INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

Expression, purification and characterization of azurin derived peptides as target sequences for P-cadherin overexpressing breast cancer cells

Ana Teresa de Carvalho Estevens

Dissertação para a obtenção do Grau de Mestre em

Engenharia Biológica

Júri

Presidente:	Doutora Maria Raquel Murias dos Santos Aires Barros
Orientadores:	Doutor Arsénio do Carmo Sales Mendes Fialho
	Mestre Nuno Filipe Santos Bernardes
Vogal:	Doutor Gabriel António Amaro Monteiro

Outubro 2011

As our case is new, we must think and act anew.

Abraham Lincoln

Aknowledgments

First, I would like to thank my supervisor Professor Arsénio Fialho for all the help and guidance during the internship. To my supervisor Nuno Bernardes for all the lab work teaching and supervising, help during writing process, support and motivation that gave me throughout the last months.

I would like to acknowledge Dr. Ana Azevedo and Luís Burlido from BERG for their help in Circular Dichroism.

I would also like to thank Bruna Mota, my desk partner that helped me so many times. Dalila Mil-Homens, Eunice Penas and everyone in the BSRG lab for all the support, help and motivation.

Finally, a special thank for my family for their inconditional support. And to my friends, especially my course colleagues for the relaxing lunch times.

A special thank to João for always being there when I most needed.

Thank you!

Obrigado!

Abstract

Therapeutic peptides are a promising class of drugs that have many advantages over therapeutic proteins, antibodies and small organic molecules. Most of these peptides are designed based on truncate fragments of therapeutic proteins and antibodies. Nowadays, these peptides are used in many pathologies like diabetes, microbial infections and oncology. (Lu *et al.*, 2006; Vlieghe *et al.*, 2010)

Azurin is a low molecular weight protein member of the cupredoxin family and produced by *Pseudomonas aeruginosa*. Several studies (Chaudhari *et al.*, 2006; Chaudhari *et al.*, 2007; Yamada *et al.*, 2002b) demonstrated that azurin has anticancer, antiparasitic and antiviral activity and that azurin behaves like a scaffold protein, establishing high affinity interactions between different molecules using different regions (Fialho *et al.*, 2008). These characteristics make azurin a good source of therapeutic peptides for oncology and microbial infections.

In this work, a production strategy was developed to obtain three different azurin derived peptides: Azu 1-50, Azu 1-77 and Azu 80-128. The recent identification of P-cadherin as a possible target protein for azurin lead to the need of having a reproducible method of production of azurin derived peptides in order to better study azurin interaction with P-cadherin (Fialho, 2009). All three peptides were cloned into the pWH844 vector that encodes a poly histidine tag in the N-terminal of the target peptide. From the three constructions, Azu 1-50 and Azu 1-77 were successfully expressed in *E. coli* BL21(DE3) (4h incubation at 37°C and inducted with 0.2 mM IPTG) albeit in lower quantities that should be desired. Both peptides were purified in a Ni-NTA column and desalted through size exclusion chromatography using the ÄKTA PRIME apparatus.

During expression both peptides showed to form dimers when in higher concentrations and were obtained in very low quantity at the end of the purification procedure.

In conclusion, the adopted method of producing azurin derived peptide demonstrated to produce low quantities of peptides and difficult to upscale. Therefore different aspects of the purification method should be optimized in order to produce more peptides and to improve the purification yield.

Key words: Therapeutic peptides, azurin, azurin derived peptides, anticancer activity, P-cadherin, purification, poly histidine tag.

Resumo

Os péptidos terapêuticos são uma classe de terapêuticos com vantagens sobre proteínas, anticorpos e pequenas moléculas orgânicas. Muitos destes péptidos são desenhados a partir de truncados de proteínas e anticorpos com propriedades terapêuticas. Hoje em dia estes péptidos são usados em muitas patologias como diabetes, infecções microbianas e oncologia. (Lu *et al.*, 2006; Vlieghe *et al.*, 2010)

A azurina é uma proteína de baixo peso molecular membro da família das cupredoxinas, produzida por *Pseudomonas aeruginosa*. Diversos estudos (Chaudhari *et al.*, 2006; Chaudhari *et al.*, 2007; Yamada *et al.*, 2002b) demonstram que a azurina tem actividade anti-cancerígena, anti-parasítica e anti-viral. A azurina comporta-se como uma proteína "andaime", estabelecendo interacção de grande afinidade com diferentes moléculas usando diferentes regiões (Fialho *et al.*, 2008). Estas características fazem da azurina uma fonte promissora de péptidos terapêuticos com aplicações em oncologia e infecções microbianas.

Neste trabalho, foi desenvolvida uma estratégia de produção de três péptidos derivados da azurina: Azu 1-50, Azu 1-77 e Azu 80-128. Recentemente, a P-caderina foi identificada como possível alvo de acção da azurina, o que impulsiona a necessidade de desenvolver um método reprodutível de produção de péptidos derivados da azurina para melhor estudar a interacção entre a azurina e a P-caderina (Fialho, 2009). Os três péptidos foram clonados no vector pWH844 que codifica uma cauda de histidinas no N-terminal do péptido. Das três construções, conseguiram-se expressar as Azu 1-50 e Azu 1-77 em *E. coli* BL21(DE3) (4h de incubação a 37°C e indução com 0.2 mM IPTG). No entanto, os níveis de expressão obtidos para ambos os péptidos foram bastante baixos. Ambos os péptidos foram purificados numa coluna Ni-NTA e trocou-se o tampão por cromatografia de exclusão molecular usando o aparelho ÄKTA PRIME.

Durante a expressão, observou-se para ambos os péptidos a formação de dímeros, quando se apresentavam em maiores concentrações. A quantidade obtida após todo o procedimento de purificação foi bastante baixa.

Em conclusão, o método adoptado de produção e purificação de péptidos derivados da azurina é pouco eficiente em termos de quantidade de péptidos produzidos e é difícil de aumentar a sua escala de produção. Devido a estas observações, é necessário estudar vários aspectos do processo implementado de modo a optimizar a produção de péptidos e melhorar o rendimento de todo o processo.

Palavras chave: Péptidos terapêuticos, azurina, péptidos derivados da azurina, actividade anti-cancerígena, P-caderina, purificação, cauda de histidinas.

Table of Contents

List of tables

List of figures

List of abbreviations

APS – Ammonium persulfate

azu 1-50 – Azurin DNA sequence fragment encoding the azurin derived peptide corresponding to the fragment from the 1st to the 50th amino acid.

azu 1-77 – Azurin DNA sequence fragment encoding the azurin derived peptide corresponding to the fragment from the 1st to the 77th amino acid.

azu 80-128 – Azurin DNA sequence fragment encoding the azurin derived peptide corresponding to the fragment from the 80th to the 128th amino acid.

Azu 1-50 – Azurin derived peptide corresponding to the fragment from the 1st to the 50th amino acid.

Azu 1-77 – Azurin derived peptide corresponding to the fragment from the 1st to the 77th amino acid.

Azu 80-128 – Azurin derived peptide corresponding to the fragment from the 80th to the 128th amino acid.

BSA – Bovine serum albumin

CD – Circular dichroism

DNA – Deoxyribonucleic acid

HIV – Human immunodeficiency virus

IMAC – Immobilized metal affinity chromatography

IPTG – Isopropyl β-D-1-thiogalactopyranidose

LB – Luria Bertani

Ni-NTA – Nickel-nitriloacetic acid

NRMSD – Normalized root mean square deviation

OD – Optic density

PBS – Phosphate buffered saline

PCR – Polymerase chain reaction

SB – Super Broth

SDS – Sodium dodecyl sulfate

SDS-PAGE – Sodium dodecyl sulfate polyacrylamide gel electrophoresis

TAE – Tris-Acetate-EDTA

TCM – Tris $CaCl_2$ $MgCl_2$ solution

TEMED – Tetramethylethylenediamine

UV – Ultra violet

1 Introduction

Therapeutic peptides are a promising class of therapeutics that present many advantages over classical drugs like proteins, antibodies and small organic molecules. Therapeutic peptides have higher affinity to the target and lower toxicity profile than small organic peptides and can penetrate better due to its smaller size. Besides, because of their small size, peptides can be modified in order to increase its stability and potency. (Lu *et al.*, 2006; Vlieghe *et al.*, 2010)

Peptides have already shown to be useful in many pathologies like diabetes, infective diseases (bacterial, fungal and viral), oncology and osteoporosis (Vlieghe *et al.*, 2010). The majority of these peptides are derived from protein active sites although some are designed based on genetic, recombinant and chemical libraries (Duncan Patrick, 2008).

Azurin, a low molecular weight cupredoxin produced by *Pseudomonas aeruginosa* is a promising source of therapeutic peptides since it has anticancer, antiparasitic and anti-HIV properties associated with different regions of the protein (Fialho *et al.*, 2008). In fact, considering its anticancer activity, studies revealed that azurin preferentially enter breast cancer cells and induce apoptosis (Yamada *et al.*, 2002b), cell cycle arrest (Chaudhari *et al.*, 2007) and inhibits angiogenesis (Mehta *et al.*, 2011) through interaction with different molecules. In particular, azurin's p28 peptide, corresponding to aminoacids 50 to 77 which include the PTD (protein transduction domain) responsible for cell entry, is able to induce apoptosis and inhibit angiogenesis and recently ended Phase I clinical trials (CDG, 2011; Mehta *et al.*, 2011; Yamada *et al.*, 2009). Another azurin derived peptide, p26, corresponding to aminoacid 88 to 113, was identified as being structurally similar to ephrinB2, the EphB2 ligand. p26 is able to competitively bind to EphB2 inhibiting cellular signaling pathways that ultimately contribute for cancer growth (Chaudhari *et al.*, 2007).

Recently, P-cadherin, which is associated with breast cancer patients with poor prognosis (Paredes *et al.*, 2007; Ribeiro *et al.*, 2010), was identified as a new target molecule for azurin (Fialho, 2009).The aim of this work is to develop a reproducible method of production of azurin derived peptides in order to identify and study the responsible region for azurin's anti P-cadherin activity. In previous works azurin and some derived peptides have been expressed and purified using a GST-based strategy (Chaudhari *et al.*, 2007; Yamada *et al.*, 2005). In this work, we aim to develop a poly-histidine based expression and purification method to produce three azurin derived peptides, Azu 1-50, Azu 1-77 and Azu 80-128.

2 Literature overview

2.1 Therapeutic peptides

2.1.1 Overview

Various peptides are nowadays used as therapeutic drugs (Lien and Lowman, 2003). The first commercial therapeutic peptide made available was Lypressin, a vasopressin analogue, developed by Novartis in the 1970's (Pichereau and Allary, 2005). The so-called therapeutic peptides have three main origins, they are isolated from nature (naturally occurring peptides or fragments of larger proteins) or developed based on chemical or genetic/recombinant libraries (Duncan Patrick, 2008; Sato *et al.*, 2006).

Traditionally, therapeutic peptides are obtained from natural sources and clearly show some advantages over protein drugs such as antibodies (better tissue penetration due to their smaller size) and small molecules (higher affinity/specificity with the target and lower toxicity profile). However, some disadvantages such as low *in vivo* stability, short half-life, lower potency comparing to antibodies and possible occurrence of immunogenic sequences have drawn back peptide industry (Sato *et al.*, 2006).

Because of their length, peptides can be produced through chemical synthesis, recombinant DNA technology, cell-free expression systems, transgenic animals and plants or enzymatic synthesis. With chemical synthesis it is possible to use unnatural aminoacids and conjugation of small molecules, which diversifies the produced peptides and increase their stability. (Vlieghe *et al.*, 2010)

2.1.2 Advantages over other drugs

As referred before, therapeutic peptides have several advantages over protein drugs, antibodies and small molecules.

Comparing proteins and antibodies to peptide drugs, peptide drugs are more advantageous for several reasons: the complex conformation of proteins and antibodies difficult its production and storage, whereas peptides don't prompt conformational issues, are more stable and have lower manufacturing costs (synthetic versus recombinant production); the risk of immunogenicity posed by whole proteins and antibodies is very reduced in peptides sequences obtained from these proteins; because of their bigger size, proteins and antibodies have more restraints moving to its target; because of their smaller size, peptides have higher activity per mass than proteins and antibodies; because of peptides' simpler intellectual property landscape, they have lower royalty costs (Lu *et al.*, 2006; Vlieghe *et al.*, 2010). Therefore, using short peptides

derived from the interacting regions of proteins and antibodies poses a great advantage over using the whole protein (Lu *et al.*, 2006).

Compared to small organic molecules, therapeutic peptides also show several advantages. Because most therapeutic peptides represent small portions of functional regions of a protein, they are more efficient, selective and specific than organic molecules. The shorter half-life of peptides reduces the risk of accumulation in tissues. Besides, degradation of peptides generates aminoacids that do not pose a risk of systemic toxicity, whereas degradation of small organic compounds may lead to the formation of toxic substances. (Vlieghe *et al.*, 2010)

Nowadays most therapeutic peptides are chemically synthesized using the solid-phase peptide synthesis (SPPS) technique. Peptides produced through this technique often show higher purity comparing to recombinant techniques. SPPS also enables the use of unnatural aminoacids and C-terminal amidation that confer higher stability to the peptide. (Vlieghe *et al.*, 2010)

Several strategies can be used in order to increase peptide stability. The most used modifications are the addition of polyethylene glycol (PEG) to the N or C-terminal, peptide cyclization, use of D-aminoacids and unnatural aminoacids, N-terminal modifications (acetylation and glycosylation), C-terminal modifications (amidation) and fusion to Fc domain of human gamma immunoglobulin (IgG). (Sato *et al.*, 2006)

2.1.3 Main applications

Peptides are used in many pathologies like allergy, asthma, arthritis, cardiovascular diseases, diabetes, gastrointestinal dysfunction, immunity diseases, infective diseases (bacterial, fungal and viral), obesity, oncology and osteoporosis (Vlieghe *et al.*, 2010). Peptides are used to enhance cellular uptake, drug targeting, immunostimulation and, in some cases, act as a drug by itself (Khandare and Minko, 2006; Lien and Lowman, 2003).

Among the referred pathologies, oncology is one of the most studied. In cancer therapy, surgery and chemotherapy are efficient in removing localized tumors but often fail to treat metastatic and spreading tumors. Because of its ability to target cancer cells, the use of peptides is a promising strategy for cancer therapy. Peptides can be used as tumor markers, immunostimulators, anticancer agents and drug targeting vehicles. (Khandare and Minko, 2006)

Tumors associated with the endocrine system (adrenal gland, thyroid gland, pancreas and others) and nervous system usually overexpress somatostatin receptor at the cell surface. Somatostatin is a naturally occurring peptide hormone. The use of OctreoScan, a somatostatin analog peptide octreotide conjugated with a radioactive marker (Indium-111), increases the sensitivity of tumor imaging and detection. (Khandare and Minko, 2006)

Cancer vaccination is often used as an indirect tumor targeting. Using cancer specific peptides as tumor antigens is a safer method of immunization than the use of tumor cells (Khandare and

Minko, 2006). The bi-specific construction ULBP2-BB4 was designed for activation of natural killer (NK) cells against malignant cells of multiple myeloma. The ULBP2 part is a designed ligand for NKG2D, a cytoxicity receptor of NK cells that induce cell targeted lysis through binding to specific ligands naturally expressed in cancer cells. The BB4 part is an antibody fragment that recognizes CD138, a tumor antigen frequently overexpressed in multiple myeloma cells and absent in other hematopoietic cells that is involved in cell adhesion, maturation and proliferation. The bispecificity of ULBP2-BB4 peptide bind both activate NK cells and exerts antitumoral effects on plasma malignant cells. (von Strandmann *et al.*, 2006)

Concerning drug targeting, peptides show many advantages comparing to other classical vehicles. Besides, it is possible to develop constructions that can simultaneously act as a targeting vehicle and as an anticancer drug. An example is the Ant-BH3 construction. The Ant part corresponds to an antennapedia internalization sequence that act as a high soluble targeting vehicle for cancer cells. The BH3 corresponds to the minimal sequence of BCL-2 homology 3 (BH3). The BH3 domain is important for the pro-apoptotic activity of some proteins of the BCL-2 family, including Bak and Bax. The conjugation of Ant with BH3 allies the advantage of the solubility of Ant with the pro-apoptotic effect of BH3. (Khandare and Minko, 2006)

In Table 1 are presented some peptides approved by FDA and their indications.

Table 1 – Examples of FDA approved peptides, their originator, indications and approval date. Adapted from (FDA, 2011; informa-healthcare, 2007).

Generic name	Trade name	Originator	Indication	Approval date
Desmopressin	DDAVP	Ferring	Enuresis	December 1990
Indium-111 pentetreotide	OctreoScan	Mallinkrodt Medical	Neuroendocrine tumor scintigraphy	June 1994
Cyclosporin	Neoral	Novartis	General transplant rejection	July 1995
Calcitonin	Miacalcin	Novartis	Osteoporosis	August 1995
Goserelin	Zoladex	AstraZeneca	Prostate cancer	January 1996
Leuprolide acetate	Lupron	QLT	Prostate cancer	July 1997
Nesiritide citrate	Natrecor	Johnson and Johnson's	Heart failure	April 1998
Eptifibatide	Integrilin	Millenium	Unstable angina	May 1998
Glucagon	GlucaGen	Lilly	General diabetes	September 1998
Octreotide	Sandostatin	Novartis	Acromegaly	December 1998

Generic name	Trade name	Originator	Indication	Approval date
Bivalirudin	Angiomax	The Medicines Company	Thrombosis	December 2000
Teriparatide	Forsteo	Lilly	Osteoporosis	November 2002
Enfuvirtide	Fuzeon	Trimeris	HIV/AIDS infection	March 2003
Leuprolelin	Leuplin	Takeda	Prostate cancer	September 2005
Degarelix acetate	Firmagon	Ferring	Prostate cancer	December 2008

2.1.4 Expression of recombinant therapeutic peptides

With the development of SPPS, chemical synthesis of therapeutic peptides is nowadays a viable technique for large scale production of therapeutic peptides. Besides it allows different modifications in the peptides that increase their *in vivo* stability (Vlieghe *et al.*, 2010). However, recombinant techniques are still very used in therapeutic peptides production. Recombinant therapeutic proteins and peptides have been successfully produced at large scale since the 1980's and continue to be produced through this technique. (Graumann and Premstaller, 2006)

Although eukaryotes and yeasts are very efficient in protein expression and are able to perform posttranslational modifications, when such modifications are not needed, *Escherichia coli* is the more attractive host for recombinant protein and peptide expression. Its genome simplicity and availability of commercial strains specialized in protein expression make *E. coli* a good expression host in almost cases (Graumann and Premstaller, 2006). Ideally, heterologous proteins and peptides are intended to be expressed in soluble form in order to facilitate its recovery and purification. However, many heterologous proteins are expressed in insoluble forms (Sorensen and Mortensen, 2005). Regarding this problem, scientists often apply one of two strategies: enhance solubilization of the expressed product or enhance the inclusion body formation of insoluble expressed product.

