendorf, C., Quax, P.H., Van der Eb, a J., and Noteborn, M.H. (1997). Apoptin induces apoptosis in human transformed and malignant cells but not in normal cells. Proc. Natl. Acad. Sci. U. S. A. *94*, 5843–5847.

Danen-van Oorschot, A. A. A. M., Voskamp, P., Seelen, M.C.M.J., van Miltenburg, M.H. a M., Bolk, M.W., Tait, S.W., Boesen-de Cock, J.G.R., Rohn, J.L., Borst, J., and Noteborn, M.H.M. (2004). Human death effector domain-associated factor interacts with the viral apoptosis agonist Apoptin and exerts tumor-preferential cell killing. Cell Death Differ. *11*, 564–573.

Danen-Van Oorschot, A.A., van der Eb, A.J., and Noteborn, M.H. (1999a). BCL-2 stimulates Apoptin-induced apoptosis. Adv. Exp. Med. Biol. *457*, 245–249.

Danen-Van Oorschot, A.A., Zhang, Y., Erkeland, S.J., Fischer, D.F., van der Eb, A.J., and Noteborn, M.H. (1999b). The effect of Bcl-2 on Apoptin in "normal" vs transformed human cells. Leukemia *13 Suppl 1*, S75–S77.

Danen-Van Oorschot, A.A.A.M., Zhang, Y.-H.H., Leliveld, S.R., Rohn, J.L., Seelen, M.C.M.J., Bolk, M.W., Van Zon, A., Erkeland, S.J., Abrahams, J.-P.P., Mumberg, D., et al. (2003). Importance of nuclear localization of apoptin for tumor-specific induction of apoptosis. J. Biol. Chem. *278*, 27729–27736.

Deininger, M., Buchdunger, E., and Druker, B.J. (2005). The development of imatinib as a therapeutic agent for chronic myeloid leukemia. Blood *105*, 2640–2653.

Dosztányi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics *21*, 3433–3434.

Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.M., Hipps, K.W., et al. (2001). Intrinsically disordered protein. J. Mol. Graph. Model. *19*, 26–59.

Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D., and Kay, L.E. (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. Biochemistry *33*, 5984–6003.

Fiser, A. (2010). Template-based protein structure modeling. Methods Mol. Biol. *673*, 73–94.

Flinterman, M., Farzaneh, F., Habib, N., Malik, F., Gäken, J., and Tavassoli, M. (2009). Delivery of therapeutic proteins as secretable TAT fusion products. Mol. Ther. *17*, 334–342.

Font, J., Benito, A., Lange, R., Ribó, M., and Vilanova, M. (2006). The contribution of the residues from the main hydrophobic core of ribonuclease A to its pressure-folding transition state. Protein Sci. *15*, 1000–1009.

Friedeberg, C., Scarlett, G., McGeehan, J., Abu-Daya, A., Guille, M., and Kneale, G. (2006). Identification of a structural and functional domain in xNAP1 involved in protein-protein interactions. Nucleic Acids Res. *34*, 4893–4899.

Gaffney, P.J., and Edgell, T.A. (1995). The International and "NIH" units for thrombin--how do they compare? Thromb. Haemost. *74*, 900–903.

Goldman, L.A., Cutrone, E.C., Kotenko, S. V., Krause, C.D., and Langer, J.A. (1996). Modifications of vectors pEF-BOS, pcDNA1 and pcDNA3 result in improved convenience and expression. Biotechniques *21*, 1013–1015.

Gras, S.L., Waddington, L.J., and Goldie, K.N. (2011). Transmission electron microscopy of amyloid fibrils. Methods Mol. Biol. *752*, 197–214.

Gualfetti, P.J., Iwakura, M., Lee, J.C., Kihara, H., Bilsel, O., Zitzewitz, J.A., and Matthews, C.R. (1999). Apparent radii of the native, stable intermediates and unfolded conformers of the alpha-subunit of tryptophan synthase from *E. coli*, a TIM barrel protein. Biochemistry *38*, 13367–13378.

Guan, G., Zhao, M., Liu, L., Jin, C., Sun, K., Zhang, D., Yu, D., Cao, H., Lu, Y., and Wen, L. (2013). *Salmonella typhimurium* mediated delivery of Apoptin in human laryngeal cancer. Int. J. Med. Sci. *10*, 1639–1648.

Guelen, L., Paterson, H., Gäken, J., Meyers, M., Farzaneh, F., and Tavassoli, M. (2004). TAT-apoptin is efficiently delivered and induces apoptosis in cancer cells. Oncogene *23*, 1153–1165.

Gupta, S.K., Gandham, R.K., Sahoo, a. P., and Tiwari, a. K. (2014). Viral genes as oncolytic agents for cancer therapy. Cell. Mol. Life Sci. *72*, 1073–1094.

Han, S.-X., Zhu, Q., Ma, J.-L., Lv, Y., Zhao, J., Huang, C., Jia, X., Ou, W., and Guo, H.-T. (2010). Apoptin sensitizes radiation-induced cell death via classic mitochondrial, caspase and p53-dependent signaling in HepG2 cells.

Heckl, S., Regenbogen, M., Sturzu, A., Gharabaghi, A., Feil, G., Beck, A., Echner, H., and Nagele, T. (2008). Value of apoptin's 40-amino-acid C-terminal fragment for the differentiation between human tumor and non-tumor cells. Apoptosis *13*, 495–508.

Heilman, D.W., Teodoro, J.G., and Green, M.R. (2006). Apoptin nucleocytoplasmic shuttling is required for cell type-specific localization, apoptosis, and recruitment of the anaphase-promoting complex/cyclosome to PML bodies. J. Virol. *80*, 7535–7545.

Herrmann, F., and Biology, A.M. (1995). Cancer gene therapy : principles , problems , and perspectives. 157–163.

Hirel, P.H., Schmitter, M.J., Dessen, P., Fayat, G., and Blanquet, S. (1989). Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. Proc. Natl. Acad. Sci. U. S. A. *86*, 8247–8251.

Ho, A., Schwarze, S.R., Mermelstein, S.J., Waksman, G., and Dowdy, S.F. (2001). Synthetic protein transduction domains: enhanced transduction potential *in vitro* and *in vivo*. Cancer Res. *61*, 474–477.

Huang, F., Oldfield, C., Meng, J., Hsu, W.-L., Xue, B., Uversky, V.N., Romero, P., and Dunker, A. K. (2012). Subclassifying disordered proteins by the CH-CDF plot method. Pacific Symp. Biocomput. 2012 128–139.

Huo, D.-H., Yi, L.-N., and Yang, J. (2008). Interaction with Ppil3 leads to the cytoplasmic localization of Apoptin in tumor cells. Biochem. Biophys. Res. Commun. *372*, 14–18.

Imai, K., and Takaoka, A. (2006). Comparing antibody and small-molecule therapies for cancer. Nat. Rev. Cancer *6*, 714–727.

Jangamreddy, J.R., Panigrahi, S., and Lotfi, K. Mapping of apoptin-interaction with BCR-ABL1, and development of apoptin-based targeted therapy . Oncotarget *5*, 7198-7211.

Janssen, K., Hofmann, T.G., Jans, D.A., Hay, R.T., Schulze-Osthoff, K., and Fischer, U. (2007). Apoptin is modified by SUMO conjugation and targeted to promyelocytic leukemia protein nuclear bodies. Oncogene *26*, 1557–1566.

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

Jiang, J., Cole, D., Westwood, N., Macpherson, L., Farzaneh, F., Mufti, G., Tavassoli, M., and Gäken, J. (2010). Crucial roles for protein kinase C isoforms in tumor-specific killing by apoptin. Cancer Res. *70*, 7242–7252.

Jin, J.L., Gong, J., Yin, T.J., Lu, Y.J., Xia, J.J., Xie, Y.Y., Di, Y., He, L., Guo, J.L., Sun, J., et al. (2011). PTD4-apoptin protein and dacarbazine show a synergistic antitumor effect on B16-F1 melanoma *in vitro* and *in vivo*. Eur. J. Pharmacol. *654*, 17–25.

Johnson, B.A. (2004). Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol. Biol. *278*, 313–352.

Jorgenson, T.C., Zhong, W., and Oberley, T.D. (2013). Redox imbalance and biochemical changes in cancer. Cancer Res. *73*, 6118–6123.

Kayed, R., Head, E., Sarsoza, F., Saing, T., Cotman, C.W., Necula, M., Margol, L., Wu, J., Breydo, L., Thompson, J.L., et al. (2007). Fibril specific, conformation dependent antibodies recognize a generic epitope common to amyloid fibrils and fibrillar oligomers that is absent in prefibrillar oligomers. Mol. Neurodegener. *2*, 18.

Khurana, R., Uversky, V.N., Nielsen, L., and Fink, A.L. (2001). Is Congo Red an amyloid-specific dye? J. Biol. Chem. *276*, 22715–22721.

Khurana, R., Coleman, C., Ionescu-Zanetti, C., Carter, S. a., Krishna, V., Grover, R.K., Roy, R., and Singh, S. (2005). Mechanism of thioflavin T binding to amyloid fibrils. J. Struct. Biol. *151*, 229–238.

Kim, Y.-S., Randolph, T.W., Manning, M.C., Stevens, F.J., and Carpenter, J.F. (2003). Congo Red populates partially unfolded states of an amyloidogenic protein to enhance aggregation and amyloid fibril formation. J. Biol. Chem. *278*, 10842–10850.

Kingston, R.E., Chen, C.A., and Rose, J.K. (2003). Calcium phosphate transfection. Curr. Protoc. Mol. Biol. *Chapter 9*, Unit 9.1.

Kirn, D., Martuza, R.L., and Zwiebel, J. (2001). Replication-selective virotherapy for cancer: Biological principles, risk management and future directions. Nat. Med. *7*, 781–787.

Kjaergaard, M., Brander, S., and Poulsen, F.M. (2011). Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. J. Biomol. NMR *49*, 139–149.

Klunk, W.E., Pettegrew, J.W., and Abraham, D.J. (1989). Quantitative evaluation of Congo Red binding to amyloid-like proteins with a beta-pleated sheet conformation. J. Histochem. Cytochem. *37*, 1273–1281.

Komatsu, H., Shinotani, N., Kimori, Y., Tokuoka, J.I., Kaseda, K., Nakagawa, H., and Kodama, T. (2006). Aggregation of partially unfolded myosin subfragment-1 into spherical oligomers with amyloid-like dye-binding properties. J. Biochem. *139*, 989–996.

Kucharski, T.J., Gamache, I., Gjoerup, O., and Teodoro, J.G. (2011). DNA damage response signaling triggers nuclear localization of the chicken anemia virus protein Apoptin. J. Virol. *85*, 12638–12649.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. *157*, 105–132.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680–685.

Lanz, H.L., Florea, B.I., Noteborn, M.H.M., and Backendorf, C. (2012). Development and application of an *in vitro* apoptin kinase assay. Anal. Biochem. *421*, 68–74.

Lee, M.-S., Sun, F.-C., Huang, C.-H., Lien, Y.-Y., Feng, S.-H., Lai, G.-H., Lee, M.-S., Chao, J., Chen, H.-J., Tzen, J.T.C., et al. (2012). Efficient production of an engineered apoptin from chicken anemia virus in a recombinant *E. coli* for tumor therapeutic applications. BMC Biotechnol. *12*, 27.

Lee, Y.-H., Cheng, C.-M., Chang, Y.-F., Wang, T.-Y., and Yuo, C.-Y. (2007). Apoptin T108 phosphorylation is not required for its tumor-specific nuclear localization but partially affects its apoptotic activity. Biochem. Biophys. Res. Commun. *354*, 391–395.

Leliveld, S.R., Zhang, Y.-H., Rohn, J.L., Noteborn, M.H.M., and Abrahams, J.P. (2003a). Apoptin induces tumor-specific apoptosis as a globular multimer. J. Biol. Chem. *278*, 9042–9051.

Leliveld, S.R., Dame, R.T., Mommaas, M.A., Koerten, H.K., Wyman, C., Danen-van Oorschot, A.A., Rohn, J.L., Noteborn, M.H.M., and Abrahams, J.P. (2003b). Apoptin protein multimers form distinct higher-order nucleoprotein complexes with DNA. Nucleic Acids Res. *31*, 4805–4813.

Leliveld, S.R., Noteborn, M.H.M., and Abrahams, J.P. (2003c). Prevalent conformations and subunit exchange in the biologically active apoptin protein multimer. Eur. J. Biochem. *270*, 3619–3627.

Leliveld, S.R., Dame, R.T., Rohn, J.L., Noteborn, M.H.M., and Abrahams, J.P. (2004). Apoptin's functional N- and C-termini independently bind DNA. FEBS Lett. *557*, 155–158.

Li, X., Liu, Y., Wen, Z., Li, C., Lu, H., Tian, M., Jin, K., Sun, L., Gao, P., Yang, E., et al. (2010). Potent anti-tumor effects of a dual specific oncolytic adenovirus expressing apoptin *in vitro* and *in vivo*. Mol. Cancer *9*, 10.

Lindgren, M., Sörgjerd, K., and Hammarström, P. (2005). Detection and characterization of aggregates, prefibrillar amyloidogenic oligomers, and protofibrils using fluorescence spectroscopy. Biophys. J. *88*, 4200–4212.

Liu, L., Wu, W., Zhu, G., Liu, L., Guan, G., Li, X., Jin, N., and Chi, B. (2012). Therapeutic efficacy of an hTERT promoter-driven oncolytic adenovirus that expresses apoptin in gastric carcinoma. Int. J. Mol. Med. *30*, 747–754.

Liu, Q., Fu, H., Xing, R., Tie, Y., Zhu, J., Sun, Z., and Zheng, X. (2008). Survivin knockdown combined with apoptin overexpression inhibits cell growth significantly. Cancer Biol. Ther. *7*, 1053–1060.

Liu, X., Elojeimy, S., El-Zawahry, A.M., Holman, D.H., Bielawska, A., Bielawski, J., Rubinchik, S., Guo, G.-W., Dong, J.-Y., Keane, T., et al. (2006). Modulation of ceramide metabolism enhances viral protein apoptin's cytotoxicity in prostate cancer. Mol. Ther. *14*, 637–646.

Los, M., Panigrahi, S., Rashedi, I., Mandal, S., Stetefeld, J., Essmann, F., and Schulze-Osthoff, K. (2009). Apoptin, a tumor-selective killer. Biochim. Biophys. Acta *1793*, 1335–1342.

Ma, J.-L., Han, S.-X., Zhao, J., Zhang, D., Wang, L., Li, Y.-D., and Zhu, Q. (2012). Systemic delivery of lentivirus-mediated secretable TAT-apoptin eradicates hepatocellular carcinoma xenografts in nude mice. Int. J. Oncol. *41*, 1013–1020.

Maddika, S., Booy, E.P., Johar, D., Gibson, S.B., Ghavami, S., and Los, M.

(2005). Cancer-specific toxicity of apoptin is independent of death receptors but involves the loss of mitochondrial membrane potential and the release of mitochondrial cell-death mediators by a Nur77-dependent pathway. J Cell Sci. *118*, 4485-4493.

Maddika, S., Mendoza, F.J., Hauff, K., Zamzow, C.R., Paranjothy, T., and Los, M. (2006). Cancer-selective therapy of the future: Apoptin and its mechanism of action. Cancer Biol. Ther. *5*, 10–19.

Maddika, S., Bay, G.H., Kroczak, T.J., Ande, S.R., Wiechec, E., Gibson, S.B., and Los, M. (2007). Akt is transferred to the nucleus of cells treated with apoptin, and it participates in apoptin-induced cell death. Cell Prolif. *40*, 835–848.

Maddika, S., Wiechec, E., Ande, S.R., Poon, I.K., Fischer, U., Wesselborg, S., Jans, D. a, Schulze-Osthoff, K., and Los, M. (2008a). Interaction with PI3-kinase contributes to the cytotoxic activity of apoptin. Oncogene *27*, 3060–3065.

Maddika, S., Ande, S.R., Wiechec, E., Hansen, L.L., Wesselborg, S., and Los, M. (2008b). Akt-mediated phosphorylation of CDK2 regulates its dual role in cell cycle progression and apoptosis. J. Cell Sci. *121*, 979–988.

Maddika, S., Panigrahi, S., Wiechec, E., Wesselborg, S., Fischer, U., Schulze-Osthoff, K., and Los, M. (2009). Unscheduled Akt-triggered activation of cyclin-dependent kinase 2 as a key effector mechanism of apoptin's anticancer toxicity. Mol. Cell. Biol. *29*, 1235–1248.

Markley, J.L., Bax, A., Arata, Y., Hilbers, C.. W., Kaptein, R., Sykes, B.D., Wright, P.E., and Wüthrich, K. (1998). Recommendations for the presentation of NMR structures of proteins and nucleic acids. J. Mol. Biol. *280*, 933–952.

Marley, J., Lu, M., and Bracken, C. (2001). A method for efficient isotopic labeling of recombinant proteins. J. Biomol. NMR *20*, 71–75.

McLellan, T. (1982). Electrophoresis buffers for polyacrylamide gels at various pH. Anal. Biochem. *126*, 94–99.

Mitrus, I., Missol-Kolka, E., Plucienniczak, A., and Szala, S. (2005). Tumour therapy with genes encoding apoptin and E4orf4. Anticancer Res. *25*, 1087–1090.

Nagahara, H., Vocero-Akbani, A.M., Snyder, E.L., Ho, A., Latham, D.G., Lissy, N.A., Becker-Hapak, M., Ezhevsky, S.A., and Dowdy, S.F. (1998). Transduction of full-length TAT fusion proteins into mammalian cells: TAT-p27Kip1 induces cell migration. Nat. Med. *4*, 1449–1452.

Naiki, H., and Gejyo, F. (1999). Kinetic analysis of amyloid fibril formation. Methods Enzymol. *309*, 305–318.

Nemunaitis, J. (1999). Oncolytic viruses. Invest. New Drugs *17*, 375–386.

Nilsson, M.R. (2004). Techniques to study amyloid fibril formation *in vitro*. Methods *34*, 151–160.

Noteborn, M.H.M. (2004). Chicken anemia virus induced apoptosis: underlying molecular mechanisms. Vet. Microbiol. *98*, 89–94.

Noteborn, M.H., and van der Eb, A. J. (1998). Apoptin-induced apoptosis: potential for antitumor therapy. Drug Resist. Updat. *1*, 99–103.

Noteborn, M.H., and Koch, G. (1995). Chicken anaemia virus infection: molecular basis of pathogenicity. Avian Pathol. *24*, 11–31.

Noteborn, M.H., de Boer, G.F., van Roozelaar, D.J., Karreman, C., Kranenburg, O., Vos, J.G., Jeurissen, S.H., Hoeben, R.C., Zantema, A., and Koch, G. (1991). Characterization of cloned chicken anemia virus DNA that contains all elements for the infectious replication cycle. J. Virol. *65*, 3131–3139.

Noteborn, M.H., Todd, D., Verschueren, C.A., de Gauw, H.W., Curran, W.L., Veldkamp, S., Douglas, A.J., McNulty, M.S., van der EB, A.J., and Koch, G. (1994). A single chicken anemia virus protein induces apoptosis. J. Virol. *68*, 346–351.

Olijslagers, S.J., Zhang, Y.-H., Backendorf, C., and Noteborn, M.H.M. (2007). Additive cytotoxic effect of apoptin and chemotherapeutic agents paclitaxel and etoposide on human tumour cells. Basic Clin. Pharmacol. Toxicol. *100*, 127–131.

Pan, Y., Fang, L., Fan, H., Luo, R., Zhao, Q., Chen, H., and Xiao, S. (2010). Antitumor effects of a recombinant pseudotype baculovirus expressing Apoptin *in vitro* and *in vivo*. Int. J. Cancer *126*, 2741–2751.

Panigrahi, S., Klonisch, T., and Los, M. (2008). The art of killing: double stroke with apoptin and survivin as a novel approach in cancer therapy. Cancer Biol. Ther. *7*, 1061–1062.

Panigrahi, S., Stetefeld, J., Jangamreddy, J.R., Mandal, S., Mandal, S.K., and Los, M. (2012). Modeling of molecular interaction between apoptin, BCR-Abl and CrkL--an alternative approach to conventional rational drug design. PLoS One *7*, e28395.

Pantoja-Uceda, D., and Santoro, J. (2008). Amino acid type identification in NMR spectra of proteins via beta- and gamma-carbon edited experiments.

J. Magn. Reson. *195*, 187–195.

Pantoja-Uceda, D., and Santoro, J. (2012). New amino acid residue type identification experiments valid for protonated and deuterated proteins. J. Biomol. NMR *54*, 145–153.

Parato, K.A., Senger, D., Forsyth, P.A.J., and Bell, J.C. (2005). Recent progress in the battle between oncolytic viruses and tumours. Nat. Rev. Cancer *5*, 965–976.

Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M., et al. (2015). A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. Cell *162*, 1066–1077.

Peng, D.-J., Sun, J., Wang, Y.-Z., Tian, J., Zhang, Y.-H., Noteborn, M.H.M., and Qu, S. (2007). Inhibition of hepatocarcinoma by systemic delivery of Apoptin gene via the hepatic asialoglycoprotein receptor. Cancer Gene Ther. *14*, 66–73.

Peters, M.A., Jackson, D.C., Crabb, B.S., and Browning, G.F. (2002). Chicken anemia virus VP2 is a novel dual specificity protein phosphatase. J. Biol. Chem. *277*, 39566–39573.