In order to enhance the solubility of the expressed peptide, several parameters can be optimized with or without engineering the target peptide. Without modifying the expressed peptide, variables like temperature, host strain, culture media, transcription rates, inclusion of molecular chaperones and tRNA complementation plasmid can be tested. If the peptide can be modified, then the fusion of affinity tags and screening of insoluble fragments and their consequent removal or modification are two hypotheses that can be adopted. (Sorensen and Mortensen, 2005)

Recently, Gauci *et al.* optimized the production of antigen fragments from Taeniid Cestode parasites in *E. coli*. Because classic optimization strategies involving culture conditions did not improve expression yield, they tested the expression of modified peptides. They expressed the TSOL18 (from *Taenia*) and EG95 (from *Echinococcus granulosus*) antigens fused with a

Glutathione *S*-transferase (GST) affinity tail and peptides lacking the N, the C and the N and C-terminal of both antigens. In the case of TSOL18, the N-terminal encodes a hydrophobic secretory signal sequence. The N-terminal of EG95 also encodes a hydrophobic secretory signal sequence and the C-terminal encodes a transmembrane domain. Comparing the expression of the full length peptides to the N-terminal truncates, in both cases there was an increase in expression levels, particularly, there was an increase of soluble peptides fraction. Comparing the expression of the C and C and N-terminal truncate of EG95 to the N-terminal truncate of EG95, the doubled truncate showed higher levels of expression and higher soluble fraction. Sheep vaccination with the peptides and truncates showed that truncates maintained the whole peptide function, since the active site (FNIII domain) remained intact in all cases. (Gauci *et al.*, 2011)

However, some peptides are toxic for the host or/and are very hydrophobic, difficulting its expression in soluble form. Despite the more complex downstream processing, in such cases expression of these peptides in inclusion bodies pose a better strategy than trying to produce it in soluble form.

A good example of this strategy is the one implemented by Sharpe *et al.* in 2005 to express truncates from Alzheimer's β-amyloid (Aβ) protein. This protein tends to be toxic to the host strain and form aggregates, as so as its derived peptides. In this case, toxicity of the peptides is the main bottleneck of the expression step, since the conjugation of a soluble fusion partner enhanced its solubility but the yield of expression was far too low due to its toxicity. The expressed peptide was $A\beta_{11-26}$ that comprise the aminoacids 11 to 26 of Aβ protein. Three repeats of $A\beta_{11-26}$ separated by Met residues were fused with a KSI (ketosteroid isomerase) carrier in its N-terminal and a His-tag in its C-terminal. The KSI carries the fusion protein to inclusion bodies, where the peptides are less toxic to the host and amyloid formation is reduced. The use of tandem repeats of short peptides increases the expression yield because a larger portion of the fusion protein corresponds to the target peptide. His-tag is used because it enables the purification under denaturant conditions that are needed for inclusion body solubilization. After purification through immobilized metal affinity chromatography (ICAM) and dialysis against water, fusion peptides were cleaved using CN Br cleavage method and purified through HPLC. (Sharpe *et al.*, 2005)

Examples like the two presented show that, although recombinant strategies are frequently preferred over chemical synthesis, these strategies are not trivial, and many optimization tests have to be done in order to achieve reasonable levels of production.

2.2 Azurin

2.2.1 Overview

Azurin is a low molecular weight protein (128 aa – 14 kDa), member of the cupredoxin family produced by *P. aeruginosa*. Cupredoxins are copper-containing water-soluble proteins involved in the electron transport chain of prokaryotes and eukaryotes.

Azurin has anticancer, antiparasitic and antiviral properties (Fialho *et al.*, 2007a). These properties are due to azurin's ability to establish high-affinity interactions with some proteins relevant in these diseases. This ability makes azurin a natural scaffold protein that has immunoglobulin-like binding characteristics (Fialho *et al.*, 2007b).

As an anticancer agent, azurin acts in three different forms: entering into the cancer cell and stabilizing the tumor suppressor protein p53, thus inducing apoptosis; interfering extracellularly in the ephrin/Eph signaling system that is involved in tumor progression, angiogenesis, migration and invasion (Chaudhari *et al.*, 2007); and inhibiting the VEGFA (vascular endothelial growth factor A) involved in angiogenesis (Mehta *et al.*, 2011).

Amino acids 50 to 77, p28, act as protein transduction domain (PTD) and are responsible for azurin entry into cells and p53 stabilization (Yamada *et al.*, 2009). The azurin PTD was refined to amino acids 50 to 67, the hydrophobic peptide p18 whereas amino acids 68 to 77, hydrophilic peptide p10, are those responsible for p28's antiproliferative activity. (Taylor *et al.*, 2009)

Azurin has antiparasitic and antiviral action against the malaria parasite *Plasmodium falciparum*, the toxoplasmosis causing protozoa, *Toxoplasma gondii*, and the HIV-1 virus. It binds to host cell ICAMs (intercellular adhesion molecules) that act as receptors for various pathogens including *P. falciparum* and HIV-1, and also binds to a number of surface proteins of parasites and virus, interfering in the entry of the infectious agents into the host cells (Chaudhari *et al.*, 2006).

Azurin is also produced by other pathogenic bacteria besides *P. aeruginosa*. The azurin-like protein produced by meningitis-causing bacteria, *Neisseria meningitides* is termed Laz. Unlike most azurin produced by other bacteria, Laz is not periplasmic but surface exposed. It has an additional 39 amino acid lipidated tail in its N-terminal called H.8 epitope. Comparing to azurin, Laz is much more efficient entering glioblastoma cells and has higher cytotoxicity level against these cells. This suggests that the H.8 epitope is important in disrupting entry barriers to glioblastoma cells. The cloning of the H.8 epitope in other toxic drug candidates might be very promising in the treatment of brain tumors. (Fialho *et al.*, 2008)

2.2.2 Structure

Members of the cupredoxin family have structural features similar to the immunoglobulin variable domains, a large group of specific antibodies. The common structure is formed by two β-sheets made up of seven or more parallel and anti-parallel strands (β-sandwich) (Fialho *et al.*, 2008; Fialho *et al.*, 2007b). The similarity of azurin with the variable domains of several immunoglobulins demonstrates its single antibody-like structure. This antibody-like structure consists in eight anti-parallel strands stabilized by a disulphide bridge that form a compact and rigid β-sandwich core (Bernardes *et al.*, 2010; Fialho *et al.*, 2008). Although azurin and immunoglobulins have low sequence identity, their secondary structure can be aligned by structural superposition, suggesting a common ancestry (Fialho *et al.*, 2007b).

In the middle part is located the p28. The p28 is an extended α-helix PTD responsible for the azurin's entry specificity in cancer cells. Azurin also has four exposed loop regions that are thought to be involved in binding with other proteins (Bernardes *et al.*, 2010). These exposed loop regions are analogous to those occurring naturally in antibodies, enforcing its property of natural scaffold protein (Fialho *et al.*, 2007b).

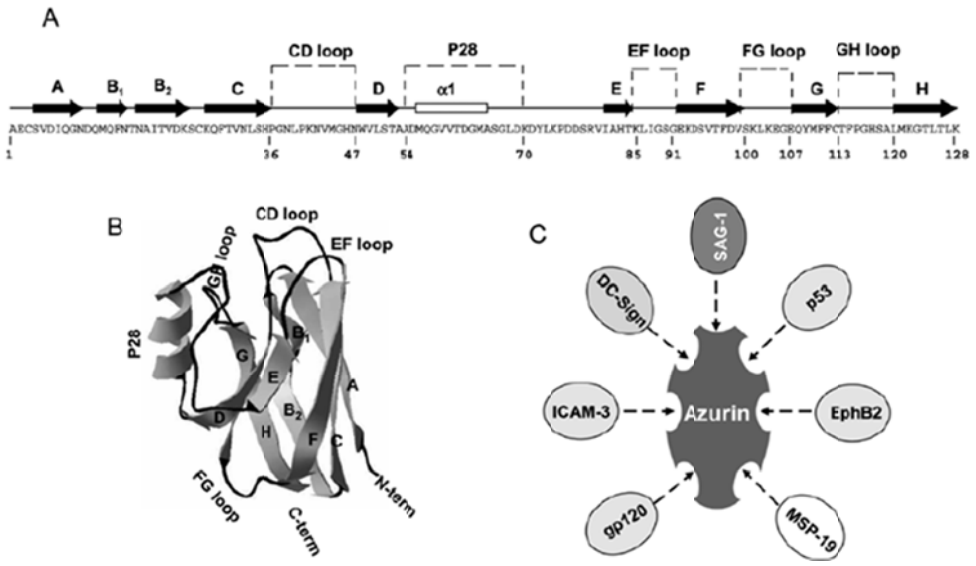

Figure 1 – Azurin's structure. A: Azurin's amino acid sequence highlighting the p28 α-helix and four loop regions, candidates for binding domains; B: Azurin single antibody-like structure, consisting in a β-sandwich core of eight anti-parallel strands and four exposed loops; C: Azurin's target molecules, p53 tumor suppressor, receptor tyrosine kinase EphB2, MSP-19 antigenic protein, gp120 HIV surface protein, ICAM-3 and DC-SIGN adhesion molecules, and SAG-1 parasitic protein. Adapted from (Fialho *et al.*, 2008).

2.2.3 Anticancer action

As previously mentioned, azurin is a multi-targeted anticancer agent that acts through three pathways: inhibition of cell cycle progression through extracellular binding to Eph receptor tyrosine kinases, induction of apoptosis through intracellular stabilization of p53 protein and

prevention of angiogenesis through inhibition of VEGFA. These pathways are independent and together lead to the inhibition of cancer growth.

2.2.3.1 Inhibition of cell cycle progression

Ephrins are a family of receptor tyrosine kinases that bind to their ephrin ligands generating signals at sites of cell-cell contact controlling cell morphology, adhesion, migration and invasion. In many tumors, Eph and ephrins are overexpressed and promote cancer growth and angiogenesis. Ephrins can be divided into two classes: A-type, that are cell membrane-linked and B-type, which possess highly conserved intracellular domains and a transmembrane domain. (Chaudhari *et al.*, 2007; Pasquale, 2010)

Azurin has structural similarity with ephrinB2, EphB2 ligand, and binds strongly to EphB2, interfering with the ephrin/Eph signaling system. This interference contributes to the induction of cancer cell death and tumor regression. (Chaudhari *et al.*, 2007)

In 2007, Chaudhari *et al.* (Chaudhari *et al.*, 2007) showed that azurin and plastocyanin, other cupredoxin, have a remarkable structural similarity with the extracellular domains of some ephrins. Surface plasmon resonance (SPR) analysis showed that azurin has the highest affinity to EphB2, followed by EphA6, EphA4 and A7. The azurin derived peptide Azu 88-113 showed similar affinities and demonstrated to interfere in EphrinB2 signaling pathway by competing with ephrinB2 on binding to the tyrosine kinase.

2.2.3.2 Induction of apoptosis

Yamada *et al.* demonstrated that azurin, and more specifically p28, can enter cancer cells and induce p53 mediated apoptosis in murine J774, human breast cancer, melanoma and osteosarcoma but not in normal liver cells or p53-negative osteosarcoma cells (Yamada *et al.*, 2005; Yamada *et al.*, 2009).

The azurin entry into cancer cells is mediated by its protein transduction domain (PTD), p18, through a caveolin-mediated pathway. Caveolins (caveolin-1, caveolin-2 and caveolin-3) are membrane proteins that act as regulators of signal transduction and are present in a subset of lipid raft invaginations called caveolae. The increased caveolin-1 expression in cancer cells over normal cells partly explains the preferential entry of azurin into cancer cells. Although, kinetic studies show that this is not the only entry pathway of azurin. (Taylor *et al.*, 2009; Yamada *et al.*, 2005)

After internalization, azurin interacts strongly with the p53 forming a stable complex of four azurins binding to one p53 monomer. This complex stabilizes and shields p53 from degrading enzymes like ubiquitin ligases, increasing its intracellular levels (Apiyo and Wittung-Stafshede, 2005; Yamada *et al.*, 2009). Regulation of p53 expression is mainly done through ubiquitin proteosomal pathways. Because some ubiquitin ligases are overexpressed in some cancers,

like breast and ovarian cancers, p53 levels are naturally low and its apoptotic and cycle arrest action is reduced (Yamada *et al.*, 2009).

In 2004, Punj *et al.* showed that azurin-induced apoptosis in MCF-7 cells occurs partly because of p53 stabilization and consequent downregulation of *bcl2* and upregulation of *bax* genes. p53 is a transcriptional regulator of Bax and Bcl2, a proapoptotic and antiapoptotic proteins, respectively. The downregulation of Bcl2 levels and upregulation of Bax levels triggers Bax translocation to mitochondria and release of cytochrome *c* that will activate a caspase cascade that induce apoptosis. (Punj *et al.*, 2004)

As referred before, p28 is an important fragment of azurin that can, by itself, enter cancer cells and induce apoptosis and cell arrest through stabilization of p53 tumor suppressor. p28 is also able to inhibit angiogenesis (Mehta *et al.*, 2011).

In 2009, Yamada *et al.* studied the mode of action of p28 in p53-positive breast cancer cells. Confocal microscopy showed that p28 preferentially enter tumorigenic cells (adenocarcinoma MCF-7, ductal T47D and carcinoma ZR75-1) comparing to non-tumorigenic cells (epithelial MCF-10A cells). The incubation of MCF-7 and MCF-10A cells with caveolae disrupting agents (methyl-β-cyclodextrin and filipin) and microtubule disrupting agents (nocodazole) inhibited significantly p28 entry, confirming the involvement of a caveolae-mediated endocytosis in p28 penetration. Clathrin-mediated endocytosis inhibitors (chlorpromazine, amiloride and sodium azide) did not affect p28 entry. MCF-7 cells were more sensitive than MCF-10A cells to caveolae disrupting agents, revealing that MCF-7 cells may have more entry sites at cell surface. Incubation of MCF-7 and ZR-75-1 cells in medium supplemented with p28 showed that p28 has a time- and dose-dependent antiproliferative effect. 30 days treatment of MCF-7 tumor xenograft in athymic mice with intraperitoneal injection of p28 significantly inhibited tumor growth without reducing the body weight and changing mice behavior.

MCF-7 treatment with p28 elevated p53 intracellular levels. GST-pull down assays using p28, p18 (aminoacids 50 to 67), p18b (aminoacids 60 to 77) and p12 (aminoacids 66 to 77) showed that p28, p18 and p18b pulled down p53, suggesting that the azurin region responsible for binding to p53 is located between azurin aminoacids 60 and 67. MCF-7 cells treatment with p28 enhanced the DNA-binding activity of p53. Reverse transcription PCR and studies of p53 ubiquitination in presence of p28, MDM2 (main responsible for p53 ubiquitination) and MG132 (proteasome inhibitor) showed that p53 increased levels due to p28 presence result from post-transcriptional stabilization of p53 by p28 and that, although p28 does not compete with MDM2, it inhibits other ubiquitination and proteasomal pathways. Therefore, p28 stabilizes p53 decreasing its ubiquitination and proteasomal degradation. Computational analyzes and western blotting assays with specific antibodies for different regions of p53 suggest that p28 binds to p53 in its DNA binding domain (DBD), between aminoacids 80 and 276.

To evaluate the extent of p28 treatment effect in the cellular cycle and apoptosis, levels of some proteins transcriptionally regulated by p53 (p21 and p27) were analyzed in treated MCF-7 and

MDD2 cells. In the case of MCF-7 cells, p21 (inhibitor of CDK activity) levels were elevated, inhibiting cell cycle progression through G_2-M phase. p21 binds to cyclin A, inhibiting CDK2 (cell cycle progression at G_1 regulator) and induces cdc2 (cell cycle progression at G_2-M regulator) inactivation through phosphorylation. Therefore, the levels of the natural activator of cdc2, cyclin B1, are elevated in p28 treated cells and levels of CDK2 and cyclin A are reduced over time of p28 exposure. p27 levels were also elevated. In MDD2 cells p27 was not detectable and p21 levels were not altered, as well as the levels of the analyzed cyclins and CDKs. (Yamada *et al.*, 2009)

CDG Therapeutics Inc, an American biotechnology company, is developing a drug consisting in the synthetic p28 peptide for the treatment of p53-positive cancer. Several patents have been issued on this matter and, at the date, p28 ended Phase I clinical trials. (CDG, 2011)

2.2.3.3 Inhibition of angiogenesis

Mehta *et al.* very recently showed that p28 preferentially enter endothelial cells and inhibit angiogenesis and tumor growth (Mehta *et al.*, 2011).

VEGFA is an angiogenic factor that activates ECM (extracellular matrix) proteins that induce quiescent endothelial cells to form a tube-like structure. The angiogenic VEGFA effect is essentially mediated by one of its receptors, VEGFR-2. This binding activates several proteins like FAK (focal adhesion kinase) that are involved in cell motility processes. p28 co-localizes with caveolin-1 and VEGFR-2 and inhibits VEGFR-2 activity. (Mehta *et al.*, 2011)

p28 possesses antiangiogenic properties by interfering with different VEGFR-2 and VEGFA target proteins and pathways like F-actin, a stress fiber that is related to cell motility and migratory ability; FAK, a non-receptor protein tyrosine kinase associated with supramolecular focal adhesion complexes that is important in cell attachment and movement; and PECAM-1 (platelet/endothelial cell adhesion molecule-1), responsible for endothelial cell contact maintenance. (Mehta *et al.*, 2011)

Different observations suggest that p28 induces reduction of cell motility, through increasing of cell rigidity and ultimately impeding migration and capillary tube formation, therefore inhibiting angiogenesis (Mehta *et al.*, 2011).

2.2.4 Azurin as a scaffold protein

As mentioned before, azurin acts as a scaffold protein because it mediates different high-affinity interactions with several proteins. These high-affinity interactions are due to azurin's structural similarity with those proteins, enabling the formation of complexes between azurin and the target proteins (Fialho and Chakrabarty, 2010).

Different protein interaction analysis show that azurin is capable of forming stable complexes (K_D in the nM range) with the different proteins referred before: p53, EphB2, ICAM-3, DC-SIGN, gp120, MSP-1 and SAG1 (Figure 2) (Fialho and Chakrabarty, 2010).

Azurin derived peptide of the 88-113 region, Azu 88-113 is involved in azurin's interaction with EphB2 and with DC-SIGN. This peptide includes the F-G loop, which is the responsible region for the complex formation with EphB2 (Fialho and Chakrabarty, 2010; Fialho *et al.*, 2008).

Azurin interaction region with p53 was identified as being its hydrophobic region, which includes the aminoacids Met-44 and Met-64 (Yamada *et al.*, 2002a). De Grandis *et al.* proposed that the hydrophobic region of azurin bind to the p53's DNA binding domain (DBD) involving the L_1 and s_7-s_8 loops (De Grandis *et al.*, 2007).

The different regions involved in azurin binding to EphB2 and p53 strongly suggests that different regions of azurin are specialized in interacting with different proteins, *i.e.*, particular domains of azurin bind to particular proteins (Fialho *et al.*, 2008).

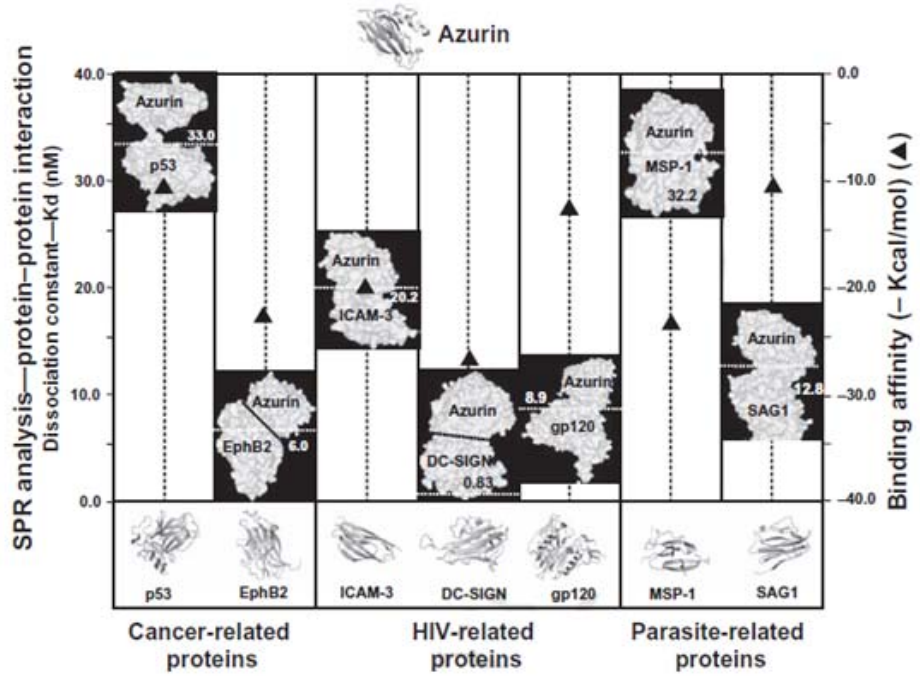

Figure 2 – Azurin is considered a scaffold protein as it can form stable complexes with various proteins involved in cancer (p53 and EphB2), HIV-1 infection (ICAM-3, DC-SIGN and gp120) and parasitemia (*P. falciparum* MSP-1 and *T. gondii* SAG1). Adapted from (Fialho and Chakrabarty, 2010).