Peters, M.A., Jackson, D.C., Crabb, B.S., and Browning, G.F. (2005). Mutation of chicken anemia virus VP2 differentially affects serine/threonine and tyrosine protein phosphatase activities. J. Gen. Virol. *86*, 623–630.

Pietersen, A.M., van der Eb, M.M., Rademaker, H.J., van den Wollenberg, D.J., Rabelink, M.J., Kuppen, P.J., van Dierendonck, J.H., van Ormondt, H., Masman, D., van de Velde, C.J., et al. (1999). Specific tumor-cell killing with adenovirus vectors containing the apoptin gene. Gene Ther. *6*, 882–892.

Pietersen, A.M., Rutjes, S.A., van Tongeren, J., Vogels, R., Wesseling, J.G., and Noteborn, M.H.M. (2004). The tumor-selective viral protein apoptin effectively kills human biliary tract cancer cells. J. Mol. Med. (Berl). *82*, 56–63.

Pontius, B.W. (1993). Close encounters: why unstructured, polymeric domains can increase rates of specific macromolecular association. Trends Biochem. Sci. *18*, 181–186.

Poon, I.K.H., Oro, C., Dias, M.M., Zhang, J.-P., and Jans, D.A. (2005a). A tumor cell-specific nuclear targeting signal within chicken anemia virus VP3/apoptin. J. Virol. *79*, 1339–1341.

Poon, I.K.H., Oro, C., Dias, M.M., Zhang, J., and Jans, D. a (2005b). Apoptin nuclear accumulation is modulated by a CRM1-recognized nuclear export

signal that is active in normal but not in tumor cells. Cancer Res. *65*, 7059–7064.

Prilusky, J., Felder, C.E., Zeev-Ben-Mordehai, T., Rydberg, E.H., Man, O., Beckmann, J.S., Silman, I., and Sussman, J.L. (2005). FoldIndex(C): a simple tool to predict whether a given protein sequence is intrinsically unfolded. Bioinformatics *21*, 3435–3438.

Pringle, C.R. (2014). Virus taxonomy at the XIth international congress of virology, Sydney, Australia, 1999. Arch. Virol. *144*, 2065–2070.

Qi, Y., Guo, H., Hu, N., He, D., Zhang, S., Chu, Y., Huang, Y., Li, X., Sun, L., and Jin, N. (2014). Preclinical pharmacology and toxicology study of Ad-hTERT-E1a-Apoptin, a novel dual cancer-specific oncolytic adenovirus. Toxicol. Appl. Pharmacol. *280*, 362–369.

Renner, C., Schleicher, M., Moroder, L., and Holak, T.A. (2002). Practical aspects of the 2D ^{15}N-[^{1}H]-NOE experiment. J. Biomol. NMR *23*, 23–33.

Rietsch, A., Belin, D., Martin, N., and Beckwith, J. (1996). An *in vivo* pathway for disulfide bond isomerization in *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. *93*, 13048–13053.

Ring, J., Seifert, J., Jesch, F., and Brendel, W. (1977). Anaphylactoid reactions due to non-immune complex serum protein aggregates. Monogr. Allergy *12*, 27–35.

Rohn, J.L., Zhang, Y.-H., Aalbers, R.I.J.M., Otto, N., Den Hertog, J., Henriquez, N. V, Van De Velde, C.J.H., Kuppen, P.J.K., Mumberg, D., Donner, P., et al. (2002). A tumor-specific kinase activity regulates the viral death protein Apoptin. J. Biol. Chem. *277*, 50820–50827.

Rohn, J.L., Zhang, Y.-H., Leliveld, S.R., Danen-van Oorschot, A.A.A.M., Henriquez, N. V, Abrahams, J.P., and Noteborn, M.H.M. (2005). Relevance of apoptin's integrity for its functional behavior. J. Virol. *79*, 1337–1338.

Rollano Peñaloza, O.M., Lewandowska, M., Stetefeld, J., Ossysek, K., Madej, M., Bereta, J., Sobczak, M., Shojaei, S., Ghavami, S., and Łos, M.J. (2014). Apoptins: selective anticancer agents. Trends Mol. Med. *20*, 519–528.

Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J., and Dunker, A.K. (2001). Sequence complexity of disordered protein. Proteins *42*, 38–48.

Roulston, A., Marcellus, R.C., and Branton, P.E. (1999). Viruses and apoptosis. Annu. Rev. Microbiol. *53*, 577–628.

Sabaté, R., Lascu, I., and Saupe, S.J. (2008). On the binding of Thioflavin-T to HET-s amyloid fibrils assembled at pH 2. J. Struct. Biol. *162*, 387–396.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U. S. A. *74*, 5463–5467.

Sauvage, V., Cheval, J., Foulongne, V., Gouilh, M.A., Pariente, K., Manuguerra, J.C., Richardson, J., Dereure, O., Lecuit, M., Burguiere, A., et al. (2011). Identification of the first human gyrovirus, a virus related to chicken anemia virus. J. Virol. *85*, 7948–7950.

Schubert, M., Ball, L.J., Oschkinat, H., and Schmieder, P. (2000). Bridging the gap: A set of selective ^{1}H-^{15}N-correlations to link sequential neighbors of prolines. J. Biomol. NMR *17*, 331–335.

Schubert, M., Labudde, D., Oschkinat, H., and Schmieder, P. (2002). A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on ^{13}C chemical shift statistics. J. Biomol. NMR *24*, 149–154.

Scott, A.M., Wolchok, J.D., and Old, L.J. (2012). Antibody therapy of cancer. Nat. Rev. *12*, 278–287.

Semisotnov, G. V, Rodionova, N.A., Razgulyaev, O.I., Uversky, V.N., Gripas', A.F., and Gilmanshin, R.I. (1991). Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. Biopolymers *31*, 119–128.

Shelkovnikova, T.A., Kulikova, A.A., Tsvetkov, P.O., Peters, O., Bachurin, S.O., Buchman, V.L., and Ninkina, N.N. (2012). Proteinopathies, neurodegenerative disorders with protein aggregation-based pathology. Mol. Biol. *46*, 362–374.

Sickmeier, M., Hamilton, J.A., LeGall, T., Vacic, V., Cortese, M.S., Tantos, A., Szabo, B., Tompa, P., Chen, J., Uversky, V.N., et al. (2007). DisProt: the Database of Disordered Proteins. Nucleic Acids Res. *35*, D786–D793.

Singh, S.M., and Panda, A.K. (2005). Solubilization and refolding of bacterial inclusion body proteins. J. Biosci. Bioeng. *99*, 303–310.

Slamon, D.J., Clark, G.M., Wong, S.G., Levin, W.J., Ullrich, A., and McGuire, W.L. (1987). Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science *235*, 177–182.

Sreerama, N., Venyaminov, S.Y., and Woody, R.W. (1999). Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy. Protein Sci. *8*, 370–380.

Studier, F.W., and Moffatt, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. *189*, 113–130.

Sun, G.J., Tong, X., Dong, Y., Mei, Z.Z., and Sun, Z.X. (2002). [Identification of a protein interacting with apoptin from human leucocyte cDNA library by using yeast two-hybrid screening]. Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai). *34*, 369–372.

Sun, J., Yan, Y., Wang, X.-T., Liu, X.-W., Peng, D.-J., Wang, M., Tian, J., Zong, Y.-Q., Zhang, Y.-H., Noteborn, M.H.M., et al. (2009). PTD4-apoptin protein therapy inhibits tumor growth *in vivo*. Int. J. Cancer *124*, 2973–2981.

Susin, S.A., Lorenzo, H.K., Zamzami, N., Marzo, I., Snow, B.E., Brothers, G.M., Mangion, J., Jacotot, E., Costantini, P., Loeffler, M., et al. (1999). Molecular characterization of mitochondrial apoptosis-inducing factor. Nature *397*, 441–446.

Tamiola, K., Acar, B., and Mulder, F.A.A. (2010). Sequence-specific random coil chemical shifts of intrinsically disordered proteins. J. Am. Chem. Soc. *132*, 18000–18003.

Teodoro, J.G., Heilman, D.W., Parker, A.E., and Green, M.R. (2004). The viral protein Apoptin associates with the anaphase-promoting complex to induce G2/M arrest and apoptosis in the absence of p53. Genes Dev. *18*, 1952–1957.

Tojkander, S., Gateva, G., and Lappalainen, P. (2012). Actin stress fibers--assembly, dynamics and biological roles. J. Cell Sci. *125*, 1855–1864.

Tompa, P., and Fersht, A. (2009). Structure and function of intrinsically disordered proteins. CRC Press.

Torrent, M., Valle, J., Nogués, M.V., Boix, E., and Andreu, D. (2011). The generation of antimicrobial peptide activity: a trade-off between charge and aggregation? Angew. Chemie *123*, 10874–10877.

Ulrich, E.L., Akutsu, H., Doreleijers, J.F., Harano, Y., Ioannidis, Y.E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z., et al. (2008). BioMagResBank. Nucleic Acids Res. *36*, D402–D408.

Uversky, V., and Longhi, S. (2011). Flexible viruses: structural disorder in viral proteins (John Wiley & Sons).

Uversky, V.N., and Fink, A.L. (2004). Conformational constraints for amyloid fibrillation: The importance of being unfolded. Biochim. Biophys. Acta - Proteins Proteomics *1698*, 131–153.

Uversky, V.N., Gillespie, J.R., and Fink, A.L. (2000). Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins *41*, 415–427.

Vera, A., González-Montalbán, N., Arís, A., and Villaverde, A. (2007). The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. Biotechnol. Bioeng. *96*, 1101–1106.

Vert, A., Castro, J., Ruiz-Martínez, S., Tubert, P., Escribano, D., Ribó, M., Vilanova, M., and Benito, A. (2012). Generation of new cytotoxic human ribonuclease variants directed to the nucleus. Mol. Pharm. *9*, 2894–2902.

Wadia, J.S., Wagner, M. V, Ezhevsky, S.A., and Dowdy, S.F. (2004). Apoptin/VP3 contains a concentration-dependent nuclear localization signal (NLS), not a tumorigenic selective NLS. J. Virol. *78*, 6077–6078.

Wang, Q.-M., Fan, G.-C., Chen, J.-Z., Chen, H.-P., and He, F.-C. (2004). A putative NES mediates cytoplasmic localization of Apoptin in normal cells. Acta Biochim. Biophys. Sin. (Shanghai). *36*, 817–823.

Xu, M.-M., Ren, W.-M., Tang, X.-C., Hu, Y.-H., and Zhang, H.-Y. (2016). Advances in development of fluorescent probes for detecting amyloid-β aggregates. Acta Pharmacol. Sin.

Xue, B., Oldfield, C.J., Dunker, A.K., and Uversky, V.N. (2009). CDF it all: consensus prediction of intrinsically disordered proteins based on various cumulative distribution functions. FEBS Lett. *583*, 1469–1474.

Yao, Q., Li, M., Yang, H., Chai, H., Fisher, W., and Chen, C. (2005). Roles of cyclophilins in cancers and other organ systems. World J. Surg. *29*, 276–280.