2.2.5 Azurin derived peptides as therapeutic agents

As referred before, the synthetic p28 peptide ended Phase I clinical trials and acts as an anticancer agent against p53 positive cancer (Jia *et al.*, 2010). Another study is being conducted using other azurin derived peptides as a targeted delivery system of radiosensitizers (Micewicz *et al.*, 2011).

Based on azurin affinity with Ephs, Micewitz *et al.* studied the development of azurin derived peptides as targeted delivery systems of nicotinamide, a potent radiosensitizer currently used in cancer treatment (Micewicz *et al.*, 2011). As previously said, azurin region responsible for binding with Ephs is the 96-113 region (Chaudhari *et al.*, 2007). The developed peptides were based on the C-terminal region of azurin and were modified to increase their stability. These modifications include the addition of D- and other unusual aminoacids. Nicotinamide was conjugated in the N-terminal of the most stable and binding effective peptide. The final product was a linear peptide named AzV36-NicL. Because some types of cancer (lung, breast and prostate cancers) overexpress Ephs at the cell surface, conjugating a radiosensitizer to a peptide with affinity to these receptors will target the radiosensitizer to cancer cells. Having a targeted way of delivering the radiosensitizer may increase the radiotherapy efficacy and reduce the required dose of radiation. (Micewicz *et al.*, 2011)

In vitro and *in vivo* studies showed that using this peptide-based-radiosensitizer-targeting strategy conjugated with radiotherapy increases the efficacy of the treatment. This strategy increases radiotherapy efficacy about ~13-fold in a artificial metastasis mouse model and delays the tumor growth in about 2.5 days in solid tumor engraftment mouse model. Other in vitro studies also showed that this peptide has low cytotoxic effect in the *in vivo* models and is very stable in human serum. (Micewicz *et al.*, 2011)

Altogether, these observations show that this azurin derived peptide has promising therapeutic properties and that the study of other azurin derived peptides can open new doors in cancer treatment whether using azurin anticancer activity or azurin high affinity targeting.

2.3 Cadherins in breast cancer

2.3.1 Overview

Cadherins are calcium-dependent cell-cell adhesion proteins that play an essential role in morphogenesis and in the development and maintenance of adult tissues (Paredes *et al.*, 2005; Soler *et al.*, 1999). Cadherins interact intracellularly with proteins called catenins. The catenins link the cadherins to the cytoskeleton and mediate signal-transduction mechanisms that control the cell polarity, differentiation, growth and death, among others (Paredes *et al.*, 2007). The homotypic binding between identical cadherins promote, during development, the sorting of cells into distinctive tissues, and at adult stage, the maintenance of the differentiated states of tissues and suppression of tumor progression (Soler *et al.*, 1999).

The best characterized cadherins are the classical cadherins that include E-cadherin (epithelial) and P-cadherin (placental). Although E-cadherin is present in all epithelial tissues, P-cadherin is only present in the lower layer of stratified epithelia including breast, skin and prostate myoepithelial cells (Paredes *et al.*, 2005). Because of the importance of cadherins in cell

adhesion, recognition, sorting and signaling, their altered function has serious implications in disease states, including cancer (Paredes *et al.*, 2007).

In breast carcinoma, several evidences indicate that tumor development and invasion is associated with low expression and function of cadherins. E-cadherin is considered a tumor suppressor in the breast and has frequently low expression in breast carcinomas. This low expression can be caused by different mechanisms like hypermethylation of the E-cadherin promoter region, loss of heterozygosity of the long arm of the chromosome 16, where the E-cadherin gene is located, and E-cadherin mutations. E-cadherin low expression in breast carcinoma is associated with dedifferentiation, increased invasiveness and high metastatic potential. (Soler *et al.*, 1999)

Very recent studies identified the cadherins family as a good candidate target for azurin (Fialho, 2009).

2.3.2 P-cadherin

P-cadherin is a 118 kDa glycoprotein with three distinct domains: extracellular, transmembrane and cytoplasmic tail. The cytoplasmic tail has two main domains: the catenin binding domain (CBD), essential for the cadherin function, and the juxtamembrane domain (JMD), which is thought to be important in cell relocation. (Paredes *et al.*, 2007)

P-cadherin expression in breast carcinoma has been shown to be a valuable indicator of poor patient survival, better than the level of E-cadherin expression (Soler *et al.*, 1999). It has been detected as altered in different types of human cancer but its role in carcinogenesis remains unclear because it behaves differently according to the type of cancer: in malignant melanoma P-cadherin behaves as tumor suppressor while in breast cancer it promotes tumor invasion and aggressiveness. (Paredes *et al.*, 2007)

P-cadherin show aberrant expression in ~30% of breast carcinomas (Paredes *et al.*, 2005) and tumors overexpressing P-cadherin usually present a high histological grade and decreased cell polarity (Ribeiro *et al.*, 2010). The P-cadherin molecule is essentially expressed in basal-like carcinomas which have triple negative phenotype (negative for oestrogen receptor, progesterone receptor and HER2 (human epidermal growth factor receptor 2)). (Paredes *et al.*, 2007)

P-cadherin expression was reported as being inversely related to hormonal receptor content (oestrogen, progesterone and epidermal growth factor), p53 expression, high proliferation rates and decreased cell differentiation. These biological conditions are strongly associated with poor survival of breast cancer patients. (Paredes *et al.*, 2007)

2.3.2.1 P-cadherin and breast cancer cells invasion, motility and migration

Recently, Ribeiro *et al.* reported that the overexpression of P-cadherin in wild-type E-cadherin breast cancer cells increases cell invasion, motility and migration (Ribeiro *et al.*, 2010). P-cadherin overexpressing breast cancer cells showed to have increased single cell motility (increased number of moving cells and their speed), increased directional migration and invasion capacity through Matrigel. This behavior was demonstrated to be directly dependent on P-cadherin as when cells are treated with blocking antibodies or its transcription is inhibited through siRNA transfection, inhibition of migration and invasion is observed. (Ribeiro *et al.*, 2010)

P-cadherin was also showed to be involved in cell motility as it can alter cell phenotype through changes in cell polarity, formation of membrane protrusions and increasing cytoplasmic area, characteristic of motile cells (Ribeiro *et al.*, 2010).

2.3.2.2 MMPs and P-cadherin shedding

Tumor growth process includes local proteolysis of the ECM and migration of the tumor cells through the degraded tissue. The ECM degradation is carried out by several proteases, including the matrix metalloproteases family (MMPs) (Bostrom *et al.*, 2011). As a family, the MMPs degrade most components of the ECM (Vu and Werb, 2000).

Ribeiro and coworkers reported that the levels of active MMP-1 and MMP-2 are significantly induced by P-cadherin expression, facilitating cell invasion. MMPs are involved in shedding of extracellular domains of membrane glycoproteins, like E-cadherin, that origin a soluble fragment with pro-invasive activity. In the case of P-cadherin, the MMPs also shed the soluble ectodomain of P-cadherin (sP-cad) which has pro-invasive activity (Ribeiro *et al.*, 2010).

In the same work was also shown that sP-cad is produced by MMP-1 and MMP-2 and that it is independently responsible for the invasive behavior of the breast cancer cells (Ribeiro *et al.*, 2010).

2.3.2.3 Azurin and P-cadherin

Very recent studies show that cadherins are good candidates for azurin binding. Because P-cadherin is associated with poor survival of breast cancer patients, it is of great interest study if azurin is able to antagonize its invasive effect. Therefore, studying azurin and azurin derived peptides interaction with P-cadherin and sP-cad is important to better understand if P-cadherin can indeed be a target for azurin and if azurin is able to act as an anticancer drug to treat breast cancer overexpressing P-cadherin. (Fialho, 2009)

2.4 Project aim

The main objective of this work is to develop a reproducible method of expression and purification of azurin derived peptides. Having a robust method of peptide production facilitates further studies towards the determination of azurin region implied in the P-cadherin inhibition. Because recent studies report that azurin interacts with P-cadherin, it is important to search for the interaction region.

Three azurin derived peptides were studied, Azu 1-50, Azu 1-77 and Azu 80-128 (Figure 3). Azu 1-50 corresponds to the first fifty azurin aminoacids and is mainly constituted by β-sheets. Azu 1-77 corresponds to the addiction of the PTD p28 to the Azu 1-50. Azu 80-128 consists in the late 48 aminoacids of azurin.

Peptides expression was performed in *E. coli* strains and purification was done through poly-Histidine affinity.

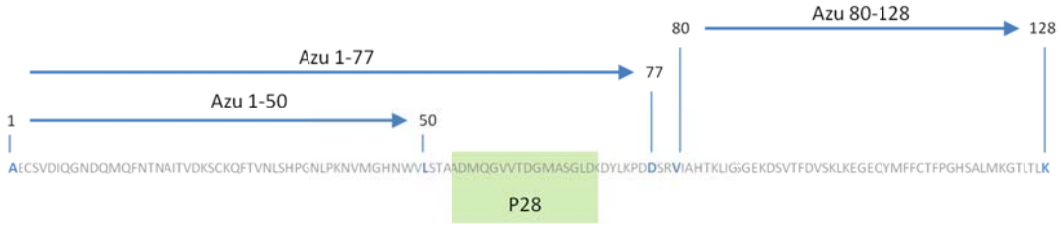

Figure 3 – Azurin and azurin derived peptides amino acid sequences. Azu 1-50 corresponds to the first 50 amino acids of azurin, Azu 1-77 include Azu 1-50 and the p28. Azu 80-128 corresponds to the last 48 amino acids.

The development of a reproducible method of producing azurin derived peptides can also be useful in other studies of azurin activity. These peptides can be tested to resolve the azurin regions involved in their different activities (antiviral, antiparasitic and anticancer).

3 Materials and methods

3.1 Bacterial strains

P. aeruginosa PAO1 DNA was used to amplify azurin gene.

pWH844 plasmid was extracted from *E. coli* XL1-Blue.

Competent *E. coli* XL1-Blue was used for transformation and plasmid amplification.

For plasmid amplification, bacterial cultures were grown at 37°C on LB broth (Lennox) (Conda). For selection, ampicillin was added to a final concentration of 150 µg/mL. For solid media, LB agar (Lennox) was used.

Competent *E. coli* BL21(DE3) and *E. coli* SURE were used to study azurin derived peptides expression.

For peptide expression, bacterial cultures were grown either overnight at 20°C or grown for 4h at 37°C on SB broth (32 g/L of tryptone (Becton, Dickinson and Company); 20 g/L of yeast extract (Himedia); 5 g/L of sodium chloride (Panreac)) supplemented with IPTG at the desired concentration. For selection, ampicillin was added to a final concentration of 150 µg/mL.

3.2 DNA and peptide quantification

DNA quantification from samples was done using the Nanodrop spectrophotometer ND 1000 (Thermo Scientific).

Protein quantification was done through the Bradford method. 10 µL of peptide samples and BSA standard solutions were loaded into microplate wells and 200 µL of diluted (1:5) Bio-Rad Protein assay dye reagent concentrate were added. The microplate was incubated for 15 min at room temperature and OD was measured at 595 nm using a VERSAmax tunable microplate reader (Molecular Devices).

3.3 DNA manipulation and cloning

3.3.1 Azurin derived peptides DNA sequences isolation by PCR DNA amplification

azu 1-50, azu 1-77 and azu 80-128 sequences were obtained by PCR DNA amplification using the *P. aeruginosa* genomic DNA as template and different primers.

The *P. aeruginosa* genomic DNA was extracted from 700 µL of a culture grown overnight in LB medium at 37°C using the DNeasy Blood & Tissue (QIAGEN) and following the manufacturer's

instructions. 700 µL of cell culture were centrifuged for 10 min at 7500 rpm. The pellet was ressuspended in 180 µL of ATL buffer, added 20 µL of proteinase K and incubated at 56°C for 1-3 hours. After incubation the mix was vortexed 15 s and mixed with 200 µL of buffer AL and ethanol. The mix was loaded into a DNeasy column and centrifuged 1 min at 8000 rpm. The column was washed with 500 µL of buffer AW2 and eluted with 50 µL of water.

Primers were designed based on the Custom Primers - OligoPerfect™ Designer from Invitrogen.

The primers sequences and the PCR mix solution are shown, respectively, in Table 2 and Table 3. The PCR cycle conditions are shown in Table 4.

Table 2 – Primers sequences for PCR.

DNA sequence	Primers
azu 1-50	Fwd: 5' – CGG GAT CCG CCG AGT GCT CGG TGG ACA T – 3'
	Rev: 5' – CCC AAG CTT CAG TAC CCA GTT GTG GCC C – 3'
azu1-77	Fwd: 5' – CGG GAT CCG CCG AGT GCT CGG TGG ACA T – 3'
	Rev: 5' – CCC AAG CTT GTC GTC GGG CTT CAG GTA – 3'
azu 80-128	Fwd: 5' – CGG GAT CCG TCA TCG CCC ACA CCA AG – 3'
	Rev: 5' – CCC AAG CTT GCA TCA CTT CAG GGT CAG GGT – 3'

Table 3 – PCR mix solution composition.

	Volume (µL)
PCR buffer (10x) (Invitrogen)	5
dNTPs (10 µM)	1
$MgCl_2$ (50 µM) (Invitrogen)	1.5
Primer fwd (100 µM)	1
Primer rev (100 µM)	1
DNA template	(corresponding to 50 ng)
Taq polymerase platinum (Invitrogen)	0.2
Water to a final volume of	50

Table 4 – PCR cycle conditions for amplification of the azurin derived peptides DNA sequences.

Step		Temperature (°C)			Number of cycles	Hold time (s)
		azu 1-50	azu 1-77	azu 80-128		
Pre incubation		94	94	94	1	120
Amplification	Denaturation	94	94	94	30	30
	Annealing	58	56	58	30	30
	Elongation	72	72	72	30	30
Cooling		72	72	72	1	600

3.3.2 DNA agarose gel electrophoresis

PCR products were observed by agarose gel electrophoresis in a 4% NuSieve 3:1 agarose (Cambrex) gel. The agarose used in this work is specific for low weight DNA analysis (≤1 kb) and has higher gel strength than common agaroses that ensures a good resolution of DNA fragments up to 1 kb.

25 µL of PCR product were loaded with 2.5 µL of 10x DNA loading buffer (Fermentas). 5 µL of O'GeneRuler™ 50bp DNA ladder (Fermentas) were also loaded. Gels were run at 100 V using TAE buffer (1x). The DNA was stained by immersion in Gel Red solution. The resulting gel was visualized with the Quantity One system (BioRad).

3.3.3 PCR results extraction from agarose gel

DNA was extracted from the gel using the Jetquick gel extraction spin kit (Genomed). The DNA bands were cut from the gel and transferred to 2 mL tubes. For each 100 g of gel band 600 µL of L1 solution were added. The gel was dissolved by 15 min incubation at 50°C. The dissolved gel was then transferred to a Jetquick column and centrifuged for 1 min at 12000 rpm. The supernatant was discharged and the column was washed with 500 µL of L1 solution and then washed with 500 µL of L2 solution. The column was eluted with 30 µL of water.

3.3.4 DNA restriction endonuclease digestion

The DNA sequences were digested for further ligation to the pWH844 vector.

The DNA sequences digestion were made mixing about 300 ng of DNA with 0.5 µg of each restriction enzyme (BamHI and HindIII from Takara), 4 µL of buffer K (Invitrogen) and water to perform a final volume of 40 µL. Mixes were incubated over night at 37°C.

The resulting digested DNA was purified by precipitation.

The DNA sequences were precipitated by adding to the digestion mix 4 µL of sodium acetate 3 M and 100 µL of cold absolute ethanol. The mix was kept for 1h at -80°C and centrifuged at 4°C (15,300 rpm, 30 min). After discharging the supernatant, the pellet was washed in 500 µL of 70% ethanol and centrifuged at 4°C (15,300 rpm, 10 min). Again, the supernatant was discharged and the pellet was dried for 15 min in a speed vacuum. The dried pellet was ressuspended in 10 µL of water.

3.3.5 Plasmid preparation

The pWH844 was extracted from *E. coli* XL1-Blue.

The vector was extracted from 24 mL of bacterial cells grown overnight at 37°C in LB medium with ampicillin (150 µg/mL) using the QIAprep Spin Miniprep Kit (Qiagen). Briefly, the cell suspension was spun down and ressuspended in buffer P1. The suspension was transferred to 2 mL tubes and mixed with buffer P2 and N3, and centrifuged for 10 min at 14000 rpm. The supernatant was applied to a QIAprep column and centrifuged for 1 min at 14000 rpm. The column was washed with PB buffer and PE buffer, and eluted with water.

The extracted vector was digested by mixing about 1 µg of plasmid with 0.5 µg of each restriction enzyme (BamHI and HindIII) (Takara), 2 µL of buffer K (Invitrogen) and water fulfilling a final volume of 20 µL. The mix was incubated over night at 37°C.

The digested vector was precipitated by the same protocol described in section 3.3.4.

The cloning of the inserts into the pWH844 were made mixing 75 ng of insert, 25 ng of plasmid, 4 µL of buffer T4 (Invitrogen), DNA ligase T4 (Invitrogen) and water to perform a final volume of 20 µL. The mix was kept for 1 h at room temperature and at 4°C overnight.

3.3.6 *E. coli* XL1-Blue transformation

Ligation mix was transformed in competent *E. coli* XL1-Blue for analysis of the constructs.

150 µL of competent cells were mixed with 50 µL of TCM and 50 µL of construction (0.2 ng/µL). The mix was kept on ice for 30 min, heat shocked for 3 min at 42°C and kept for 5 min on ice. 800 µL of LB medium were added and the mix was incubated at 37°C for 1 h in a incubate shaker. The cells were centrifuged (8000 rpm, 2 min) and ressuspended in 100 µL of LB medium. 50 µL were plated on LB plates with ampicillin (150 µg/mL) and incubated overnight at 37°C.

3.3.7 Colony PCR

Colony PCR was performed to confirm the transformation of *E. coli* XL1-Blue.

9 µL of a cell culture grown overnight in LB medium with ampicillin (150 µg/mL) at 37°C was mixed with 1 µL of triton 10% and boiled for 5 min in PCR tubes. The boiled cell culture was added to other reagents necessary for colony PCR showed in Table 5. The colony PCR cycle conditions are shown in Table 6.

Table 5 – Colony PCR mix solution composition.

	Volume (µL)
PCR buffer (10x) (Citomed)	2.5
dNTPs (10 µM)	0.5
$MgCl_2$ (20 µM) (Citomed)	2.5
Primer fwd (100 µM)	0.5
Primer rev (100 µM)	0.5
Boiled cell culture in 1% triton	10
Taq polymerase Taqmed (Citomed)	0.1
Water to a final volume of	25

Table 6 – Colony PCR cycle conditions for confirmation of E. coli XL1-Blue transformation.

Step		**Temperature (°C)**			**Number of cycles**	**Hold time (s)**
		azu 1-50	azu 1-77	azu 80-128		
Pre incubation		94	94	94	1	300
Amplification	Denaturation	94	94	94	30	30
	Annealing	58	56	58	30	20
	Elongation	72	72	72	30	30
Cooling		72	72	72	1	420

3.3.8 Plasmidique constructions extraction

The plasmidique constructions extractions from *E. coli* XL1-Blue were done using the QIAprep Spin Miniprep Kit (Qiagen).

3.4 Peptide production

3.4.1 *E. coli* SURE and *E. coli* BL21(DE3) transformation

The constructions were transformed into competent *E. coli* SURE in order to express the azurin derived peptides. The transformation was done using the same protocol described in section 3.3.6.

E. coli BL21(DE3) were transformed with the constructions by electroporation using approximately 10 ng of DNA, a capacitance of 25 μF, 2.5 kV and 400 Ω of resistance.

3.4.2 Superexpression

Transformed cells were incubated over night at 37°C with continuous shaking in LB medium with ampicillin (150 μg/mL). In the following day, SB medium with ampicillin (150 μg/mL) was inoculated with a volume of the previous grown cells to perform a OD_{640} of 0.1 and incubated at 37°C with continuous shaking until OD_{640} reached 0.6-0.8. At this point, IPTG was added to the medium to obtain the desired concentration. The induced inoculums were incubated for 4-6h (at 37°C) or overnight (at 20°C).

3.4.3 Tricine-SDS-PAGE

For superexpression samples, OD_{640nm} was measured and the collected sample volume was calculated as:

$$\mathrm{V\ (mL)} = \frac{1.2}{\mathrm{OD_{640nm}}} \qquad \text{Eq. 1}$$

Samples were centrifuged 1 min at 8000 rpm, the supernatant was discharged and the pellet ressuspended in 80 μL of 2x SDS gel-loading buffer.

For NiNTA purification fraction samples, 10 μL of the collected fractions were mixed with 10 μL of 2x SDS gel-loading buffer.

If not used immediately, samples were kept at -20°C.

The gel was assembled in an Amersham Biosciences gel caster. About 3 mL of resolving gel were casted and overlaid with isopropanol 70% to create a smooth surface. The gel was left to polymerize for about 30 min. The isopropanol was removed, the comb inserted (10 wells) and the stacking gel was casted to the top of the plates. The stacking gel was left to polymerize for about 30 min.

The composition of the stacking and resolving gel are shown in Table 7 and Table 8.

Table 7 – 5% stacking gel composition for one gel (8 cm x 10 cm x 0.7 mm). (Sambrook and Russel, 2001)

	5% Stacking gel
Water (mL)	1.7
Acrylamide-bisacrylamide 30% (µL)	415
1,0 M Tris base (pH 6,8) (µL)	315
SDS (10%) (µL)	25
APS (10%) (µL)	25
TEMED (µL)	2.5

Table 8 – 16% / 6 M urea resolving gel composition for one gel (8 cm x 10 cm x 0.7 mm). (Schagger, 2006)

	16% / 6 M Urea Resolving gel
AB-3 (mL)(*)	1.665
Gel buffer (3x) (mL)	1.665
Urea (g)	1.8
Water to the final volume (mL)	5
APS (10%) (µL)	16.5
TEMED (µL)	1.65

(*)**AB-3 solution:** 48 g of acrylamide and 1.5 g of bisacrylamide in 100 mL of water.

For electrophoresis, the comb was removed and the gel cassette was placed in the electrophoresis apparatus from Amersham Biosciences. Anode buffer (1x) and Cathode buffer (1x) were added to the apparatus as the lower and upper buffer, respectively.

Samples were boiled at 100°C for 3 min and centrifuged (9000 rpm, 1 min), 10 µL of each sample were loaded to the gel wells. 3 µL of PageRuler Prestained Protein Ladder (Fermentas) were also loaded.

The electrophoresis started at 30 V and this voltage was maintained until the samples entered the resolving gel, after that, the voltage was set to 200 V and the maximum current was set to 80 mA.

The gel was then stained with Coomassie Brilliant Blue. The gel was stained for at least 1h with Coomassie blue staining solution. The gel was destained with acetic acid destaining solution

until the background of the gel was transparent and the protein bands were clear. To remove the destaining solution, the gel was washed with water.

The compositions of the electrode and gel buffers used in Tricine-SDS-PAGE are shown in Table 9.

Table 9 – Electrode and gel buffers for Tricine-SDS-PAGE. (Schagger, 2006)

	Anode buffer (10x)	**Cathode buffer (10x)**	**Gel buffer (3x)**
Tris (M)	1.0	1.0	3.0
Tricine (M)	-	1.0	-
HCl (M)	0.225	-	1.0
SDS (%)	-	1.0	0.3
pH	8.9	~8.25	8.45

3.4.4 Western blotting

For Western blot, tricine-SDS-PAGE was performed without staining the gel. The Western blot sandwich was assembled as instructed by the manufacturer (Bio-Rad) using two sponges, four filter papers and one nitrocellulose blotting membrane (BioTrace NT from Life Sciences). All the components of the sandwich were pre-wet in transfer buffer before sandwich assembly. The sandwich was assembled in the transfer device and covered with transfer buffer. The transference was carried out at 300 mA for 1.5h. The membrane was then blocked for 1h with a solution of 5% non-fat milk in PBS-Tween (PBS with 0.5% Tween 20). The milk was removed and the membrane was incubated with the primary antibody anti poly-His in mouse (Santa Cruz Biosciences) (1:500 in 5% non-fat milk in PBS-Tween) overnight at 4°C. After removal of the primary antibody, the membrane was washed three times with PBS-Tween, 5 min each wash. The membrane was incubated 1h with the secondary antibody goat anti-mouse IgG-HRP (Santa Cruz Biosciences) (1:2000 in PBS-Tween) and washed five times with PBS-Tween, 5 min each wash.

The membrane was covered with SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific), placed inside a cassette (Amersham Biosciences) wrapped in plastic and exposed to X-ray film (Amersham Hyperfilm ECL from GE Healthcare) for the desired times and processed.

Table 10 – PBS 1x composition.

PBS 1x	Concentration (mM)
NaCl	137
KCl	2.7
$Na_2HPO_4.2H_2O$	4.3
KH_2PO_4	1.47

Table 11 – Transfer buffer composition for a final volume of 1 L.

Transfer buffer	Composition
Trizma-base (mM)	48
Glycine (mM)	39
SDS (%)	0.04
Methanol (%)	20

3.4.5 Peptide purification

After superexpression, cells were collected in 500 mL centrifuge tubes and centrifuged 5 min at 8000 rpm and 4°C. The pellet was ressuspended in START buffer (4 mL of START buffer for 100 mL of starting cell culture). If not used immediately, the ressuspended pellets were kept at -80°C until further use.

The cells were sonicated with a Branson sonifier 250 (9 cycles, with 5 min intervals, of 15 bursts, duty cycle of 50% and output control 9).

The sonicated cells were centrifuged 5 min at 4°C and 17,600 g, the supernatant collected and centrifuged 1h at the same conditions. The final supernatant was then filtered with a Whatman puradisc 0.2 µm filter (GE Healthcare).

A Ni-NTA (nickel nitrilotriacetic acid) purification 1 mL column (His Trap HP from GE Healthcare) was equilibrated with 5 mL of START buffer, loaded with the filtered supernatant and washed with 15 mL START buffer. The column was eluted progressively with 5 mL of elution buffer with increasing concentrations of imidazole (20, 40, 60, 100, 200, 300, 400 and 500 mM). The eluted fractions containing the purified peptides were collected.

Table 12 – Phosphate buffer 8x composition.

	Phosphate buffer 8x
$Na_2HPO_4.2H_2O$ (M)	0.08
$NaH_2PO_4.H_2O$ (M)	0.08
NaCl (M)	4
pH	7.4

Table 13 – START and elution buffer compositions.

	START buffer	Elution buffers
Phosphate buffer 8x (mL)	12	3
Imidazole (mM)	10	20; 40; 60; 100; 200; 300; 400; 500
Water until final volume of (mL)	100	24
pH	7.4	7.4

The collected fractions were desalted through size-exclusion chromatography using the ÄKTA prime apparatus (Amersham Biosciences) or centrifuging (3000 rpm at 4°C) several times in a 3 kDa cut-off column (Ultracell-3K Amicon Ultra centrifugal filter from Millipore). In the centrifugal filter, the fractions are concentrated until a volume about 5 mL, then the concentrated protein is washed several times with phosphate buffer to desalt through centrifugation. In Table 14 is shown the elution program using the ÄKTA apparatus with a HiPrep 26/10 Desalting column (GE Healthcare).

Table 14 – ÄKTA elution program for protein sample desalting through molecular exclusion chromatography.

Step	Volume (mL)	Flow rate (mL/min)	Inject valv
Tube wash	50	40	waste
Column wash	5	5	load
Protein injection	Protein sample volume	5	inject
Protein elution	100	5	load
Column wash	30	5	load

The ÄKTA eluted fractions were collected according to the chromatogram given by the program that measures the desalted solution OD_{280nm}, adequate for protein detection.

When using the ÄKTA apparatus, the collected fractions from the desalting step were concentrated in a 3 kDa cut-off column (Microsep Advance Centrifugal Device 3K MWCO from Pall Corporation) by centrifuging at 3000 rpm (at 4°C) for several times until the volume was approximately 1 mL. The concentrated solution was kept at 4°C until further use.

3.5 Spectroscopic analysis

Purified peptides structure was analyzed through spectroscopic analysis. UV-visible and far UV Circular Dichroism (CD) spectra were traced.

UV-visible spectra were obtained using a PharmaSpec UV-1700 (Shimadzu) UV-visible spectrophotometer. Spectra cover wave lengths from 250 to 800 nm.

Far UV CD spectra were traced using a π*-180 spectropolarimeter from Applied Photophysics using the parameters shown in Table 15. For each peptide 10 measures were made.

The obtained spectra were analyzed using the online DICHROWEB server in order to predict the secondary structure of the peptides (Whitmore and Wallace, 2004).

Table 15 – Spectropolarimeter parameters used for circular dichroism analysis.

	Value
Time per point	1s
Path length	10
Wavelength range	190 to 250 nm
Step size	1 nm
Bandwidth	2 nm
Number of repeats	10

4 Results

4.1 DNA manipulation and cloning

The strategy to produce azurin derived peptides used in this work was to isolate the respective genes through PCR, clone them into pWH884 and transform the constructs into *E. coli* strains adequate for protein expression.

The main relevant characteristics for gene cloning and protein expression of each azurin derived peptide are present in Table 16.

Table 16 – Peptide DNA and protein sequence and their respective sizes.

Azu 1-50	DNA sequence	GCC GAG TGC TCG GTG GAC ATC CAG GGT AAC GAC CAG ATG CAG TTC AAC ACC AAT GCC ATC ACC GTC GAC AAG AGC TGC AAG CAG TTC ACC GTC AAC CTG TCC CAC CCC GGC AAC CTG CCG AAG AAC GTC ATG GGC CAC AAC TGG GTA CTG
	DNA size	150 bp
	Aminoacid sequence	AECSVDIQGNDQMQFNTNAITVDKSCKQFTVNLSHPGNLPKNVMGHN WVL
	Peptide size	5.55 kDa
Azu 1-77	DNA sequence	GCC GAG TGC TCG GTG GAC ATC CAG GGT AAC GAC CAG ATG CAG TTC AAC ACC AAT GCC ATC ACC GTC GAC AAG AGC TGC AAG CAG TTC ACC GTC AAC CTG TCC CAC CCC GGC AAC CTG CCG AAG AAC GTC ATG GGC CAC AAC TGG GTA CTG AGC ACC GCC GCC GAC ATG CAG GGC GTG GTC ACC GAC GGC ATG GCT TCC GGC CTG GAC AAG GAT TAC CTG AAG CCC GAC GAC
	DNA size	231bp
	Aminoacid sequence	AECSVDIQGNDQMQFNTNAITVDKSCKQFTVNLSHPGNLPKNVMGHN WVLSTAADMQGVVTDGMASGLDKDYLKPDD
	Peptide size	8.33 kDa
Azu 80-128	DNA sequence	CAC ACC AAG CTG ATC GGC TCG GGC GAG AAG GAC TCG GTG ACC TTC GAC GTC TCC AAG CTG AAG GAA GGC GAG CAG TAC ATG TTC TTC TGC ACC TTC CCG GGC CAC TCC GCG CTG ATG AAG GGC ACC CTG ACC CTG AAG TGA TGC
	DNA size	144 bp
	Aminoacid sequence	IAHTKLIGSGEKDSVTFDVSKLKEGEQYMFFCTFPGHSALMKGTLTLK
	Peptide size	5.29 kDa

4.1.1 PCR

In order to obtain the DNA fragments of azurin gene coding the three desired peptides, PCR was performed.

The amplification of the desired fragments was carried out using the designed primers and the program shown in section 3.3.1. As shown in Table 16, the azu1-50, azu 1-77 and azu 80-128 sizes are expected to be 150, 231 and 144 bp, respectively. The PCR results are shown in Figure 4.

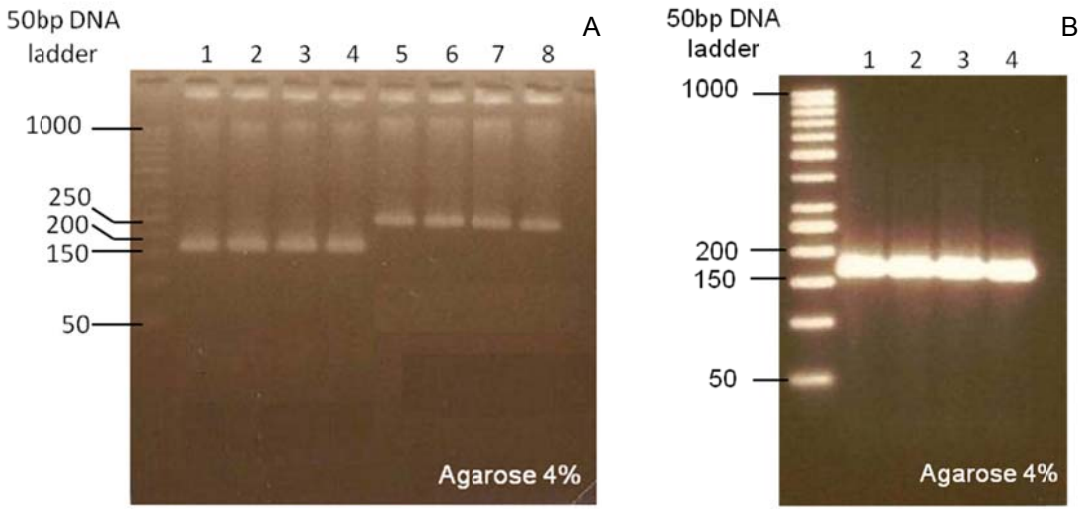

Figure 4 – 4% agarose gel with the amplified fragments by PCR. A: lane 1 to 4 - azu 1-50 (about 150 bp); lane 5 to 8 – azu 1-77 (about 230 bp); B: lane 1 to 4 – azu 80-128 (about 150 bp).

Observing Figure 4, it is verified that all fragments amplified with the expected size. azu 1-50 and azu 80-128 DNA bands are slightly above the 150 bp band, the azu 1-77 DNA band is between the 200 and 250 bp bands. Sizes are slightly above those indicated in Table 16 because the primers used in the PCR add a small portion of DNA to the N- and C-terminal intended to enable DNA restriction and cloning.

After confirming that the DNA fragments were at the correct size, fragments were extracted from the agarose gel and proceeded to cloning into pWH844 plasmid.

4.1.2 Transformations

PCR products were cloned into the pWH844 plasmid after digestion with BamHI and HindIII, resulting in the pTEazu150, pTEazu177 and pTEazu80128 constructions shown in Figure 5, Figure 6 and Figure 7. The selection marker is ampicillin and, because it has a T5 promoter, the induction is performed using IPTG and plasmid sizes are about 5 kb.

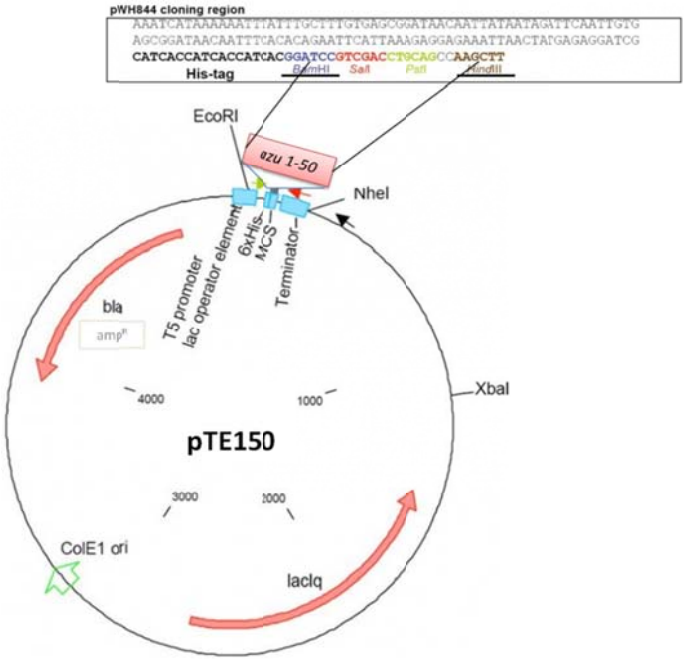

Figure 5 – Schematic representation of pTE150 plasmid. The plasmid is based on the cloning of azu 1-50 gene in the pWH844 vector using the BamHI and HindIII restriction enzymes. The translated peptide will have a poly-Histidine tag in its N-terminal. The selection marker is ampicillin and the T5 promoter is induced by IPTG.

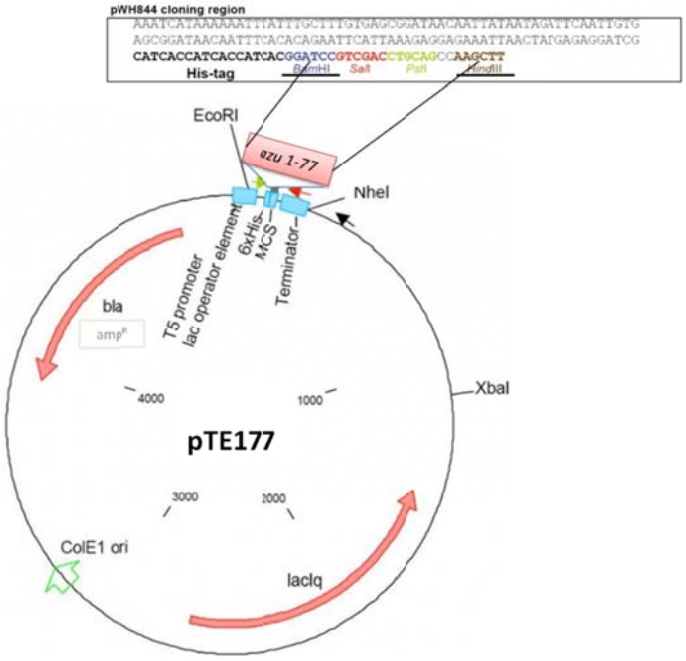

Figure 6 – Schematic representation of pTE177 plasmid. The plasmid is based on the cloning of azu 1-77 gene in the pWH844 vector using the BamHI and HindIII restriction enzymes. The translated peptide will have a poly-Histidine tag in its N-terminal. The selection marker is ampicillin and the T5 promoter is induced by IPTG.