Yi, Y., Jong Noh, M., and Hee Lee, K. (2011). Current advances in retroviral gene therapy. Curr. Gene Ther. *11*, 218–228.

Yuasa, N., Yoshida, I., and Taniguchi, T. (1976). Isolation of a reticuloendotheliosis virus from chickens inoculated with Marek's disease vaccine. Natl. Inst. Anim. Health Q. (Tokyo). *16*, 141–151.

Yuasa, N., Taniguchi, T., and Yoshida, I. (1979). Isolation and some characteristics of an agent inducing anemia in chicks. Avian Diseases, *23*, 366-385.

Zhang, M., Guller, S., and Huang, Y. (2007). Method to enhance transfection efficiency of cell lines and placental fibroblasts. Placenta *28*, 779–782.

Zhang, M., Wang, J., Li, C., Hu, N., Wang, K., Ji, H., He, D., Quan, C., Li, X., Jin, N., et al. (2013). Potent growth-inhibitory effect of a dual cancer-specific oncolytic adenovirus expressing apoptin on prostate carcinoma. Int. J. Oncol. *42*, 1052–1060.

Zhang, Y.H., Leliveld, S.R., Kooistra, K., Molenaar, C., Rohn, J.L., Tanke, H.J., Abrahams, J.P., and Noteborn, M.H.M. (2003). Recombinant apoptin multimers kill tumor cells but are nontoxic and epitope-shielded in a normal-cell-specific fashion. Exp. Cell Res. *289*, 36–46.

Zhao, J., Gao, P., Xiao, W., Fan, L.Q., Wang, F.J., Li, S.X., and Liu, J.W. (2011). A novel human derived cell-penetrating peptide in drug delivery. Mol. Biol. Rep. *38*, 2649–2656.

Zhao, M., Hu, B., Gu, Z., Joo, K.-I., Wang, P., and Tang, Y. (2013). Degradable polymeric nanocapsule for efficient intracellular delivery of a high molecular weight tumor-selective protein complex. Nano Today *8*, 11–20.

Zheng, L., Schickling, O., Peter, M.E., and Lenardo, M.J. (2001). The death effector domain-associated factor plays distinct regulatory roles in the nucleus and cytoplasm. J. Biol. Chem. *276*, 31945–31952.

Zhou, W., and Bao, S. (2014). PML-mediated signaling and its role in cancer stem cells. Oncogene *33*, 1475–1484.

Zhuang, S.M., Shvarts, A., Van Ormondt, H., Jochemsen, A.G., Van Der Eb, A.J., and Noteborn, M.H.M. (1995). Apoptin, a protein derived from chicken anemia virus, induces p53-independent apoptosis in human osteosarcoma cells. Cancer Res. *55*, 486–489.

Zimmerman, R., Peng, D.J., Lanz, H., Zhang, Y.H., Danen-Van Oorschot, A., Qu, S., Backendorf, C., and Noteborn, M. (2012). PP2A inactivation is a crucial step in triggering apoptin-induced tumor-selective cell killing. Cell Death Dis. *3*, e291.

Annex

residue sequence for the table:
MGSSHHHHHHSSGLVPRGSHMSLCGCANARAPTLRSATADNSESTGFKNVPDLRTDQPKPPSKKRSCDPSEYR VSELKESLITTTPSRPRTAKRRIRL

Chemical Shift Ambiguity Index Value Definitions

The values other than 1 are used for those atoms with different chemical shifts that cannot be assigned to stereospecific atoms or to specific residues or chains.

Index Value	Definition
1	Unique (including isolated methyl protons, geminal atoms, and geminal methyl groups with identical chemical shifts) (e.g. ILE HD11, HD12, HD13 protons)
2	Ambiguity of geminal atoms or geminal methyl proton groups (e.g. ASP HB2 and HB3 protons, LEU CD1 and CD2 carbons, or LEU HD11, HD12, HD13 and HD21, HD22, HD23 methyl protons)
3	Aromatic atoms on opposite sides of symmetrical rings (e.g. TYR HE1 and HE2 protons)
4	Intraresidue ambiguities (e.g. LYS HG and HD protons or TRP HZ2 and HZ3 protons)
5	Interresidue ambiguities (LYS 12 vs. LYS 27)
6	Intermolecular ambiguities (e.g. ASP 31 CA in monomer 1 and ASP 31 CA in monomer 2 of an asymmetrical homodimer, duplex DNA assignments, or other assignments that may apply to atoms in one or more molecule in the molecular assembly)
9	Ambiguous, specific ambiguity not

```
loop_
  _Atom_chem_shift.ID
  _Atom_chem_shift.Entity_assembly_ID
  _Atom_chem_shift.Comp_index_ID
  _Atom_chem_shift.Comp_ID
  _Atom_chem_shift.Atom_ID
  _Atom_chem_shift.Atom_type
  _Atom_chem_shift.Atom_isotope_number
  _Atom_chem_shift.Val
  _Atom_chem_shift.Val_err
  _Atom_chem_shift.Assign_fig_of_merit
  _Atom_chem_shift.Ambiguity_code
  _Atom_chem_shift.Occupancy
  _Atom_chem_shift.Details
```

Entity Comp Atom Atom isot. Val fig. of Ambig.

#ID ID ID ID ID type num Val err merit code Occup.

Details

1	.	1	GLY	H	H	1	8.729	0.0050	.	1	51	.	8	HIS	CA	C	13	55.42	0.0480	.	1
2	.	1	GLY	QA	H	1	3.928	0.0050	.	2	52	.	8	HIS	CB	C	13	29.18	0.0458	.	1
3	.	1	GLY	C	C	13	170.4	0.0280		1	53	.	8	HIS	N	N	15	120.9	0.0241	.	1
4	.	1	GLY	CA	C	13	43.43	0.0928	.	1	54	.	9	HIS	H	H	1	8.930	0.0011	.	1
5	.	1	GLY	N	N	15	111.2	0.0050		1	55	.	9	HIS	HA	H	1	4.757	0.0050	.	1
6	.	2	SER	H	H	1	8.848	0.0032	.	1	56	.	9	HIS	QB	H	1	3.231	0.0050	.	2
7	.	2	SER	HA	H	1	4.523	0.0050	.	1	57	.	9	HIS	C	C	13	174.5	0.0030	.	1
8	.	2	SER	QB	H	1	3.855	0.0050	.	2	58	.	9	HIS	CA	C	13	55.38	0.0566	.	1
9	.	2	SER	C	C	13	174.7	0.0082	.	1	59	.	9	HIS	CB	C	13	29.25	0.0332	.	1
10	.	2	SER	CA	C	13	58.26	0.0380	.	1	60	.	9	HIS	N	N	15	121.5	0.0240	.	1
11	.	2	SER	CB	C	13	63.97	0.1139	.	1	61	.	10	SER	H	H	1	8.744	0.0034	.	1
12	.	2	SER	N	N	15	115.8	0.0210	.	1	62	.	10	SER	HA	H	1	4.500	0.0050	.	1
13	.	3	SER	H	H	1	8.667	0.0029	.	1	63	.	10	SER	QB	H	1	3.882	0.0050	.	2
14	.	3	SER	HA	H	1	4.410	0.0050	.	1	64	.	10	SER	C	C	13	174.7	0.0043	.	1
15	.	3	SER	QB	H	1	3.796	0.0050	.	2	65	.	10	SER	CA	C	13	58.35	0.0363	.	1
16	.	3	SER	C	C	13	174.4	0.0050	.	1	66	.	10	SER	CB	C	13	63.81	0.1156	.	1
17	.	3	SER	CA	C	13	58.39	0.0600	.	1	67	.	10	SER	N	N	15	118.6	0.0231	.	1
18	.	3	SER	CB	C	13	63.67	0.0697	.	1	68	.	11	SER	H	H	1	8.721	0.0029	.	1
19	.	3	SER	N	N	15	118.2	0.0252	.	1	69	.	11	SER	HA	H	1	4.497	0.0050	.	1
20	.	4	HIS	H	H	1	8.646	0.0329	.	1	70	.	11	SER	QB	H	1	3.920	0.0050	.	2
21	.	4	HIS	HA	H	1	4.694	0.0050	.	1	71	.	11	SER	C	C	13	175.0	0.0069	.	1
22	.	4	HIS	QB	H	1	3.126	0.0050	.	2	72	.	11	SER	CA	C	13	58.68	0.1144	.	1
23	.	4	HIS	C	C	13	174.2	0.0051	.	1	73	.	11	SER	CB	C	13	63.77	0.0920	.	1
24	.	4	HIS	CA	C	13	55.09	0.0601	.	1	74	.	11	SER	N	N	15	118.7	0.0083	.	1
25	.	4	HIS	CB	C	13	28.79	0.0439	.	1	75	.	12	GLY	H	H	1	8.539	0.0025	.	1
26	.	4	HIS	N	N	15	120.4	0.0871	.	1	76	.	12	GLY	QA	H	1	3.959	0.0050	.	2
27	.	5	HIS	H	H	1	8.667	0.0020	.	1	77	.	12	GLY	C	C	13	173.9	0.0051	.	1
28	.	5	HIS	HA	H	1	4.675	0.0050	.	1	78	.	12	GLY	CA	C	13	45.17	0.0441	.	1
29	.	5	HIS	QB	H	1	3.115	0.0050	.	2	79	.	12	GLY	N	N	15	110.9	0.0154	.	1
30	.	5	HIS	C	C	13	174.2	0.0143	.	1	80	.	13	LEU	H	H	1	8.232	0.0022	.	1
31	.	5	HIS	CA	C	13	55.12	0.0780	.	1	81	.	13	LEU	HA	H	1	4.328	0.0050	.	1
32	.	5	HIS	CB	C	13	29.02	0.0249	.	1	82	.	13	LEU	HB	H	1	1.576	0.0050	.	2
33	.	5	HIS	N	N	15	119.6	0.0093	.	1	83	.	13	LEU	C	C	13	177.4	0.0131	.	1
34	.	6	HIS	H	H	1	8.849	0.0022	.	1	84	.	13	LEU	CA	C	13	55.02	0.0431	.	1
35	.	6	HIS	QB	H	1	3.138	0.0050	.	2	85	.	13	LEU	CB	C	13	42.27	0.0145	.	1
36	.	6	HIS	C	C	13	174.2	0.0152	.	1	86	.	13	LEU	CG	C	13	26.92	0.1000	.	1
37	.	6	HIS	CA	C	13	55.28	0.0459	.	1	87	.	13	LEU	CD1	C	13	23.53	0.1000	.	2
38	.	6	HIS	CB	C	13	29.04	0.0658	.	1	88	.	13	LEU	CD2	C	13	24.97	0.1000	.	2
39	.	6	HIS	N	N	15	120.2	0.0054	.	1	89	.	13	LEU	N	N	15	121.7	0.0125	.	1
40	.	7	HIS	H	H	1	8.907	0.0018	.	1	90	.	14	VAL	H	H	1	8.333	0.0025	.	1
41	.	7	HIS	HA	H	1	4.686	0.0050	.	1	91	.	14	VAL	C	C	13	174.4	0.1000	.	1
42	.	7	HIS	QB	H	1	3.149	0.0050	.	2	92	.	14	VAL	CA	C	13	59.84	0.0193	.	1
43	.	7	HIS	C	C	13	174.3	0.0107	.	1	93	.	14	VAL	CB	C	13	32.51	0.1000	.	1
44	.	7	HIS	CA	C	13	55.20	0.0659	.	1	94	.	14	VAL	N	N	15	123.6	0.0583	.	1
45	.	7	HIS	CB	C	13	29.09	0.0163	.	1	95	.	15	PRO	HA	H	1	4.408	0.0050	.	1
46	.	7	HIS	N	N	15	120.5	0.0160	.	1	96	.	15	PRO	HB2	H	1	1.883	0.0050	.	2
47	.	8	HIS	H	H	1	8.929	0.0022	.	1	97	.	15	PRO	HB3	H	1	2.297	0.0050	.	2
48	.	8	HIS	HA	H	1	4.697	0.0050	.	1	98	.	15	PRO	C	C	13	176.9	0.0125	.	1
49	.	8	HIS	QB	H	1	3.189	0.0050	.	2	99	.	15	PRO	CA	C	13	63.08	0.0499	.	1
50	.	8	HIS	C	C	13	174.3	0.0120	.	1	100	.	15	PRO	CB	C	13	32.18	0.0611	.	1

101	.	15	PRO	CG	C	13	27.51	0.1000	.	1
102	.	15	PRO	CD	C	13	51.18	0.1000	.	1
103	.	16	ARG	H	H	1	8.678	0.0032	.	1
104	.	16	ARG	HA	H	1	4.313	0.0050	.	1
105	.	16	ARG	QB	H	1	1.845	0.0050	.	2
106	.	16	ARG	C	C	13	177.2	0.0184	.	1
107	.	16	ARG	CA	C	13	56.47	0.0371	.	1
108	.	16	ARG	CB	C	13	30.81	0.0269	.	1
109	.	16	ARG	CG	C	13	27.03	0.1000	.	1
110	.	16	ARG	CD	C	13	43.38	0.1000	.	1
111	.	16	ARG	N	N	15	122.5	0.0376	.	1
112	.	17	GLY	H	H	1	8.644	0.0023	.	1
113	.	17	GLY	QA	H	1	3.989	0.0050	.	2
114	.	17	GLY	C	C	13	174.1	0.0276	.	1
115	.	17	GLY	CA	C	13	45.16	0.0494	.	1
116	.	17	GLY	N	N	15	110.7	0.0184	.	1
117	.	18	SER	H	H	1	8.350	0.0031	.	1
118	.	18	SER	HA	H	1	4.424	0.0050	.	1
119	.	18	SER	QB	H	1	3.833	0.0050	.	2
120	.	18	SER	C	C	13	174.6	0.0083	.	1
121	.	18	SER	CA	C	13	58.38	0.0678	.	1
122	.	18	SER	CB	C	13	63.84	0.0992	.	1
123	.	18	SER	N	N	15	115.7	0.0173	.	1
124	.	19	HIS	H	H	1	8.796	0.0032	.	1
125	.	19	HIS	HA	H	1	4.737	0.0050	.	1
126	.	19	HIS	QB	H	1	3.254	0.0050	.	2
127	.	19	HIS	C	C	13	174.5	0.0290	.	1
128	.	19	HIS	CA	C	13	55.47	0.0748	.	1
129	.	19	HIS	CB	C	13	28.78	0.0261	.	1
130	.	19	HIS	N	N	15	120.5	0.0173	.	1
131	.	20	MET	H	H	1	8.521	0.0043	.	1
132	.	20	MET	HA	H	1	4.496	0.0050	.	1
133	.	20	MET	QB	H	1	2.029	0.0050	.	2
134	.	20	MET	C	C	13	176.2	0.0791	.	1
135	.	20	MET	CA	C	13	55.52	0.0915	.	1
136	.	20	MET	CB	C	13	32.77	0.0233	.	1
137	.	20	MET	N	N	15	121.5	0.0496	.	1
138	.	21	SER	H	H	1	8.542	0.0031	.	1
139	.	21	SER	C	C	13	174.7	0.0000	.	1
140	.	21	SER	CA	C	13	58.42	0.0215	.	1
141	.	21	SER	CB	C	13	63.65	0.0000	.	1
142	.	21	SER	N	N	15	117.7	0.0153	.	1
143	.	25	CYS	HA	H	1	4.723	0.0050	.	1
144	.	25	CYS	HB	H	1	3.314	0.0050	.	2
145	.	25	CYS	C	C	13	175.0	0.0030	.	1
146	.	25	CYS	CA	C	13	55.19	0.0324	.	1
147	.	25	CYS	CB	C	13	40.58	0.0065	.	1
148	.	26	ALA	H	H	1	8.658	0.0024	.	1
149	.	26	ALA	HA	H	1	4.292	0.0050	.	1
150	.	26	ALA	MB	H	1	1.394	0.0050	.	1
151	.	26	ALA	C	C	13	177.7	0.0377	.	1
152	.	26	ALA	CA	C	13	53.15	0.0943	.	1
153	.	26	ALA	CB	C	13	18.90	0.0325	.	1
154	.	26	ALA	N	N	15	126.2	0.0173	.	1
155	.	27	ASN	H	H	1	8.414	0.0126	.	1
156	.	27	ASN	HA	H	1	4.672	0.0050	.	1
157	.	27	ASN	QB	H	1	2.799	0.0050	.	2
158	.	27	ASN	C	C	13	175.0	0.0136	.	1
159	.	27	ASN	CA	C	13	53.18	0.0808	.	1
160	.	27	ASN	CB	C	13	36.67	0.0513	.	1
161	.	27	ASN	N	N	15	117.3	0.0325	.	1
162	.	28	ALA	H	H	1	8.233	0.0060	.	1
163	.	28	ALA	HA	H	1	4.304	0.0050	.	1
164	.	28	ALA	MB	H	1	1.391	0.0050	.	1
165	.	28	ALA	C	C	13	177.6	0.0217	.	1
166	.	28	ALA	CA	C	13	52.65	0.0378	.	1
167	.	28	ALA	CB	C	13	19.23	0.0651	.	1
168	.	28	ALA	N	N	15	124.2	0.0174	.	1
169	.	29	ARG	H	H	1	8.339	0.0021	.	1
170	.	29	ARG	HA	H	1	4.320	0.0050	.	1
171	.	29	ARG	QB	H	1	1.809	0.0050	.	2
172	.	29	ARG	C	C	13	175.9	0.0274	.	1
173	.	29	ARG	CA	C	13	55.66	0.0436	.	1
174	.	29	ARG	CB	C	13	30.90	0.0448	.	1
175	.	29	ARG	CG	C	13	27.07	0.1000	.	1
176	.	29	ARG	CD	C	13	43.41	0.1000	.	1
177	.	29	ARG	N	N	15	120.2	0.0131	.	1
178	.	30	ALA	H	H	1	8.426	0.0030	.	1
179	.	30	ALA	C	C	13	175.6	0.1000	.	1
180	.	30	ALA	CA	C	13	50.65	0.0419	.	1
181	.	30	ALA	N	N	15	127.1	0.0123	.	1
182	.	31	PRO	HA	H	1	4.489	0.0050	.	1
183	.	31	PRO	HB2	H	1	1.910	0.0050	.	2
184	.	31	PRO	HB3	H	1	2.316	0.0050	.	2
185	.	31	PRO	C	C	13	177.0	0.0010	.	1
186	.	31	PRO	CA	C	13	63.08	0.0655	.	1
187	.	31	PRO	CB	C	13	32.15	0.0607	.	1
188	.	31	PRO	CG	C	13	27.45	0.1000	.	1
189	.	31	PRO	CD	C	13	50.70	0.1000	.	1
190	.	32	THR	H	H	1	8.397	0.0022	.	1
191	.	32	THR	HA	H	1	4.283	0.0050	.	1
192	.	32	THR	HB	H	1	4.181	0.0050	.	1
193	.	32	THR	C	C	13	174.6	0.0114	.	1
194	.	32	THR	CA	C	13	61.89	0.1387	.	1
195	.	32	THR	CB	C	13	69.80	0.0682	.	1
196	.	32	THR	CG2	C	13	22.07	0.1000	.	1
197	.	32	THR	N	N	15	115.4	0.0260	.	1
198	.	33	LEU	H	H	1	8.448	0.0044	.	1
199	.	33	LEU	HA	H	1	4.322	0.0050	.	1
200	.	33	LEU	HB	H	1	1.624	0.0050	.	2
201	.	33	LEU	C	C	13	177.3	0.0272	.	1
202	.	33	LEU	CA	C	13	55.09	0.0678	.	1
203	.	33	LEU	CB	C	13	42.29	0.0240	.	1
204	.	33	LEU	CG	C	13	26.87	0.1000	.	1
205	.	33	LEU	CD1	C	13	23.78	0.1000	.	2
206	.	33	LEU	CD2	C	13	24.90	0.1000	.	2
207	.	33	LEU	N	N	15	125.7	0.0194	.	1
208	.	34	ARG	H	H	1	8.550	0.0029	.	1
209	.	34	ARG	HA	H	1	4.360	0.0050	.	1
210	.	34	ARG	QB	H	1	1.818	0.0050	.	2
211	.	34	ARG	C	C	13	176.3	0.0367	.	1
212	.	34	ARG	CA	C	13	56.05	0.0265	.	1
213	.	34	ARG	CB	C	13	30.81	0.0640	.	1
214	.	34	ARG	CG	C	13	27.04	0.1000	.	1
215	.	34	ARG	CD	C	13	43.34	0.1000	.	1
216	.	34	ARG	N	N	15	123.1	0.0298	.	1
217	.	35	SER	H	H	1	8.503	0.0016	.	1
218	.	35	SER	HA	H	1	4.428	0.0050	.	1