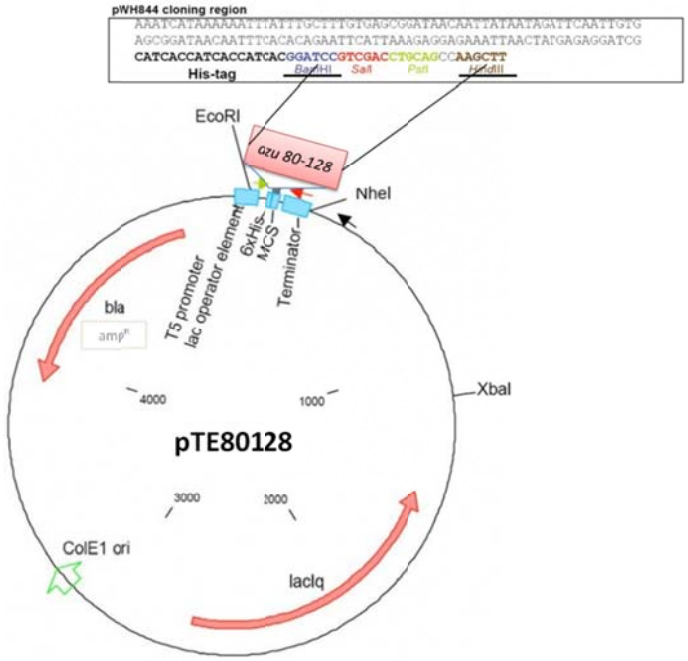

Figure 7 – Schematic representation of pTE80128 plasmid. The plasmid is based on the cloning of azu 80-128 gene in the pWH844 vector using the BamHI and HindIII restriction enzymes. The translated peptide will have a poly-Histidine tag in its N-terminal. The selection marker is ampicillin and the T5 promoter is induced by IPTG.

E. coli XL1-Blue was transformed with the constructions through classical transformation. For each transformation, twelve candidates were selected and colony PCR was performed to confirm transformation in the selected candidates. The colony PCR resulting agarose gels are presented in Figure 8, Figure 9 and Figure 10.

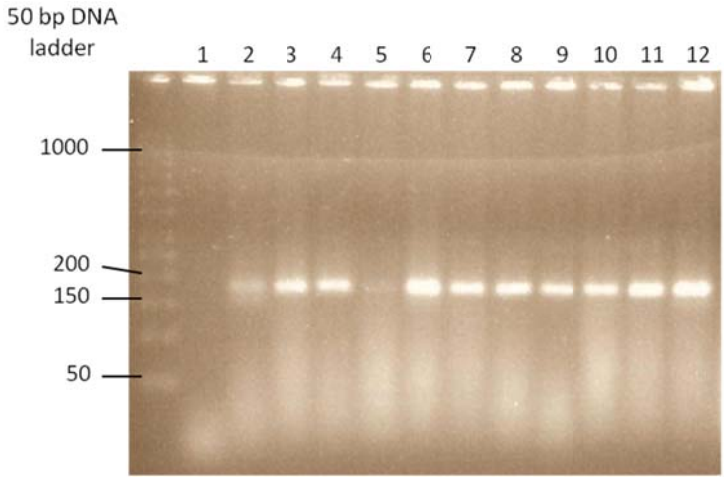

Figure 8 – 4% agarose gel with Colony PCR results over the twelve selected transformant azu 1-50 candidates. The number of the lane corresponds to the number of the candidate. Lane 1 has no amplification, which doesn't confirm transformation. Lane 2 and 5 have weak bands at the expected size. The other lanes all have good amplification signals at the expected size, which means that they are true transformants. Among these transformants, transformant 6, at lane 6 was selected for further work.

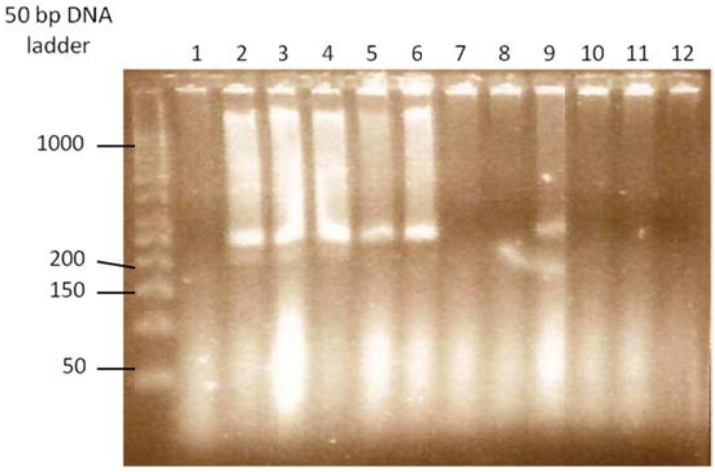

Figure 9 – 4% agarose gel with Colony PCR results over the twelve selected transformant azu 1-77 candidates. The number of the lane corresponds to the number of the candidate. Lane 1, 7, 8, 10, 11 and 12 have no amplification, which doesn't confirm transformation. Lane 9 has a weak band at the expected size. Lanes 2 to 6 have good amplification signals at the expected size, which means that they are true transformants. Among these transformants, transformant 6, at lane 6 was selected for further work.

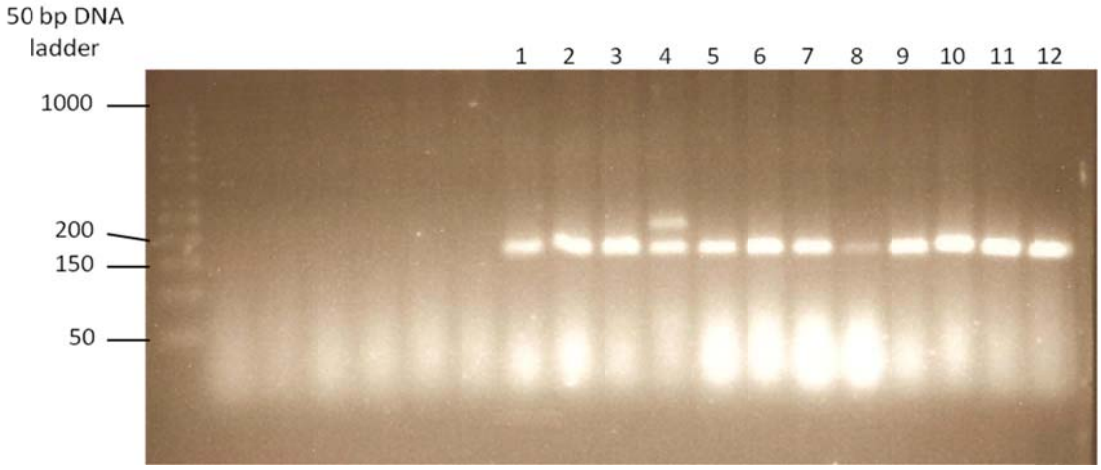

Figure 10 – 4% agarose gel with Colony PCR results over the twelve selected transformant azu 80-128 candidates. The number of the lane corresponds to the number of the candidate. All candidates present a band at the expected size, although, lane 4 has an unspecific amplification at 200 bp and lane 8 has a weak band. Among all the other true transformants, transformant 6 at lane 6 was selected for further work.

In these figures it is observed that not all selected candidates gave a positive colony PCR, which means that possibly not all candidates are true transformants. To each construction, one true candidate was selected giving attention to ensure that the colony PCR band is at the correct size and has no unspecific amplifications. The selected transformants were the candidate "6" in all transformations.

To confirm that the selected transformants have the inserted plasmids, the plasmids were extracted using the QIAprep Spin Miniprep Kit and an agarose gel was run with the extracted samples (Figure 11).

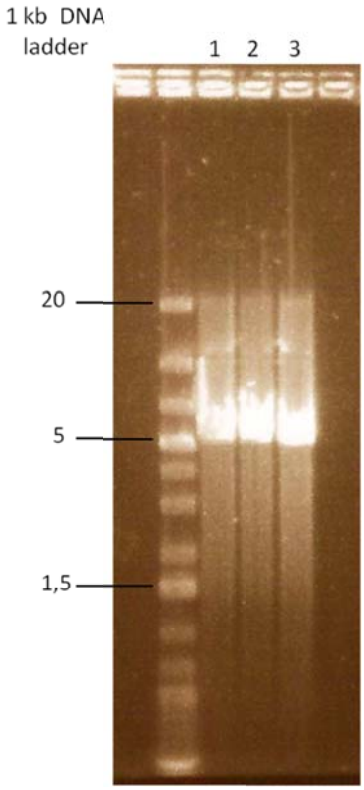

Figure 11 – 0,8% agarose gel with the extracted plasmids from the selected transformants. Lane 1 – pTE150, lane 2 – pTE177, lane 3 – pTE80128. In the three cases the plasmid has a strong signal at the expected size (approximately 5 kb).

Figure 11 shows that the plasmid is present in the three selected transformants and it has the expected size (about 5 kb).

4.2 Peptide production

Having the confirmation that the constructions were successfully cloned into *E. coli* XL1-Blue, the next step was to transform *E. coli* expression strains with these constructions in order to have peptide production.

The starting expression strains tested were *E. coli* BL21(DE3) and *E. coli* SURE. *E. coli* BL21(DE3) was selected too because is one of the most successful strain used for protein expression and needs little levels of inducer to produce high levels of protein. *E. coli* SURE was selected because it is indicated for cloning unstable plasmids. The fact that the constructed plasmids have fragments of the azurin gene and not a complete gene may create unstability, justifying the use of *E. coli* SURE.

4.2.1 Superexpression optimization

Transformed *E. coli* SURE and *E. coli* BL21(DE3) were grown in SB medium supplemented with different concentrations of IPTG at 37°C for 5h and at 20°C overnight (16h). 16% SDS PAGE gels and Western blots were run to determine the optimal conditions for peptides superexpression.

Coomassie blue stained gels of Azu 1-50 expression assays are shown in Figure 12.

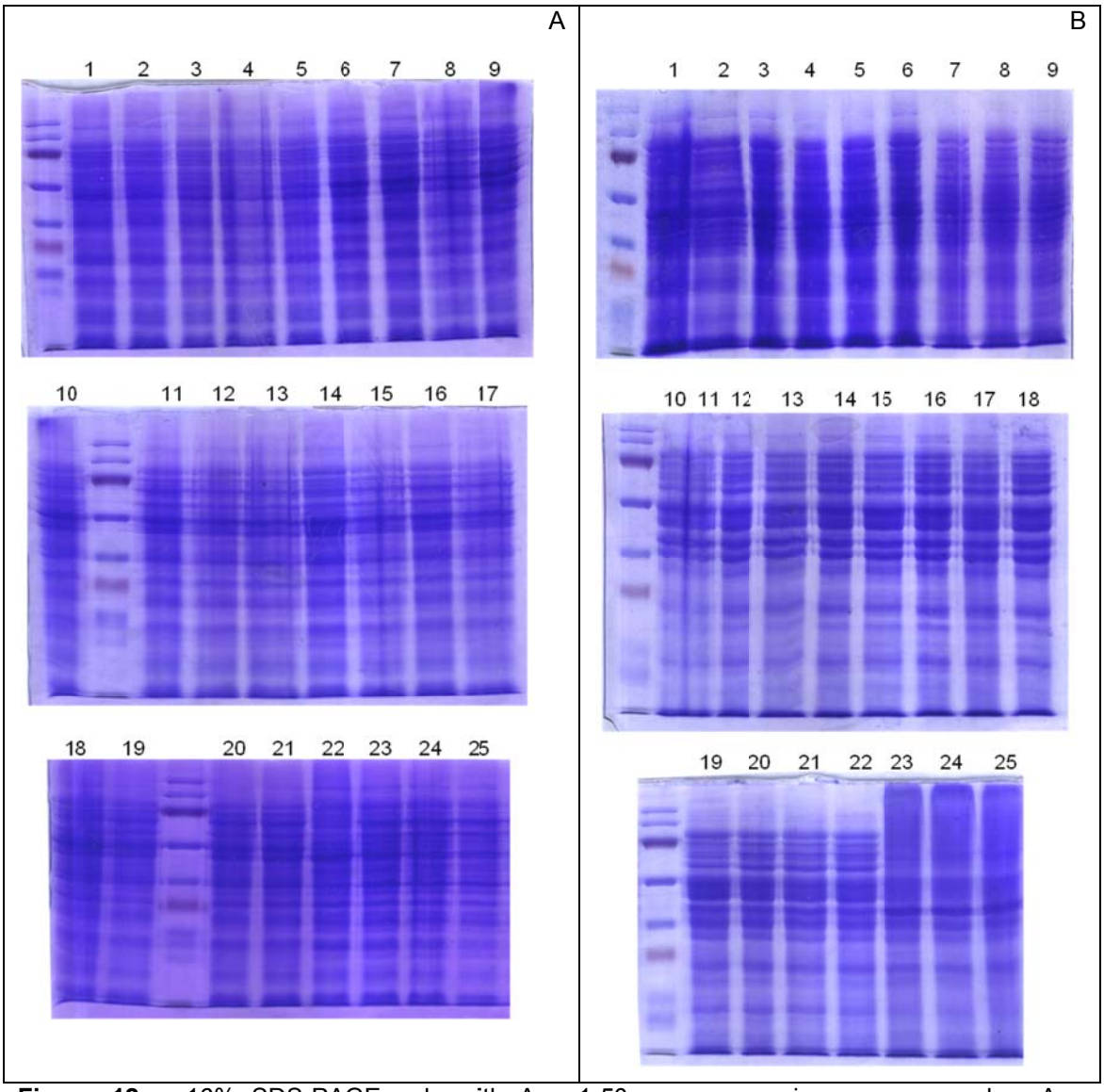

Figure 12 – 16% SDS-PAGE gels with Azu 1-50 superexpression assay samples. A – Superexpression assay samples obtained from E. coli SURE; B - Superexpression assay samples obtained from E. coli BL21(DE3). Lane description: 1- 0 mM IPTG t=0h (37°C); 2- 0 mM IPTG t=1h (37°C); 3- 0.1 mM IPTG t=1h (37°C); 4- 0.2 mM IPTG t=1h (37°C); 5- 0.3 mM IPTG t=1h (37°C); 6- 0 mM IPTG t=2h (37°C); 7- 0.1 mM IPTG t=2h (37°C); 8- 0.2 mM IPTG t=2h (37°C); 9- 0.3 mM IPTG t=2h (37°C); 10- 0 mM IPTG t=3h (37°C); 11- 0.1 mM IPTG t=3h (37°C); 12- 0.2 mM IPTG t=3h (37°C); 13- 0.3 mM IPTG t=3h (37°C); 14- 0 mM IPTG t=4h (37°C); 15- 0.1 mM IPTG t=4h (37°C); 16- 0.2 mM IPTG t=4h (37°C); 17- 0.3 mM IPTG t=4h (37°C); 18- 0 mM IPTG t=5h (37°C); 19- 0.1 mM IPTG t=5h (37°C); 20- 0.2 mM IPTG t=5h (37°C); 21- 0.3 mM IPTG t=5h (37°C); 22- 0 mM IPTG t=16h (20°C); 23- 0.1 mM IPTG t=16h (20°C); 24- 0.2 mM IPTG t=16h (20°C); 25- 0.3 mM IPTG t=16h (20°C). In all gels, no clear superexpression band is visible.

Analyzing the SDS-PAGE gels (Figure 12), Azu 1-50 superexpression is not clearly visible, as no evident band intensifies during the superexpression assay. Azu 1-50 is expected to have about 6.5 kDa (poly-His tag adds approximately 1 kDa to the peptide size). Western blotting was performed as a more sensitive method for detecting Azu 1-50. The Western blot was performed using the samples of *E. coli* BL21(DE3) grown for 4h and overnight with different levels of IPTG induction. The samples are the same that were used in the SDS PAGE gels. Results are shown in Figure 13.

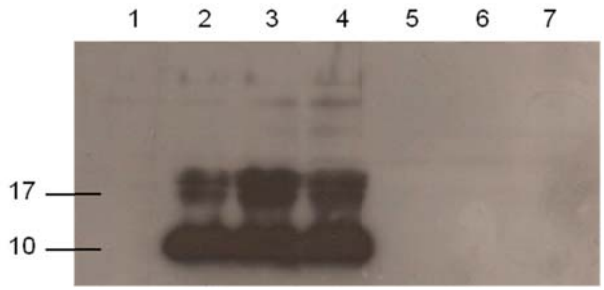

Figure 13 – Western blotting analysis of Azu 1-50 superexpression in E. coli BL21(DE3) at 37°C and 20°C and with different concentrations of IPTG. Lane description: 1 – 0 mM IPTG t=0h; 2 – 0.1 mM IPTG t=4h (37°C); 3 – 0.2 mM IPTG t=4h (37°C); 4 – 0.3 mM IPTG t=4h (37°C); 5 – 0 mM IPTG t=16h (20°C); 6 – 0.2 mM IPTG t=16h (20°C); 7 – 0.3 mM IPTG t=16h (20°C). From lane 2 to 4 several bands are visible, indicating that there is the formation of dimers and trimers. From lane 5 to 7 no band is visible, indicating that none or very few peptide was expressed over night at 20°C.

The Western blot (Figure 13) shows that, although it is not visible in SDS PAGE, Azu 1-50 is being overexpressed at 37°C but not at 20°C. It also shows that some portion of the Azu 1-50 forms dimers (higher bands at 17 kDa). The distribution of Azu 1-50 in the monomer and dimer forms partly explains why it is not visible in Coomassie blue staining SDS-PAGE.

It is also noted that the Azu 1-50 bands, more specifically the monomer band is above the expected peptide size. A possible explanation is the differential migration of peptides in a high percentage acrylamide gel comparing to bigger proteins and the applied ladder. The higher percentage of acrylamide may add a molecular exclusion effect to the peptide migration that "delays" its migration and consequently, makes us see it in a higher molecular weight.

These two figures (Figure 12 and Figure 13) show that there is expression of Azu 1-50 in *E. coli* BL21(DE3) when grown at 37°C for 4h but it is very low compared to what was desired. Through Western bloting analysis, the superexpression conditions were established as 4h growth with 0.2 mM IPTG induction in *E. coli* BL21(DE3).

Coomassie blue stained gels of Azu 1-77 expression assays are shown in Figure 14.

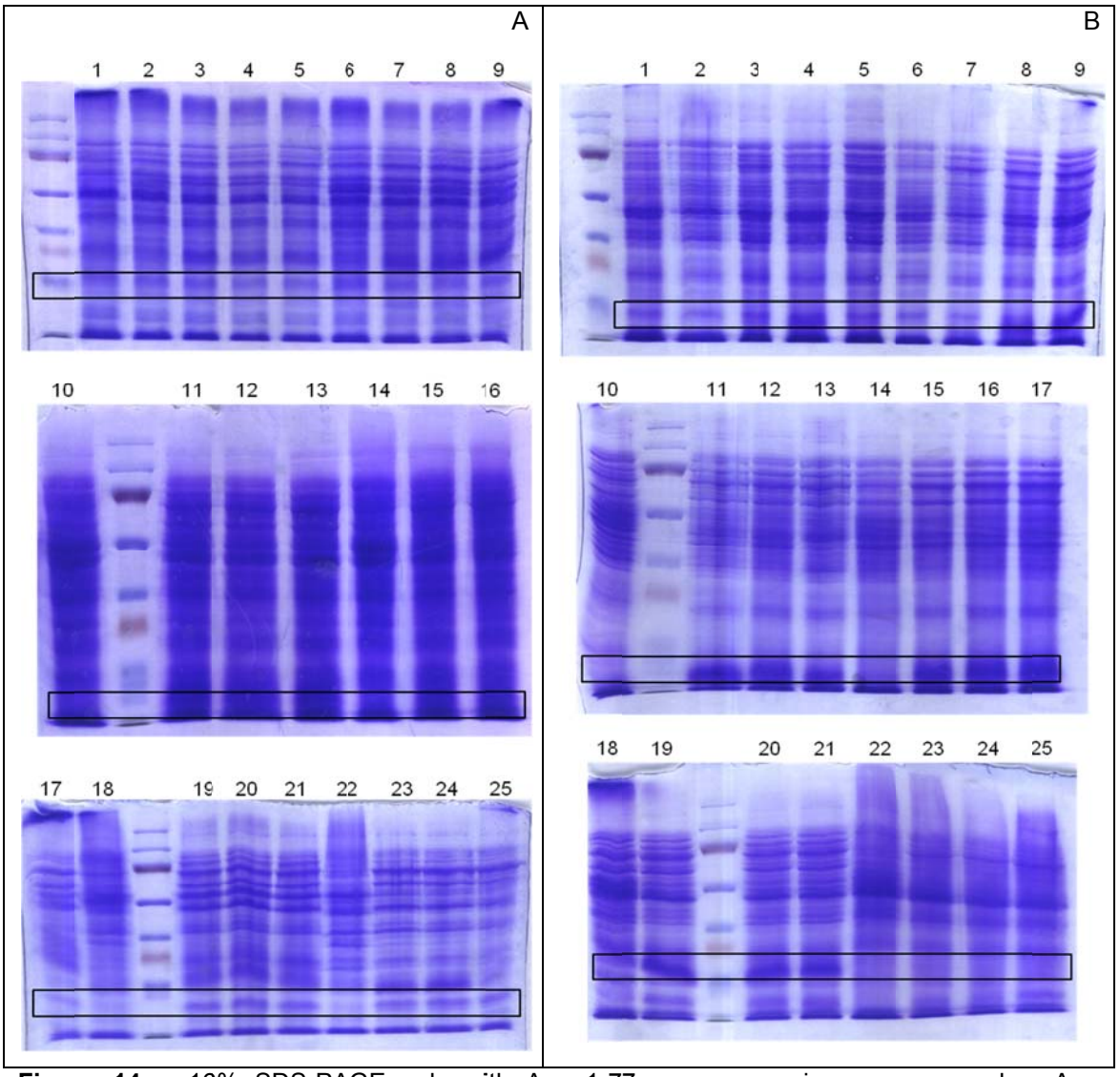

Figure 14 – 16% SDS-PAGE gels with Azu 1-77 superexpression assay samples. A – Superexpression assay samples obtained from *E. coli* SURE; B - Superexpression assay samples obtained from *E. coli* BL21(DE3). Lane description: 1- 0 mM IPTG t=0h (37°C); 2- 0 mM IPTG t=1h (37°C); 3- 0.1 mM IPTG t=1h (37°C); 4- 0.2 mM IPTG t=1h (37°C); 5- 0.3 mM IPTG t=1h (37°C); 6- 0 mM IPTG t=2h (37°C); 7- 0.1 mM IPTG t=2h (37°C); 8- 0.2 mM IPTG t=2h (37°C); 9- 0.3 mM IPTG t=2h (37°C); 10- 0 mM IPTG t=3h (37°C); 11- 0.1 mM IPTG t=3h (37°C); 12- 0.2 mM IPTG t=3h (37°C); 13- 0.3 mM IPTG t=3h (37°C); 14- 0 mM IPTG t=4h (37°C); 15- 0.1 mM IPTG t=4h (37°C); 16- 0.2 mM IPTG t=4h (37°C); 17- 0.3 mM IPTG t=4h (37°C); 18- 0 mM IPTG t=5h (37°C); 19- 0.1 mM IPTG t=5h (37°C); 20- 0.2 mM IPTG t=5h (37°C); 21- 0.3 mM IPTG t=5h (37°C); 22- 0 mM IPTG t=16h (20°C); 23- 0.1 mM IPTG t=16h (20°C); 24- 0.2 mM IPTG t=16h (20°C); 25- 0.3 mM IPTG t=16h (20°C). Superexpression bands are present approximately at the 17 kDa size, which indicates that may exist the formation of dimers.

Observing the Figure 14, Azu 1-77 expression is visible in Coomassie blue stained SDS-PAGE gels, which by itself shows that Azu 1-77 is being expressed at higher levels than Azu 1-50. It is also visible that Azu 1-77 expression has higher levels in *E. coli* BL21(DE3) than in *E.coli* SURE. When analyzing *E. coli* SURE expression assays it is visible that Azu 1-77 is expressed both in 37°C and 20°C growth. In E. coli BL21(DE3) gels, Azu 1-77 expression is only visible in 37°C growth and different concentrations of IPTG upon induction do not generate great differences in expression levels. Expression is only clearly visible at 3h, 4h and 5h growth.

To confirm the expression of Azu 1-77, Western blotting was performed using *E. coli* BL21(DE3) grown for 4h and overnight with different levels of IPTG induction. The samples were the same used in SDS-PAGE. Results are shown in Figure 15.

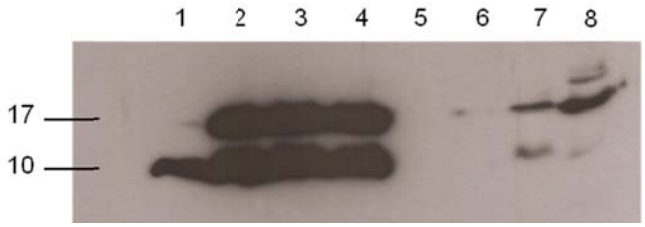

Figure 15 – Western Blotting analysis of Azu 1-77 superexpression in E. coli BL21(DE3) at 37°C and 20°C and with different concentrations of IPTG. Lane description: 1 – 0 mM IPTG t=0h; 2 – 0.1 mM IPTG t=4h (37°C); 3 – 0.2 mM IPTG t=4h (37°C); 4 – 0.3 mM IPTG t=4h (37°C); 5 – 0 mM IPTG t=16h (20°C); 6 – 0.1 mM IPTG t=16h (20°C); 7 – 0.2 mM IPTG t=16h (20°C); 8 – 0.3 mM IPTG t=16h (20°C). The presence of two bands in the gel suggests that Azu 1-77 may form dimers or trimers when in higher concentrations (as in lane 1 only one band is visible).

The Western blot obtained for Azu 1-77 expression shows that the identified band in Coomassie blue stained SDS-PAGE is effectively Azu 1-77 and that there is a very low expression of Azu 1-77 at 20°C. It also shows that it is not the only band present. Like Azu 1-50, Azu 1-77 forms dimers when in higher concentrations. The SDS-PAGE visible bands correspond to the dimers band (at 17 kDa). The monomer band (at 10 kDa) is at the expected size. Because Azu 1-77 is bigger than Azu 1-50, the phenomenon hypothesized for Azu 1-50 is less relevant and Azu 1-77 can migrate more like the ladder proteins.

Because *E. coli* BL21(DE3) demonstrated to achieve higher levels of Azu 1-77 expression, it was selected as the expression host. The selected conditions were a 4h growth at 37°C with 0,2 mM IPTG induction.

Coomassie blue stained gels of Azu 80-128 expression assays are shown in Figure 16.

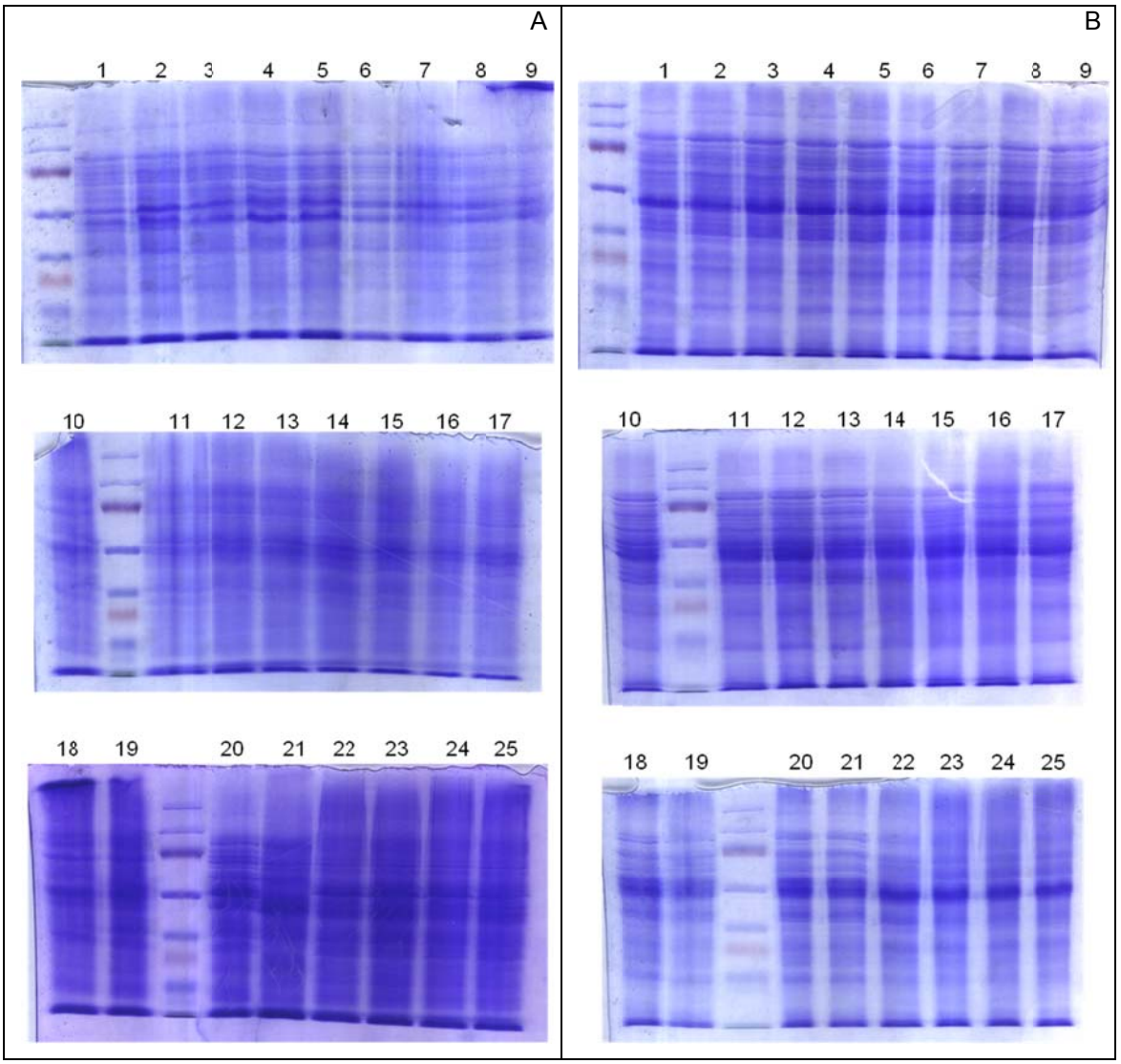

Figure 16 – 16% SDS-PAGE gels with Azu 80-128 superexpression assay samples. A – Superexpression assay samples obtained from *E. coli* SURE; B - Superexpression assay samples obtained from *E. coli* BL21(DE3). Lane description: 1- 0 mM IPTG t=0h (37°C); 2- 0 mM IPTG t=1h (37°C); 3- 0.1 mM IPTG t=1h (37°C); 4- 0.2 mM IPTG t=1h (37°C); 5- 0.3 mM IPTG t=1h (37°C); 6- 0 mM IPTG t=2h (37°C); 7- 0.1 mM IPTG t=2h (37°C); 8- 0.2 mM IPTG t=2h (37°C); 9- 0.3 mM IPTG t=2h (37°C); 10- 0 mM IPTG t=3h (37°C); 11- 0.1 mM IPTG t=3h (37°C); 12- 0.2 mM IPTG t=3h (37°C); 13- 0.3 mM IPTG t=3h (37°C); 14- 0 mM IPTG t=4h (37°C); 15- 0.1 mM IPTG t=4h (37°C); 16- 0.2 mM IPTG t=4h (37°C); 17- 0.3 mM IPTG t=4h (37°C); 18- 0 mM IPTG t=5h (37°C); 19- 0.1 mM IPTG t=5h (37°C); 20- 0.2 mM IPTG t=5h (37°C); 21- 0.3 mM IPTG t=5h (37°C); 22- 0 mM IPTG t=16h (20°C); 23- 0.1 mM IPTG t=16h (20°C); 24- 0.2 mM IPTG t=16h (20°C); 25- 0.3 mM IPTG t=16h (20°C). In all gels, no clear superexpression band is visible.

Like for Azu 1-50, for Azu 80-128 no superexpression bands were visible in the Coomassie blue stained SDS-PAGE gels. Because SDS-PAGE gels were inconclusive, Western blotting analysis of the same samples was performed. Western blotting analysis did not show any signal in any sample, which demonstrate that no expression of Azu 80-128 was achieved. No further work was done with this peptide because no expression was achieved.

To confirm that the bands identified as dimers in both Azu 1-50 and Azu 1-77 Western blots (Figure 13 and Figure 15) were in fact dimers and not aggregates, Western blotting analysis

using more denaturant conditions was performed. To achieve more denaturant conditions, the samples were added 2x SDS gel-loading buffer supplemented with 8 M urea and the SDS-PAGE had 10 mM urea content. The resulting Western blots are shown in Figure 17.

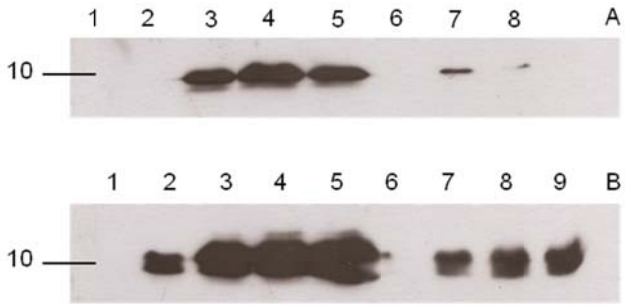

Figure 17 – Western blot analysis of Azu 1-50 (A) and Azu 1-77 (B) superexpression under denaturant conditions. Lane descripton of A: 1 – 0 mM IPTG t=0h; 2 – 0 mM IPTG t=4h; 3 – 0.1 mM IPTG t=4h (37°C); 4 – 0.2 mM IPTG t=4h (37°C); 5 – 0.3 mM IPTG t=4h (37°C); 6 – 0 mM IPTG t=16h (20°C); 7 – 0.1 mM IPTG t=16h (20°C); 8 – 0.2 mM IPTG t=16h (20°C); 9 – 0.3 mM IPTG t=16h (20°C). Lane descripton of B: 1 – 0 mM IPTG t=0h; 2 – 0 mM IPTG t=4h; 3 – 0.1 mM IPTG t=4h (37°C); 4 – 0.2 mM IPTG t=4h (37°C); 5 – 0.3 mM IPTG t=4h (37°C); 6 – 0 mM IPTG t=16h (20°C); 7 – 0.2 mM IPTG t=16h (20°C); 8 – 0.3 mM IPTG t=16h (20°C). The presence of only one band in gel A and the proximity of the bands in gel B demonstrate that there is the formation of dimers.

In Figure 17 is visible that in the case of Azu 1-50 only one band is visible, indicating that all the dimers were denatured and all Azu 1-50 is in monomers. In the case of Azu 1-77 there are still two bands visible but very close to each other. Because Azu 1-77 is present in higher concentrations and because the used conditions do not totally denaturate all the proteins, the proximity of the two bands indicate that the higher band correspond to dimers that could not be totally denaturated. From Figure 17 it is concluded that both Azu 1-50 and Azu 1-77 form dimers and that these dimers can be reduced to monomers over denaturant conditions.

4.2.2 Purification

Because low levels of peptide expression are achieved, for purification big volumes of cell culture were needed, from 1 to 2L.

As described in section 3.4.5, bacterial cells were incubated 4h at 37°C with 0.2 mM IPTG induction, centrifuged and ressuspended in START buffer. After confirmation of peptide expression through Western blotting, the bacterial ressuspended pellets were sonicated. After sonication, cell extracts were again centrifuged and filtered. The filtered cell extract were loaded into a 1 mL Ni-NTA column for peptide purification. The imidazole range used was from 20 to 500 mM.

For Azu 1-50 the superexpression sample and eluted fractions from the Ni-NTA column were analyzed by SDS-PAGE and Western blotting (Figure 18).

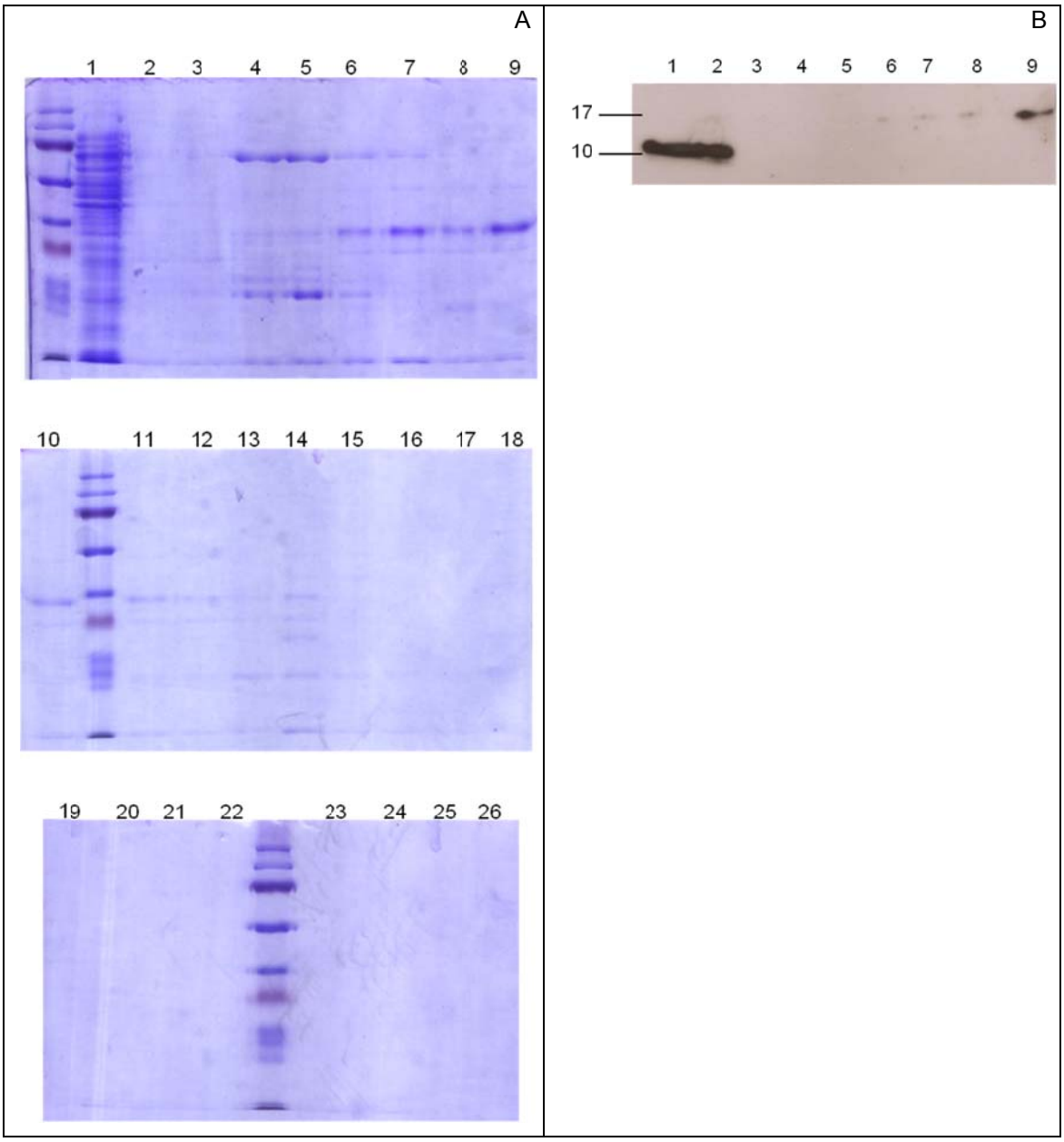

Figure 18 – 16% SDS-PAGE (A) and Western blot (B) of Azu 1-50 purification in Ni-NTA column. Lane description is the same for figure A: 1 – superexpression bacterial culture at 4h (0.2 mM IPTG); lane 2 and 3 – fractions 2 and 3 of 20 mM imidazole elution; lane 4 and 5 – fractions 2 and 3 of 40 mM imidazole elution; lane 6 and 7 – fractions 2 and 3 of 60 mM imidazole elution; lanes 8 to 12 – fractions 1, 2, 3, 4 and 5 of 100 mM imidazole elution; lanes 13 to 17 – fractions 1, 2, 3, 4 and 5 of 200 mM elution; lanes 18 to 22 – fractions 1, 2, 3, 4 and 5 of 300 mM imidazole elution; lane 23 and 24 – fractions 2 and 3 of 400 mM imidazole elution; lane 25 and 26 – fractions 2 and 3 of 500 mM imidazole elution. Lane description for figure B: 1 and 2– superexpression of two distinctive bacterial culture at 4h (0,2 mM IPTG); lane 3 to 7 – fractions 1 to 5 of 100 mM imidazole elution; lane 8 and 9 – fractions 1 and 2 of 200 mM imidazole elution. Western blot of the missing fractions (from fraction 3 of 200 mM imidazole to fraction 5 of 500 mM imidazole) did not show any signal.