219	.	35	SER	QB	H	1	3.881	0.0050	.	2	278	.	43	SER	C	C	13	175.2	0.0164	.	1
220	.	35	SER	C	C	13	174.5	0.0066	.	1	279	.	43	SER	CA	C	13	58.64	0.1102	.	1
221	.	35	SER	CA	C	13	58.21	0.0325	.	1	280	.	43	SER	CB	C	13	63.61	0.0881	.	1
222	.	35	SER	CB	C	13	63.70	0.1085	.	1	281	.	43	SER	N	N	15	116.5	0.0503	.	1
223	.	35	SER	N	N	15	117.7	0.0151	.	1	282	.	44	THR	H	H	1	8.260	0.0031	.	1
224	.	36	ALA	H	H	1	8.617	0.0017	.	1	283	.	44	THR	HA	H	1	4.338	0.0050	.	1
225	.	36	ALA	HA	H	1	4.402	0.0050	.	1	284	.	44	THR	HB	H	1	4.261	0.0050	.	1
226	.	36	ALA	MB	H	1	1.443	0.0050	.	1	285	.	44	THR	C	C	13	175.3	0.0099	.	1
227	.	36	ALA	C	C	13	178.1	0.0128	.	1	286	.	44	THR	CA	C	13	62.11	0.1025	.	1
228	.	36	ALA	CA	C	13	52.83	0.0223	.	1	287	.	44	THR	CB	C	13	69.60	0.0873	.	1
229	.	36	ALA	CB	C	13	19.31	0.0530	.	1	288	.	44	THR	N	N	15	115.3	0.0079	.	1
230	.	36	ALA	N	N	15	126.7	0.0135	.	1	289	.	45	GLY	H	H	1	8.392	0.0022	.	1
231	.	37	THR	H	H	1	8.212	0.0041	.	1	290	.	45	GLY	QA	H	1	3.907	0.0050	.	2
232	.	37	THR	HA	H	1	4.338	0.0050	.	1	291	.	45	GLY	C	C	13	173.9	0.0253	.	1
233	.	37	THR	HB	H	1	4.266	0.0050	.	1	292	.	45	GLY	CA	C	13	45.20	0.0383	.	1
234	.	37	THR	C	C	13	174.5	0.0033	.	1	293	.	45	GLY	N	N	15	110.8	0.0091	.	1
235	.	37	THR	CA	C	13	61.68	0.1005	.	1	294	.	46	PHE	H	H	1	8.189	0.0040	.	1
236	.	37	THR	CB	C	13	69.80	0.0625	.	1	295	.	46	PHE	HA	H	1	4.595	0.0050	.	1
237	.	37	THR	CG2	C	13	22.00	0.1000	.	1	296	.	46	PHE	QB	H	1	3.068	0.0050	.	2
238	.	37	THR	N	N	15	112.9	0.0315	.	1	297	.	46	PHE	C	C	13	175.7	0.0282	.	1
239	.	38	ALA	H	H	1	8.366	0.0028	.	1	298	.	46	PHE	CA	C	13	57.86	0.0376	.	1
240	.	38	ALA	HA	H	1	4.314	0.0050	.	1	299	.	46	PHE	CB	C	13	39.53	0.0503	.	1
241	.	38	ALA	MB	H	1	1.402	0.0050	.	1	300	.	46	PHE	N	N	15	120.3	0.0045	.	1
242	.	38	ALA	C	C	13	177.5	0.0278	.	1	301	.	47	LYS	H	H	1	8.317	0.0024	.	1
243	.	38	ALA	CA	C	13	52.62	0.0433	.	1	302	.	47	LYS	HA	H	1	4.294	0.0050	.	1
244	.	38	ALA	CB	C	13	19.21	0.0524	.	1	303	.	47	LYS	QB	H	1	1.722	0.0050	.	2
245	.	38	ALA	N	N	15	126.4	0.0146	.	1	304	.	47	LYS	C	C	13	175.7	0.0104	.	1
246	.	39	ASP	H	H	1	8.416	0.0023	.	1	305	.	47	LYS	CA	C	13	55.93	0.0370	.	1
247	.	39	ASP	HA	H	1	4.598	0.0050	.	1	306	.	47	LYS	CB	C	13	33.18	0.0370	.	1
248	.	39	ASP	QB	H	1	2.703	0.0050	.	2	307	.	47	LYS	CG	C	13	24.66	0.1000	.	1
249	.	39	ASP	C	C	13	176.2	0.0091	.	1	308	.	47	LYS	CD	C	13	29.02	0.1000	.	1
250	.	39	ASP	CA	C	13	54.30	0.1613	.	1	309	.	47	LYS	N	N	15	123.5	0.0427	.	1
251	.	39	ASP	CB	C	13	40.98	0.0384	.	1	310	.	48	ASN	H	H	1	8.507	0.0028	.	1
252	.	39	ASP	N	N	15	119.7	0.0289	.	1	311	.	48	ASN	HA	H	1	4.650	0.0050	.	1
253	.	40	ASN	H	H	1	8.493	0.0025	.	1	312	.	48	ASN	QB	H	1	2.756	0.0050	.	2
254	.	40	ASN	HA	H	1	4.788	0.0050	.	1	313	.	48	ASN	C	C	13	174.8	0.0025	.	1
255	.	40	ASN	QB	H	1	2.843	0.0050	.	2	314	.	48	ASN	CA	C	13	53.32	0.1268	.	1
256	.	40	ASN	C	C	13	175.8	0.0133	.	1	315	.	48	ASN	CB	C	13	38.85	0.0306	.	1
257	.	40	ASN	CA	C	13	53.18	0.0761	.	1	316	.	48	ASN	N	N	15	120.4	0.0148	.	1
258	.	40	ASN	CB	C	13	38.72	0.0207	.	1	317	.	49	VAL	H	H	1	8.235	0.0037	.	1
259	.	40	ASN	N	N	15	120.1	0.0106	.	1	318	.	49	VAL	C	C	13	174.4	0.1000	.	1
260	.	41	SER	H	H	1	8.468	0.0020	.	1	319	.	49	VAL	CA	C	13	59.83	0.0229	.	1
261	.	41	SER	HA	H	1	4.345	0.0050	.	1	320	.	49	VAL	CB	C	13	32.49	0.1000	.	1
262	.	41	SER	QB	H	1	3.908	0.0050	.	2	321	.	49	VAL	N	N	15	122.1	0.0634	.	1
263	.	41	SER	C	C	13	175.0	0.0119	.	1	322	.	50	PRO	HA	H	1	4.376	0.0050	.	1
264	.	41	SER	CA	C	13	59.37	0.0267	.	1	323	.	50	PRO	HB2	H	1	1.861	0.0050	.	2
265	.	41	SER	CB	C	13	63.68	0.1091	.	1	324	.	50	PRO	HB3	H	1	2.290	0.0050	.	2
266	.	41	SER	N	N	15	116.6	0.0102	.	1	325	.	50	PRO	C	C	13	176.3	0.1000	.	1
267	.	42	GLU	H	H	1	8.554	0.0024	.	1	326	.	50	PRO	CA	C	13	63.26	0.0369	.	1
268	.	42	GLU	HA	H	1	4.330	0.0050	.	1	327	.	50	PRO	CB	C	13	32.25	0.0330	.	1
269	.	42	GLU	QB	H	1	2.048	0.0050	.	2	328	.	50	PRO	CG	C	13	27.33	0.1000	.	1
270	.	42	GLU	C	C	13	176.9	0.0172	.	1	329	.	50	PRO	CD	C	13	51.02	0.1000	.	1
271	.	42	GLU	CA	C	13	56.80	0.0425	.	1	330	.	51	ASP	H	H	1	8.552	0.0025	.	1
272	.	42	GLU	CB	C	13	29.85	0.0547	.	1	331	.	51	ASP	HA	H	1	4.581	0.0050	.	1
273	.	42	GLU	CG	C	13	35.98	0.0000	.	1	332	.	51	ASP	HB2	H	1	2.585	0.0050	.	2
274	.	42	GLU	N	N	15	122.3	0.0187	.	1	333	.	51	ASP	HB3	H	1	2.768	0.0050	.	2
275	.	43	SER	H	H	1	8.408	0.0019	.	1	334	.	51	ASP	C	C	13	176.3	0.0085	.	1
276	.	43	SER	HA	H	1	4.501	0.0050	.	1	335	.	51	ASP	CA	C	13	54.05	0.0330	.	1
277	.	43	SER	QB	H	1	3.910	0.0050	.	2	336	.	51	ASP	CB	C	13	40.96	0.0337	.	1

337	.	51	ASP	N	N	15	121.1	0.0462	.	1	396	.	60	PRO	CG	C	13	27.34	0.1000	.	1
338	.	52	LEU	H	H	1	8.423	0.0021	.	1	397	.	60	PRO	CD	C	13	51.15	0.1000	.	1
339	.	52	LEU	HA	H	1	4.374	0.0050	.	1	398	.	61	SER	H	H	1	8.594	0.0020	.	1
340	.	52	LEU	HB	H	1	1.643	0.0050	.	2	399	.	61	SER	HA	H	1	4.411	0.0050	.	1
341	.	52	LEU	C	C	13	177.6	0.0079	.	1	400	.	61	SER	QB	H	1	3.864	0.0050	.	2
342	.	52	LEU	CA	C	13	55.08	0.0420	.	1	401	.	61	SER	C	C	13	174.5	0.1864	.	1
343	.	52	LEU	CB	C	13	42.01	0.0198	.	1	402	.	61	SER	CA	C	13	58.24	0.0574	.	1
344	.	52	LEU	CG	C	13	26.88	0.1000	.	1	403	.	61	SER	CB	C	13	63.83	0.1302	.	1
345	.	52	LEU	CD1	C	13	23.44	0.1000	.	2	404	.	61	SER	N	N	15	117.0	0.1023	.	1
346	.	52	LEU	CD2	C	13	24.70	0.1000	.	2	405	.	62	LYS	H	H	1	8.571	0.0029	.	1
347	.	52	LEU	N	N	15	123.9	0.0184	.	1	406	.	62	LYS	HA	H	1	4.341	0.0050	.	1
348	.	53	ARG	H	H	1	8.506	0.0023	.	1	407	.	62	LYS	QB	H	1	1.790	0.0050	.	2
349	.	53	ARG	HA	H	1	4.378	0.0050	.	1	408	.	62	LYS	C	C	13	176.5	0.0045	.	1
350	.	53	ARG	QB	H	1	1.888	0.0050	.	2	409	.	62	LYS	CA	C	13	56.19	0.0597	.	1
351	.	53	ARG	C	C	13	177.0	0.0041	.	1	410	.	62	LYS	CB	C	13	33.10	0.0091	.	1
352	.	53	ARG	CA	C	13	56.37	0.0459	.	1	411	.	62	LYS	CG	C	13	24.73	0.1000	.	1
353	.	53	ARG	CB	C	13	30.65	0.0207	.	1	412	.	62	LYS	CD	C	13	29.09	0.1000	.	1
354	.	53	ARG	CG	C	13	26.91	0.1000	.	1	413	.	62	LYS	N	N	15	124.1	0.0311	.	1
355	.	53	ARG	CD	C	13	43.41	0.1000	.	1	414	.	63	LYS	H	H	1	8.515	0.0016	.	1
356	.	53	ARG	N	N	15	121.9	0.0540	.	1	415	.	63	LYS	HA	H	1	4.310	0.0050	.	1
357	.	54	THR	H	H	1	8.349	0.0013	.	1	416	.	63	LYS	QB	H	1	1.781	0.0050	.	2
358	.	54	THR	HA	H	1	4.373	0.0050	.	1	417	.	63	LYS	C	C	13	176.6	0.0078	.	1
359	.	54	THR	HB	H	1	4.299	0.0050	.	1	418	.	63	LYS	CA	C	13	56.18	0.0883	.	1
360	.	54	THR	C	C	13	174.4	0.0202	.	1	419	.	63	LYS	CB	C	13	33.04	0.0464	.	1
361	.	54	THR	CA	C	13	61.77	0.0458	.	1	420	.	63	LYS	CG	C	13	24.71	0.1000	.	1
362	.	54	THR	CB	C	13	69.65	0.0856	.	1	421	.	63	LYS	N	N	15	123.7	0.0188	.	1
363	.	54	THR	CG2	C	13	21.89	0.1000	.	1	422	.	64	ARG	H	H	1	8.661	0.0044	.	1
364	.	54	THR	N	N	15	114.4	0.0291	.	1	423	.	64	ARG	HA	H	1	4.381	0.0050	.	1
365	.	55	ASP	H	H	1	8.496	0.0020	.	1	424	.	64	ARG	QB	H	1	1.834	0.0050	.	2
366	.	55	ASP	HA	H	1	4.645	0.0050	.	1	425	.	64	ARG	C	C	13	176.3	0.0140	.	1
367	.	55	ASP	QB	H	1	2.675	0.0050	.	2	426	.	64	ARG	CA	C	13	55.99	0.0554	.	1
368	.	55	ASP	C	C	13	175.8	0.0253	.	1	427	.	64	ARG	CB	C	13	31.09	0.0166	.	1
369	.	55	ASP	CA	C	13	54.08	0.0796	.	1	428	.	64	ARG	CG	C	13	27.06	0.1000	.	1
370	.	55	ASP	CB	C	13	40.78	0.0335	.	1	429	.	64	ARG	CD	C	13	43.41	0.1000	.	1
371	.	55	ASP	N	N	15	122.3	0.0082	.	1	430	.	64	ARG	N	N	15	123.9	0.0844	.	1
372	.	56	GLN	H	H	1	8.302	0.0023	.	1	431	.	65	SER	H	H	1	8.626	0.0013	.	1
373	.	56	GLN	C	C	13	173.8	0.1000	.	1	432	.	65	SER	HA	H	1	4.477	0.0050	.	1
374	.	56	GLN	CA	C	13	53.54	0.0233	.	1	433	.	65	SER	QB	H	1	3.881	0.0050	.	2
375	.	56	GLN	CB	C	13	28.90	0.1000	.	1	434	.	65	SER	C	C	13	174.5	0.0194	.	1
376	.	56	GLN	N	N	15	121.3	0.0154	.	1	435	.	65	SER	CA	C	13	58.36	0.0406	.	1
377	.	57	PRO	HA	H	1	4.403	0.0050	.	1	436	.	65	SER	CB	C	13	63.73	0.1115	.	1
378	.	57	PRO	HB2	H	1	1.862	0.0050	.	2	437	.	65	SER	N	N	15	117.9	0.0125	.	1
379	.	57	PRO	HB3	H	1	2.293	0.0050	.	2	438	.	66	CYS	H	H	1	8.651	0.0034	.	1
380	.	57	PRO	C	C	13	176.6	0.0680	.	1	439	.	66	CYS	HA	H	1	4.746	0.0050	.	1
381	.	57	PRO	CA	C	13	62.95	0.0558	.	1	440	.	66	CYS	HB	H	1	3.238	0.0050	.	2
382	.	57	PRO	CB	C	13	32.16	0.0399	.	1	441	.	66	CYS	C	C	13	174.0	0.0092	.	1
383	.	57	PRO	CG	C	13	27.40	0.0000	.	1	442	.	66	CYS	CA	C	13	55.12	0.1484	.	1
384	.	57	PRO	CD	C	13	50.75	0.0000	.	1	443	.	66	CYS	CB	C	13	41.00	0.0170	.	1
385	.	58	LYS	H	H	1	8.638	0.0017	.	1	444	.	66	CYS	N	N	15	120.9	0.0330	.	1
386	.	58	LYS	C	C	13	174.3	0.0000	.	1	445	.	67	ASP	H	H	1	8.662	0.0045	.	1
387	.	58	LYS	CA	C	13	54.11	0.0186	.	1	446	.	67	ASP	C	C	13	174.8	0.1000	.	1
388	.	58	LYS	CB	C	13	30.16	0.0000	.	1	447	.	67	ASP	CA	C	13	52.14	0.0477	.	1
389	.	58	LYS	N	N	15	123.7	0.0361	.	1	448	.	67	ASP	N	N	15	124.7	0.0089	.	1
390	.	60	PRO	HA	H	1	4.445	0.0050	.	1	449	.	68	PRO	HA	H	1	4.354	0.0050	.	1
391	.	60	PRO	HB2	H	1	1.919	0.0050	.	2	450	.	68	PRO	HB2	H	1	1.996	0.0050	.	2
392	.	60	PRO	HB3	H	1	2.336	0.0050	.	2	451	.	68	PRO	HB3	H	1	2.338	0.0000	.	2
393	.	60	PRO	C	C	13	177.0	0.0095	.	1	452	.	68	PRO	C	C	13	177.5	0.0205	.	1
394	.	60	PRO	CA	C	13	62.90	0.1643	.	1	453	.	68	PRO	CA	C	13	63.97	0.0684	.	1
395	.	60	PRO	CB	C	13	32.16	0.0710	.	1	454	.	68	PRO	CB	C	13	32.12	0.0156	.	1