Figure 18 B show very weak signals of Azu 1-50 at the collected fractions from the Ni-NTA column. Azu 1-50 only is detected in fractions 4 and 5 from 100 mM imidazole elution and 1 and 2 fractions from 200 mM imidazole elution. These signals are at the dimers molecular weight previously observed in Figure 13 (17 kDa) but the signals given in superexpression samples are at the monomer molecular weight (10 kDa) also observed in Figure 13. The presence of

monomers and absence of dimers in the superexpression samples show that the expression level was lower than what would be expected to achieve. However, the exit of Azu 1-50 from the Ni-NTA column in form of dimers, show that the relative concentration of Azu 1-50 in the column is somewhat high, otherwise no dimers would be formed.

Observing the Figure 18 A at the expected size of Azu 1-50 given by the Western blot, 17 kDa, there are visible bands in more collected fractions than the detected by Western blot. Bands at 17 kDa are visible in lanes 3 to 6, lanes 11 to 18. It is expected that these bands correspond to Azu 1-50 because these are obtained apparently pure from lane 15 to 18 and are present in almost all collected fractions. Apparently Western blot does not recognize these lanes as Azu 1-50. From lane 19 to 26, no protein is seen in Coomassie blue stained SDS-PAGE gel or in Western blot therefore apparently no Azu 1-50 is present in these fractions.

Azu 1-50 is collected pure in lanes 15 to 18 (fractions 3, 4 and 5 of 200 mM imidazole and 1 of 300 mM imidazole). From all the collected fractions, fractions 200, 300 and 400 mM imidazole were desalted using the ÄKTA apparatus.

In a first approach, 1L of bacterial culture was purified and loaded into the ÄKTA apparatus. The resulting chromatogram is shown in Figure 19.

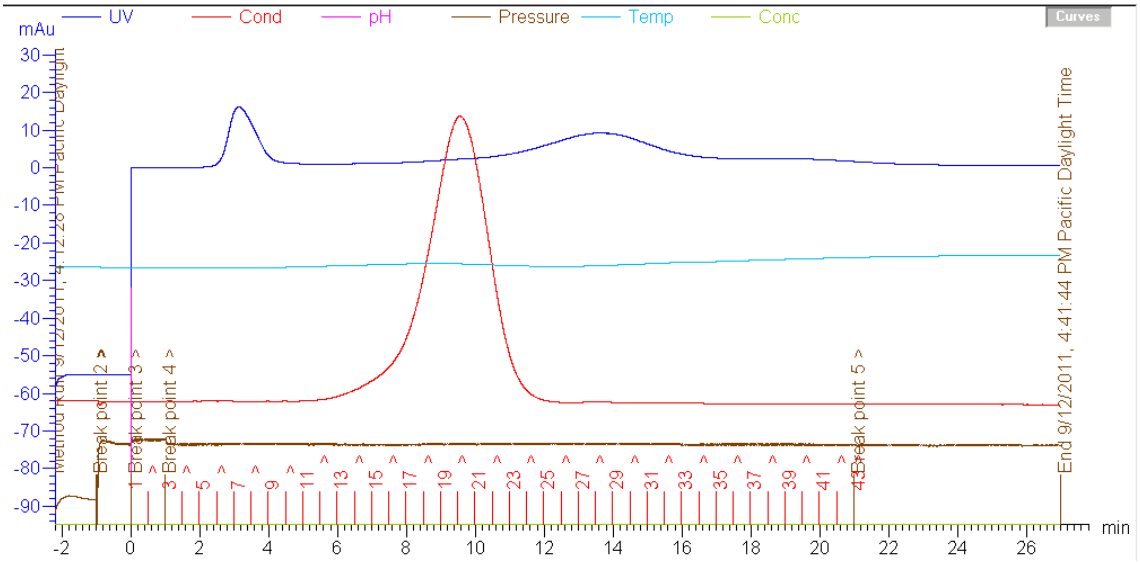

Figure 19 – ÄKTA chromatogram obtained from purification of 1L of cell culture expressing Azu 1-50. Absorbance curve at 280 nm (dark blue line) indicate that Azu 1-50 is present in fractions 6 to 9, having its higher concentration at fraction 7. The absorbance peak is approximately 15 mAu. The latter peak corresponds to the imidazole presence.

It is observed in Figure 19 that the 280 nm peak is well resolved and is about 15 mAu. Usual peaks obtained for azurin are broader and are 200 to 300 mAu. Comparing the values, the final Azu 1-50 obtained has very low concentration compared to azurin obtained through the same method of purification and similar expression method. In order to obtain higher concentrations, expression was upscaled to 2L. The resulting chromatogram of 2L cell culture purification is shown in Figure 20.

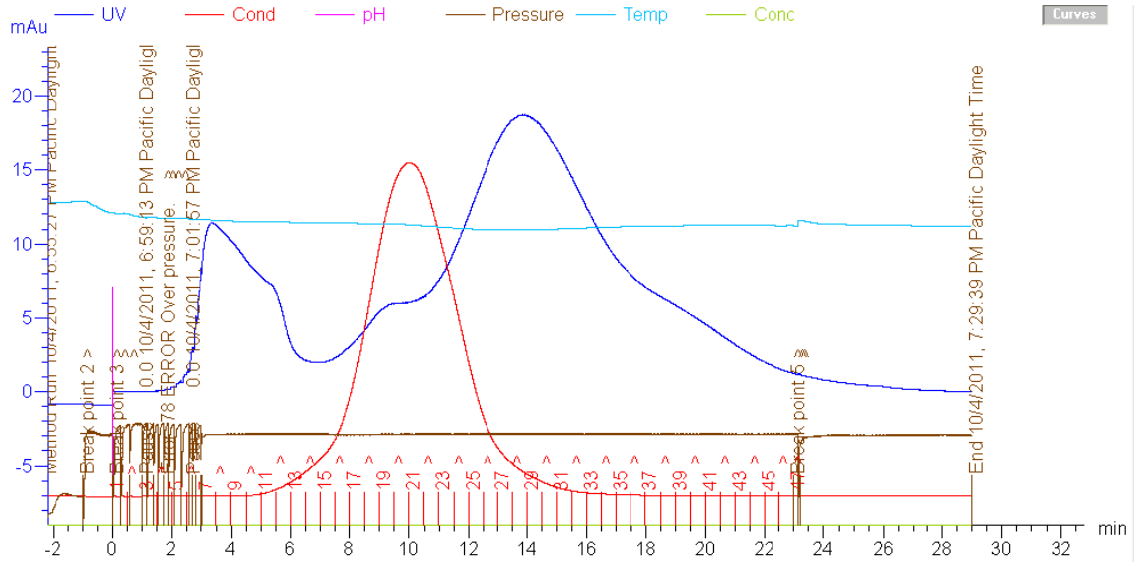

Figure 20 – ÄKTA chromatogram obtained from purification of 2L of cell culture expressing Azu 1-50. Absorbance curve at 280 nm (dark blue line) indicate that Azu 1-50 is present in fractions 7 to 13, having its higher concentration at fraction 7. The absorbance peak is approximately 12 mAu. The latter peak corresponds to the imidazole presence.

Comparing Figure 20 to Figure 19 it is observed that the level of Azu 1-50 peak is slightly lower but is also broader. This indicates that a higher quantity of Azu 1-50 was achieved but not as much as it was expected, *i. e.*, Azu 1-50 purification is not upscalable in the conditions tested.

Fractions 7 to 12 were collected and analyzed through Western Blotting analysis. The resulting Western blot is shown in Figure 21.

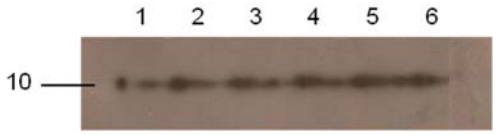

Figure 21 – Collected fractions from ÄKTA of the purification of Azu 1-50 from 2L of cell culture. Lane description: lane 1 – fraction 7; lane 2 – fraction 8; lane 3 – fraction 9; lane 4 – fraction 10; lane 5 – fraction 11; lane 6 – fraction 12.

It is confirmed in Figure 21 the presence of Azu 1-50 and is observed that the fractions with higher concentrations of the peptide are the latter collected fractions. The intensities of the bands are very low comparing to the intensity of the samples of superexpression samples, showing that during purification lots of peptide is lost.

Comparing Figure 21 to Figure 18, it is confirmed that Azu 1-50 is being collected and that the Western blot signals in Figure 18 B are too weak and do not totally represent Azu 1-50 presence. The lack of Azu 1-50 signal may be because, when exiting the Ni-NTA column, Azu 1-50 comes in agglomerates that hide the His-tags from the primary antibodies used in Western blotting. This poly-His tag hiding may partly explain the low yield obtained in the purification step.

Like for Azu 1-50, Azu 1-77 was purified by the same method. First, for 1L of cell culture, Azu 1-77 superexpression sample and purified collected fractions from the Ni-NTA column were analysed by SDS PAGE and Western Blotting. The results are shown in Figure 22.

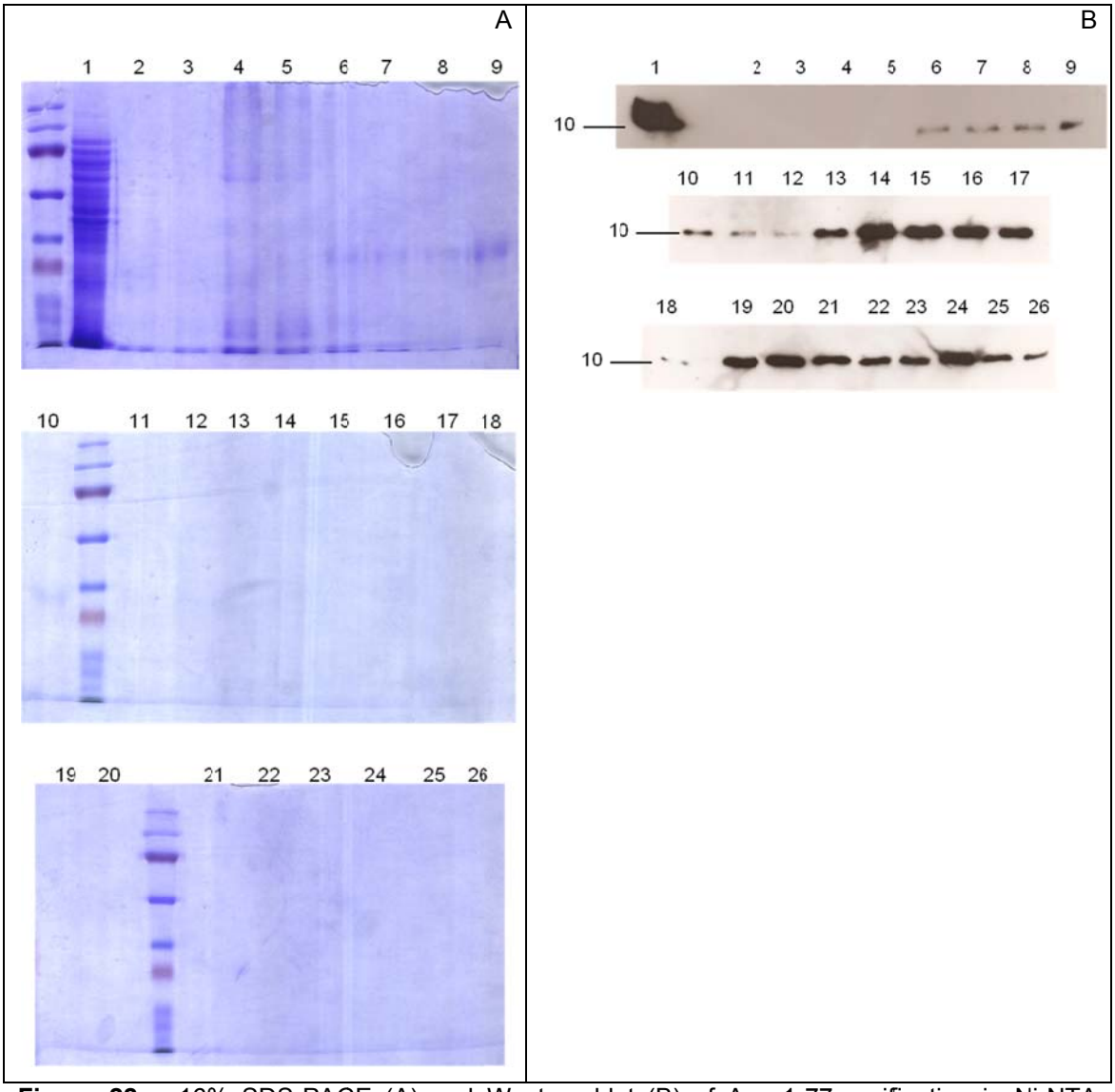

Figure 22 – 16% SDS-PAGE (A) and Western blot (B) of Azu 1-77 purification in Ni-NTA column. Lane description is the same for both A and B figures: 1 – superexpression bacterial culture at 4h (0.2 mM IPTG); lane 2 and 3 – fractions 2 and 3 of 20 mM imidazole elution; lane 4 and 5 – fractions 2 and 3 of 40 mM imidazole elution; lane 6 and 7 – fractions 2 and 3 of 60 mM imidazole elution; lanes 8 to 12 – fractions 1, 2, 3, 4 and 5 of 100 mM imidazole elution; lanes 13 to 17 – fractions 1, 2, 3, 4 and 5 of 200 mM elution; lanes 18 to 22 – fractions 1, 2, 3, 4 and 5 of 300 mM imidazole elution; lane 23 and 24 – fractions 2 and 3 of 400 mM imidazole elution; lane 25 and 26 – fractions 2 and 3 of 500 mM imidazole elution.

Figure 22 shows that Azu 1-77 is obtained pure at all fractions from 200 to 400 mM imidazole elution, as there is no visible bands in the Coomassie blue stained SDS-PAGE gels, but there are visible bands in the Western blot. Therefore these fractions were collected and loaded into the ÄKTA apparatus for peptide desalting. The obtained chromatograms obtained for 1L and 2L of starting cell culture volumes are present in Figure 24 and Figure 24, respectively.

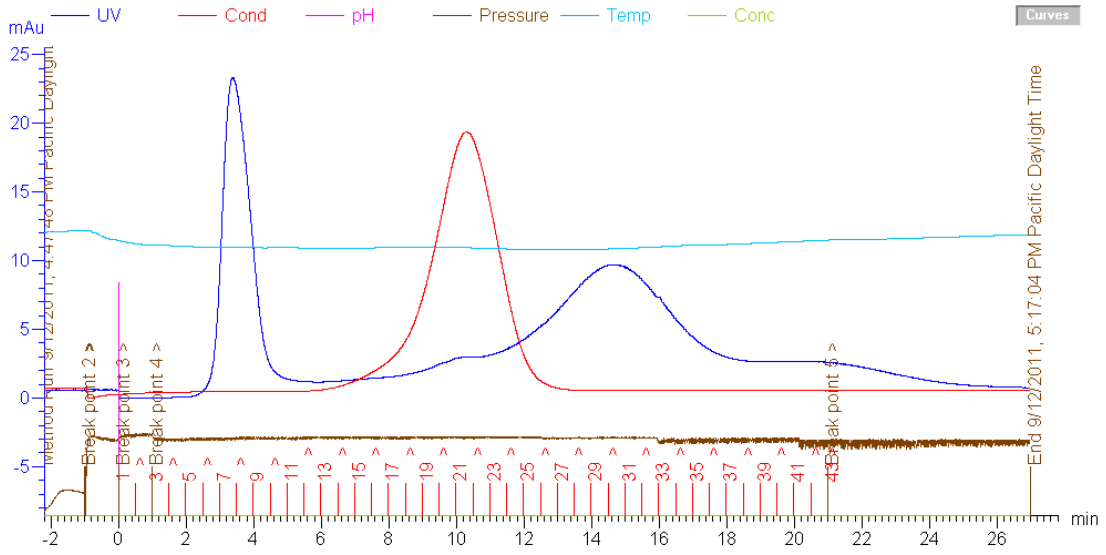

Figure 23 – ÄKTA chromatogram obtained from purification of 1L of cell culture expressing Azu 1-77. Absorbance curve at 280 nm (dark blue line) indicate that Azu 1-77 is present in fractions 6 to 10 having its higher concentration at fraction 8. The absorbance peak is approximately 24 mAu. The latter peak corresponds to the imidazole presence.

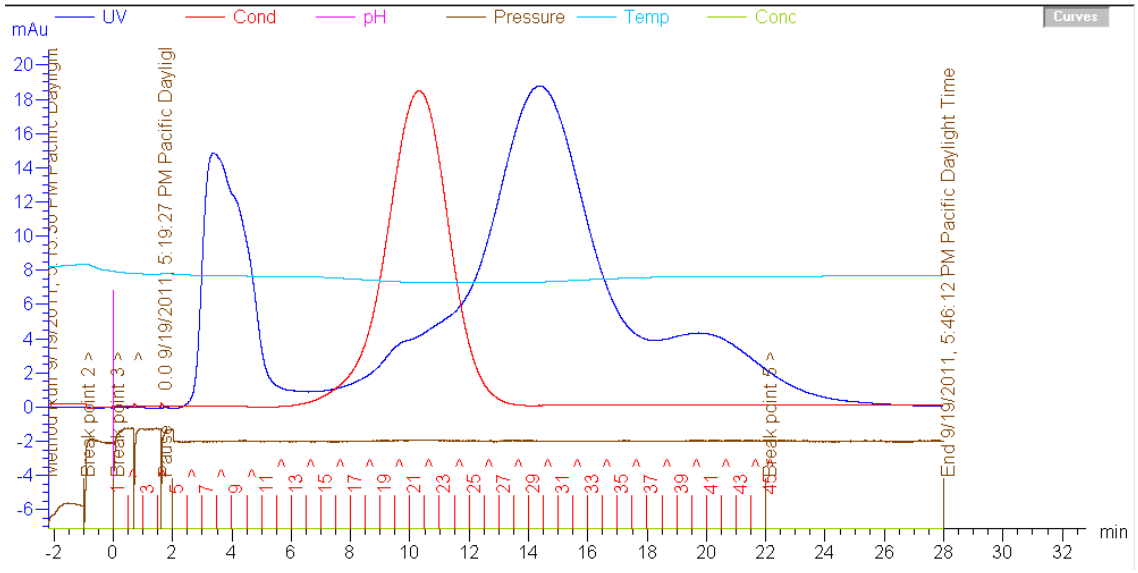

Figure 24 – ÄKTA chromatogram obtained from purification of 2L of cell culture expressing Azu 1-77. Absorbance curve at 280 nm (dark blue line) indicate that Azu 1-77 is present in fractions 6 to 12 having its higher concentration at fraction 7. The absorbance peak is approximately 15 mAu. The latter peak corresponds to the imidazole presence.