455	.	68	PRO	CG	C	13	27.21	0.0000	.	1	514	.	76	LEU	C	C	13	177.4	0.1720	.	1
456	.	68	PRO	CD	C	13	51.03	0.0000	.	1	515	.	76	LEU	CA	C	13	55.22	0.1351	.	1
457	.	69	SER	H	H	1	8.572	0.0113	.	1	516	.	76	LEU	CB	C	13	42.10	0.0321	.	1
458	.	69	SER	HA	H	1	4.314	0.0050	.	1	517	.	76	LEU	CG	C	13	26.83	0.1000	.	1
459	.	69	SER	QB	H	1	3.915	0.0050	.	2	518	.	76	LEU	CD1	C	13	23.86	0.1000	.	2
460	.	69	SER	C	C	13	174.9	0.1075	.	1	519	.	76	LEU	CD2	C	13	24.66	0.1000	.	2
461	.	69	SER	CA	C	13	59.50	0.0586	.	1	520	.	76	LEU	N	N	15	123.3	0.0296	.	1
462	.	69	SER	CB	C	13	63.39	0.1134	.	1	521	.	77	LYS	H	H	1	8.435	0.0163	.	1
463	.	69	SER	N	N	15	115.4	0.0166	.	1	522	.	77	LYS	HA	H	1	4.292	0.0050	.	1
464	.	70	GLU	H	H	1	8.208	0.0046	.	1	523	.	77	LYS	QB	H	1	1.822	0.0050	.	2
465	.	70	GLU	HA	H	1	4.251	0.0050	.	1	524	.	77	LYS	C	C	13	176.7	0.2138	.	1
466	.	70	GLU	QB	H	1	1.971	0.0050	.	2	525	.	77	LYS	CA	C	13	56.17	0.2428	.	1
467	.	70	GLU	C	C	13	176.2	0.0612	.	1	526	.	77	LYS	CB	C	13	33.00	0.0971	.	1
468	.	70	GLU	CA	C	13	56.64	0.0755	.	1	527	.	77	LYS	CG	C	13	24.88	0.1000	.	1
469	.	70	GLU	CB	C	13	30.10	0.0112	.	1	528	.	77	LYS	CD	C	13	28.95	0.1000	.	1
470	.	70	GLU	CG	C	13	35.88	0.0000	.	1	529	.	77	LYS	N	N	15	122.9	0.2384	.	1
471	.	70	GLU	N	N	15	121.9	0.0459	.	1	530	.	78	GLU	H	H	1	8.578	0.0525	.	1
472	.	71	TYR	H	H	1	8.068	0.0030	.	1	531	.	78	GLU	HA	H	1	4.287	0.0050	.	1
473	.	71	TYR	HA	H	1	4.556	0.0050	.	1	532	.	78	GLU	QB	H	1	2.015	0.0050	.	2
474	.	71	TYR	QB	H	1	2.994	0.0050	.	.	533	.	78	GLU	C	C	13	176.6	0.0184	.	1
475	.	71	TYR	C	C	13	175.6	0.0170	.	1	534	.	78	GLU	CA	C	13	56.66	0.1624	.	1
476	.	71	TYR	CA	C	13	57.96	0.0795	.	1	535	.	78	GLU	CB	C	13	30.15	0.0501	.	1
477	.	71	TYR	CB	C	13	38.57	0.0794	.	1	536	.	78	GLU	CG	C	13	36.01	0.1000	.	1
478	.	71	TYR	N	N	15	120.9	0.0134	.	1	537	.	78	GLU	N	N	15	122.5	0.5467	.	1
479	.	72	ARG	H	H	1	8.141	0.0026	.	1	538	.	79	SER	H	H	1	8.454	0.0028	.	1
480	.	72	ARG	HA	H	1	4.342	0.0050	.	1	539	.	79	SER	HA	H	1	4.431	0.0050	.	1
481	.	72	ARG	QB	H	1	1.743	0.0050	.	2	540	.	79	SER	QB	H	1	3.878	0.0050	.	2
482	.	72	ARG	C	C	13	176.0	0.0612	.	1	541	.	79	SER	C	C	13	174.5	0.0059	.	1
483	.	72	ARG	CA	C	13	55.70	0.0867	.	1	542	.	79	SER	CA	C	13	58.49	0.1095	.	1
484	.	72	ARG	CB	C	13	31.09	0.0175	.	1	543	.	79	SER	CB	C	13	63.66	0.0887	.	1
485	.	72	ARG	CG	C	13	27.02	0.1000	.	1	544	.	79	SER	N	N	15	117.1	0.0156	.	1
486	.	72	ARG	CD	C	13	43.19	0.1000	.	1	545	.	80	LEU	H	H	1	8.359	0.0019	.	1
487	.	72	ARG	N	N	15	123.8	0.0565	.	1	546	.	80	LEU	HA	H	1	4.357	0.0050	.	1
488	.	73	VAL	H	H	1	8.300	0.0036	.	1	547	.	80	LEU	HB	H	1	1.641	0.0050	.	2
489	.	73	VAL	HA	H	1	4.039	0.0050	.	1	548	.	80	LEU	C	C	13	177.4	0.0099	.	1
490	.	73	VAL	HB	H	1	2.084	0.0050	.	1	549	.	80	LEU	CA	C	13	55.14	0.1074	.	1
491	.	73	VAL	C	C	13	176.4	0.0102	.	1	550	.	80	LEU	CB	C	13	42.15	0.0585	.	1
492	.	73	VAL	CA	C	13	62.72	0.0691	.	1	551	.	80	LEU	CG	C	13	26.94	0.1000	.	1
493	.	73	VAL	CB	C	13	32.55	0.0647	.	1	552	.	80	LEU	CD1	C	13	23.23	0.1000	.	2
494	.	73	VAL	CQG	C	13	20.99	0.1000	.	2	553	.	80	LEU	CD2	C	13	24.88	0.1000	.	2
495	.	73	VAL	N	N	15	122.0	0.0168	.	1	554	.	80	LEU	N	N	15	124.4	0.0182	.	1
496	.	74	SER	H	H	1	8.469	0.0021	.	1	555	.	81	ILE	H	H	1	8.239	0.0050	.	1
497	.	74	SER	HA	H	1	4.418	0.0050	.	1	556	.	81	ILE	HA	H	1	4.253	0.0050	.	1
498	.	74	SER	QB	H	1	3.868	0.0050	.	2	557	.	81	ILE	HB	H	1	1.897	0.0050	.	1
499	.	74	SER	C	C	13	174.6	0.0167	.	1	558	.	81	ILE	C	C	13	176.6	0.0280	.	1
500	.	74	SER	CA	C	13	58.55	0.1853	.	1	559	.	81	ILE	CA	C	13	61.10	0.0845	.	1
501	.	74	SER	CB	C	13	63.70	0.0863	.	1	560	.	81	ILE	CB	C	13	38.42	0.1427	.	1
502	.	74	SER	N	N	15	119.3	0.0236	.	1	561	.	81	ILE	CG1	C	13	27.33	0.1000	.	1
503	.	75	GLU	H	H	1	8.553	0.0066	.	1	562	.	81	ILE	CG2	C	13	17.56	0.1000	.	1
504	.	75	GLU	HA	H	1	4.302	0.0050	.	1	563	.	81	ILE	CD1	C	13	12.84	0.1000	.	1
505	.	75	GLU	QB	H	1	1.993	0.0050	.	2	564	.	81	ILE	N	N	15	122.1	0.1606	.	1
506	.	75	GLU	C	C	13	176.5	0.1000	.	1	565	.	82	THR	H	H	1	8.411	0.0033	.	1
507	.	75	GLU	CA	C	13	56.38	0.1748	.	1	566	.	82	THR	HA	H	1	4.461	0.0050	.	1
508	.	75	GLU	CB	C	13	30.06	0.0020	.	1	567	.	82	THR	HB	H	1	4.256	0.0050	.	1
509	.	75	GLU	CG	C	13	35.94	0.1000	.	1	568	.	82	THR	C	C	13	174.6	0.0098	.	1
510	.	75	GLU	N	N	15	123.4	0.0312	.	1	569	.	82	THR	CA	C	13	61.57	0.0409	.	1
511	.	76	LEU	H	H	1	8.319	0.0028	.	1	570	.	82	THR	CB	C	13	69.74	0.0937	.	1
512	.	76	LEU	HA	H	1	4.321	0.0050	.	1	571	.	82	THR	CG2	C	13	21.79	0.1000	.	1
513	.	76	LEU	HB	H	1	1.631	0.0050	.	2	572	.	82	THR	N	N	15	119.1	0.1148	.	1

```
573 . 83 THR H  H  1 8.308  0.0020 . 1
574 . 83 THR HA H  1 4.349  0.0050 . 1
575 . 83 THR HB H  1 4.267  0.0050 . 1
576 . 83 THR C  C  13 177.5 0.0070 . 1
577 . 83 THR CA C  13 61.56 0.1810 . 1
578 . 83 THR CB C  13 69.83 0.0387 . 1
579 . 83 THR CG2 C 13 21.88 0.1000 . 1
580 . 83 THR N  N  15 116.9 0.0303 . 1
581 . 84 THR H  H  1 8.408  0.0028 . 1
582 . 84 THR CA C  13 59.88 0.0399 . 1
583 . 84 THR CB C  13 69.62 0.1000 . 1
584 . 84 THR N  N  15 119.9 0.0313 . 1
585 . 85 PRO HA H  1 4.458  0.0050 . 1
586 . 85 PRO HB2 H 1 1.930  0.0050 . 2
587 . 85 PRO HB3 H 1 2.333  0.0050 . 2
588 . 85 PRO C  C  13 177.0 0.0045 . 1
589 . 85 PRO CA C  13 62.93 0.1000 . 1
590 . 85 PRO CB C  13 32.13 0.0765 . 1
591 . 85 PRO CG C  13 27.38 0.1000 . 1
592 . 85 PRO CD C  13 50.68 0.1000 . 1
593 . 86 SER H  H  1 8.581  0.0151 . 1
594 . 86 SER HA H  1 4.410  0.0050 . 1
595 . 86 SER QB H  1 3.859  0.0050 . 2
596 . 86 SER C  C  13 174.4 0.2043 . 1
597 . 86 SER CA C  13 58.36 0.0611 . 1
598 . 86 SER CB C  13 63.67 0.0735 . 1
599 . 86 SER N  N  15 116.7 0.2466 . 1
600 . 87 ARG H  H  1 8.482  0.0025 . 1
601 . 87 ARG C  C  13 174.1 0.1000 . 1
602 . 87 ARG CA C  13 53.89 0.0315 . 1
603 . 87 ARG CB C  13 30.22 0.1000 . 1
604 . 87 ARG N  N  15 124.2 0.0243 . 1
605 . 88 PRO HA H  1 4.437  0.0050 . 1
606 . 88 PRO HB2 H 1 1.904  0.0050 . 2
607 . 88 PRO HB3 H 1 2.318  0.0050 . 2
608 . 88 PRO C  C  13 176.9 0.0035 . 1
609 . 88 PRO CA C  13 63.06 0.0442 . 1
610 . 88 PRO CB C  13 32.18 0.0496 . 1
611 . 88 PRO CG C  13 27.43 0.0000 . 1
612 . 88 PRO CD C  13 50.78 0.1000 . 1
613 . 89 ARG H  H  1 8.742  0.0017 . 1
614 . 89 ARG HA H  1 4.368  0.0050 . 1
615 . 89 ARG QB H  1 1.845  0.0050 . 2
616 . 89 ARG C  C  13 176.7 0.0181 . 1
617 . 89 ARG CA C  13 56.24 0.0312 . 1
618 . 89 ARG CB C  13 30.75 0.0439 . 1
619 . 89 ARG CG C  13 27.09 0.1000 . 1
620 . 89 ARG CD C  13 43.36 0.1000 . 1
621 . 89 ARG N  N  15 122.3 0.0246 . 1
622 . 90 THR H  H  1 8.338  0.0030 . 1
623 . 90 THR HA H  1 4.310  0.0050 . 1
624 . 90 THR HB H  1 4.254  0.0050 . 1
625 . 90 THR C  C  13 174.2 0.0117 . 1
626 . 90 THR CA C  13 61.71 0.0636 . 1
627 . 90 THR CB C  13 69.90 0.1158 . 1
628 . 90 THR CG2 C 13 22.00 0.0880 . 1
629 . 90 THR N  N  15 116.1 0.0134 . 1
630 . 91 ALA H  H  1 8.498  0.0112 . 1
631 . 91 ALA HA H  1 4.349  0.0050 . 1
632 . 91 ALA MB H  1 1.386  0.0050 . 1
633 . 91 ALA C  C  13 177.6 0.0243 . 1
634 . 91 ALA CA C  13 52.34 0.0395 . 1
635 . 91 ALA CB C  13 19.38 0.0892 . 1
636 . 91 ALA N  N  15 127.3 0.0182 . 1
637 . 92 LYS H  H  1 8.494  0.0034 . 1
638 . 92 LYS HA H  1 4.338  0.0050 . 1
639 . 92 LYS QB H  1 1.785  0.0050 . 2
640 . 92 LYS C  C  13 176.6 0.0050 . 1
641 . 92 LYS CA C  13 56.33 0.0258 . 1
642 . 92 LYS CB C  13 33.01 0.0143 . 1
643 . 92 LYS CG C  13 24.83 0.0000 . 1
644 . 92 LYS CD C  13 29.17 0.0000 . 1
645 . 92 LYS N  N  15 121.8 0.0419 . 1
646 . 93 ARG H  H  1 8.549  0.0211 . 1
647 . 93 ARG HA H  1 4.291  0.0050 . 1
648 . 93 ARG QB H  1 1.774  0.0050 . 2
649 . 93 ARG C  C  13 176.1 0.0152 . 1
650 . 93 ARG CA C  13 56.07 0.0278 . 1
651 . 93 ARG CB C  13 30.92 0.0119 . 1
652 . 93 ARG CG C  13 27.16 0.1000 . 1
653 . 93 ARG N  N  15 123.4 0.0160 . 1
654 . 94 ARG H  H  1 8.647  0.0047 . 1
655 . 94 ARG HA H  1 4.344  0.0050 . 1
656 . 94 ARG QB H  1 1.789  0.0050 . 2
657 . 94 ARG C  C  13 176.1 0.0120 . 1
658 . 94 ARG CA C  13 56.00 0.0710 . 1
659 . 94 ARG CB C  13 30.76 0.0461 . 1
660 . 94 ARG CG C  13 27.13 0.1000 . 1
661 . 94 ARG CD C  13 43.43 0.1000 . 1
662 . 94 ARG N  N  15 124.1 0.0789 . 1
663 . 95 ILE H  H  1 8.452  0.0034 . 1
664 . 95 ILE HA H  1 4.143  0.0050 . 1
665 . 95 ILE HB H  1 1.831  0.0050 . 1
666 . 95 ILE C  C  13 175.9 0.0116 . 1
667 . 95 ILE CA C  13 61.04 0.0734 . 1
668 . 95 ILE CB C  13 38.63 0.0800 . 1
669 . 95 ILE CG1 C 13 27.37 0.1000 . 1
670 . 95 ILE CG2 C 13 17.46 0.1000 . 1
671 . 95 ILE CD1 C 13 12.83 0.1000 . 1
672 . 95 ILE N  N  15 124.0 0.0249 . 1
673 . 96 ARG H  H  1 8.551  0.0022 . 1
674 . 96 ARG HA H  1 4.394  0.0050 . 1
675 . 96 ARG QB H  1 1.820  0.0050 . 2
676 . 96 ARG C  C  13 175.1 0.0182 . 1
678 . 96 ARG CA C  13 55.77 0.0359 . 1
679 . 96 ARG CB C  13 30.85 0.0387 . 1
680 . 96 ARG CG C  13 27.06 0.1000 . 1
681 . 96 ARG CD C  13 43.43 0.1000 . 1
682 . 96 ARG N  N  15 126.9 0.0397 . 1
683 . 97 LEU H  H  1 8.177  0.0039 . 1
684 . 97 LEU C  C  13 173.6 0.1000 . 1
685 . 97 LEU CA C  13 56.69 0.0288 . 1
686 . 97 LEU CB C  13 43.09 0.1000 . 1
687 . 97 LEU N  N  15 129.9 0.2  . 1

stop_
```