In Figure 23 and Figure 24 is observed a first 280 nm peak of 24 and 15 mAu, respectively, corresponding to Azu 1-77 presence in the eluted fractions. The peak obtained from 1L of initial cell culture is higher than the one obtained from 2L but is also thinner. Both peaks are well resolved. The latter peak corresponds to the imidazole elution. Comparing to Azu 1-50, 2L of cell culture expressing Azu 1-77 generates a higher peak with a similar thickness. This shows, as it was previously said, that Azu 1-77 has better levels of expression than Azu 1-50. Although, these values are still very far from those obtained for azurin, and the quantity of Azu 1-77 produced in each batch is still very small comparing to what should be desired.

Concerning the Azu 1-77 produced from 2L of cell culture, it is present in fractions 6 to 11 (fractions that correspond to the peptide peak in Figure 24). These fractions were collected and analyzed through Western blotting.

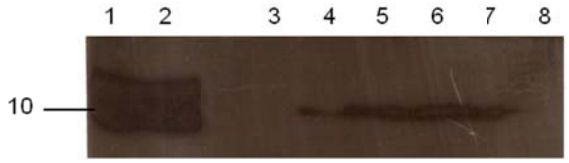

Figure 25 – Western blot analysis of ÄKTA fractions of purified 1-77 desalting. Lane description: lane 1 and 2 – distinct superexpression bacterial cultures at 4h (0.2 mM IPTG); lane 3 to 8 – fraction 6 to 11.

Figure 25 show that 1-77 is present in fractions 7 to 10 and that it has higher concentration in the latter two fractions. This Western blot confirms that Azu 1-77 is being successfully purified but at very low quantities.

The collected fractions were concentrated in a 3 kDa cut-off column until a final volume of approximately 1 mL. The final concentrations were about 5 µM for both Azu 1-50 and Azu 1-77.

4.2.3 Spectroscopic analysis

After obtaining both Azu 1-50 and Azu 1-77 purified, spectroscopic analysis were performed. In a first approach, the spectrum from 250 to 800 nm was traced to visualize the aminoacid characteristic peaks located in the near-UV region (from 350 to 250 nm). To better analyze the peptide secondary structure, far UV circular dichroism (CD) spectrum (from 190 to 250 nm) was obtained.

Azurin UV-Vis spectrum was traced to compare with azurin derived peptides spectra. The UV-Vis spectra are shown in Figure 26, Figure 27 and Figure 28.

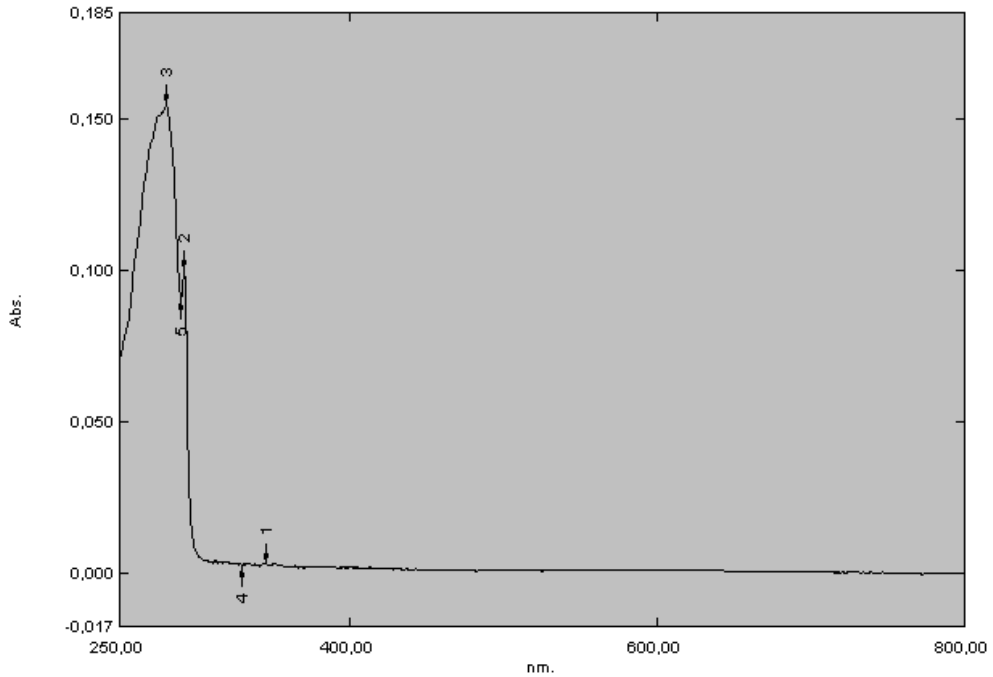

Figure 26 – Azurin 250 to 800 nm spectrum. The maximum absorption peak is located at 280 nm and a well resolved elbow at 292 nm.

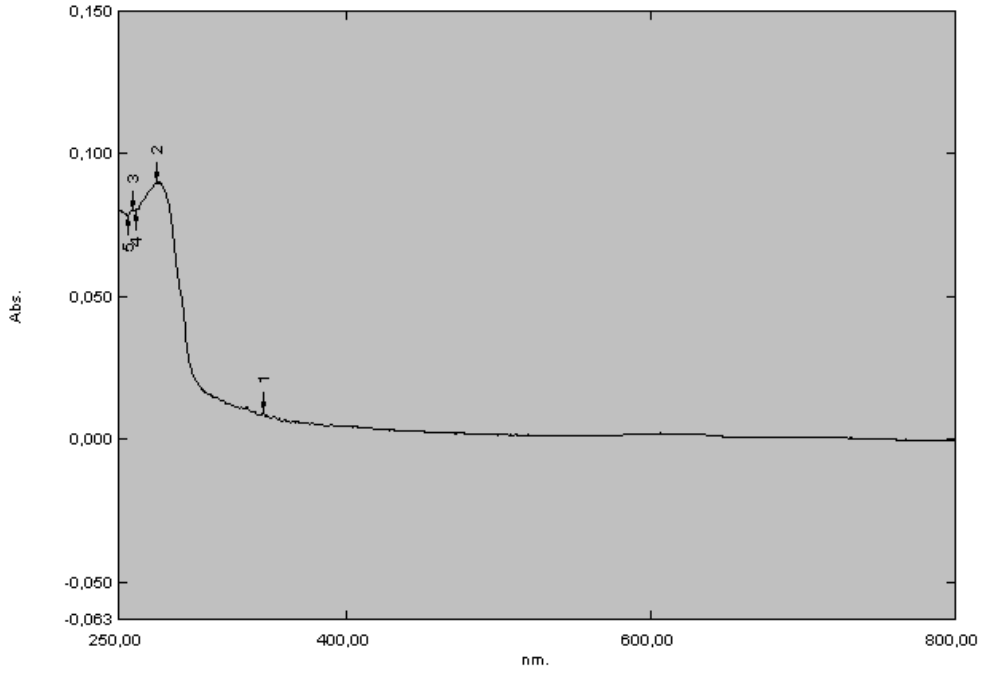

Figure 27 – Azu 1-50 250 to 800 nm spectrum. The maximum absorption peak is located at 280 nm and a smooth elbow at 292 nm.

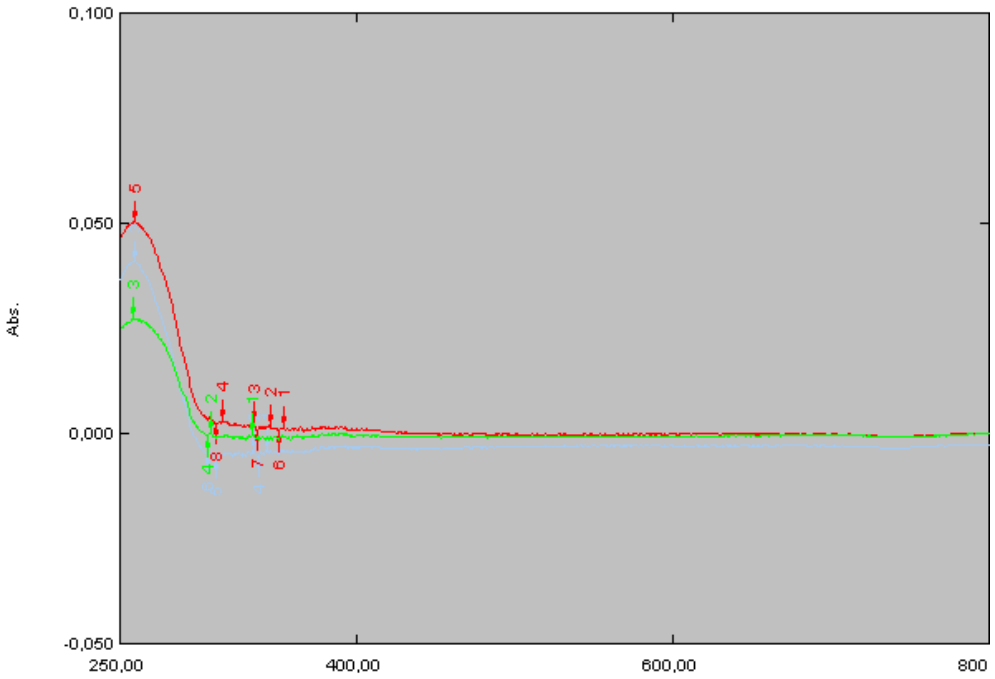

Figure 28 – Azu 1-77 250 to 800 nm spectrum. The maximum absorption peak is located at 258 nm.

Figure 26, Figure 27 and Figure 28, show that both azurin and Azu 1-50 have the maximum absorption peak at 280 nm and Azu 1-77 has it at 258 nm. While azurin has a resolved peak at 292 nm corresponding to the tryptophan (48) residue, Azu 1-50 has a smooth elbow in the same region. On the other hand, Azu 1-77 maximum absorption peak is located at 258 nm. This difference was not expected and explains the low signals detected during the ÄKTA desalting step.

In order to analyze the peptide secondary structure, far UV spectra were obtained for both peptides. The spectra are presented in Figure 29.

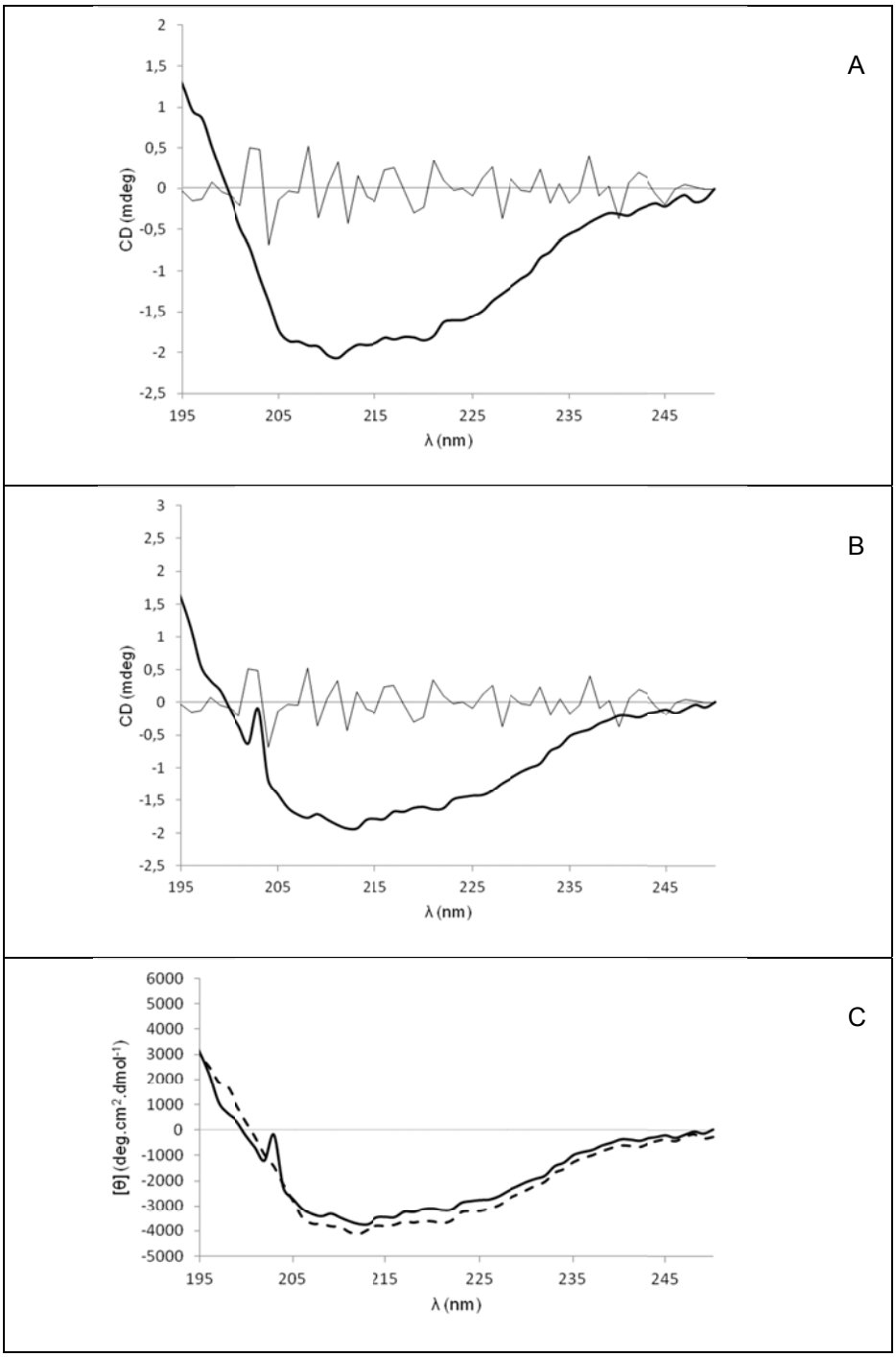

Figure 29 – Far UV CD spectra of Azu 1-50 (A) and Azu 1-77 (B) peptides. The thicker line corresponds to the CD signal. This spectra corresponds to the smoothed media of 10 different measurements with a window size of 5. The thinner line corresponds to the mean residues. The normalized data (C), in molar ellipticity ([θ]) show similar spectra for both peptides. Dashed line corresponds to the Azu 1-50 and solid line corresponds to Azu 1-77.

Figure 29 shows that both peptides have a defined secondary structure, since it is observed a negative depression between 200 and 240 nm, typical of both α-helix and β-sheet structure. Both spectra suggest that Azu 1-50 and Azu 1-77 have a similar structure. The main difference observed between both spectra is the existence of a local maximum at 202 nm in Azu 1-77 spectra. A pure β-sheet structure has a minimum at 215 nm while a pure α-helix has two local

minima at 210 and 220 nm. The resulting spectra do not clearly show the minima expected in a pure structure, which show that none of the peptides has a pure structure.

The DICHROWEB server was used to analyze the far-UV spectra. The DICHROWEB server provides access to different programs (CONTIN, SELCON3, VARSLC, CDSTTR and K2d) and databases (reference data base 1 to 7 and SPF175 (Lees *et al.*, 2006)) used for CD spectra analysis. Along with the predicted structure provided by the programs, DICHROWEB retrieves the NRMSD (normalized root mean square deviation) that is used as a goodness-of-fit parameter. NRMSD indicates how similar the calculated and experimental spectra are. Ideally, NRMSD values should be lower than 0.1. (Whitmore and Wallace, 2004, 2008)

The predicted fractions of secondary structures in both peptides using different programs and reference data sets are presented in Table 17.

Table 17 – Secondary structure prediction of Azu 1-50 and Azu 1-77 using different programs and reference data sets.

	Analysis program	**Reference data set**	**Predicted fractions of secondary structure**				**NRMSD**
			α-helix	**β-sheet**	**turns**	**unordered**	
Azu 1-50	SELCON3	4	0.200	0.271	0.224	0.305	0.260
		7	0.214	0.217	0.192	0.377	0.227
		SP175	0.234	0.246	0.133	0.389	0.397
	CONTIN	4	0.239	0.230	0.230	0.301	0.217
		7	0.230	0.190	0.187	0.393	0.217
		SP175	0.217	0.256	0.132	0.394	0.225
	K2d	–	0.30	0.13	(n.d.)	0.57	0.182
Azu 1-77	SELCON3	4	0.182	0.359	0.196	0.265	0.306
		7	0.123	0.303	0.164	0.365	0.289
		SP175	0.251	0.280	0.143	0.326	0.322
	CONTIN	4	0.170	0.293	0.241	0.295	0.256
		7	0.162	0.261	0.206	0.370	0.256
		SP175	0.122	0.345	0.132	0.401	0.203
	K2d	–	0.17	0.31	(n.d.)	0.52	0.538

(n.d.): not defined by the program

Having in consideration the NRMSD values, the more relevant predictions are given by the K2d program for Azu 1-50 and CONTIN (using the SP175 reference data set) for Azu 1-77. Although, in both cases the NRMSD values are higher than what is desired for accurate predictions (<0.1).

In both cases there is a predominance of unordered structure over α-helix and β-strands structure. Azu 1-50 has predominance of α-helix (0.30) over β-sheet (0.13), while for Azu 1-77 is observed the opposite, β-sheet (0.345) predominance over α-helix (0. 122).

5 General discussion

The main objective of this work was to develop a reproducible technique of production of azurin derived peptides Azu 1-50, Azu 1-77 and Azu 80-128. Recent work (Fialho, 2009) identified P-cadherin as a possible target for azurin. Because P-cadherin is associated with the aggressiveness of breast cancers (Ribeiro *et al.*, 2010), it is of great interest to study azurin interaction with this protein. Therefore, producing azurin derived peptides is of great interest so further studies can be done on this matter. The study of Azu 1-50, Azu 1-77 and Azu 80-128 may elucidate which region of azurin is involved in P-cadherin interaction, N- or C-terminal region. Between Azu 1-50 and Azu 1-77, the main difference is the presence and absence of the PTD domain, p28. Comparing the effect of both peptides on breast cancer cells may elucidate the importance of the penetration of azurin on its activity against P-cadherin, *i. e.*, it may elucidate whether azurin effect on P-cadherin is mainly done extracellularly or on the interior of breast cancer cells. The study of Azu 80-128 is important to clarify if the C-terminal has or not a role in the interaction.

Previous studies show that azurin and azurin derived peptides can be successfully produced in *E. coli* expression strains and purified using a GST based purification system (Chaudhari *et al.*, 2007; Yamada *et al.*, 2002b). The adopted strategy of this work uses a poly-His tag based purification method. Azurin has been successfully produced in our lab using this method (Bernardes *et al.*, unpublished), therefore the same method was implemented for production of the azurin derived peptides. The described methods of producing azurin derived peptides by other authors (Yamada *et al.*, 2002b) use *E. coli* strains as expression systems. *E. coli* was chosen as expression host for azurin derived peptides production in this work. The evaluated strains were *E. coli* SURE and *E. coli* BL21(DE3). For purification, poly-His tag strategy was chosen because, besides it is successfully used for azurin production, GST has some disadvantages that may pose some difficulties in the production process. GST tag (218 aminoacids, 26 kDa) is a metabolic burden for the host cell and has a great tendency to form dimers and insoluble aggregates whereas poly-His tag (6 aminoacids, 1 kDa) poses a low metabolic burden for the host and can be used in denaturant conditions (Waugh, 2005). On the other hand, because poly-His tag is small, it has little effect on peptide and protein conformation and usually does not need to be cleaved from the target protein, whereas GST tag poses greater problems on this matter (Waugh, 2005).

After successfully cloning the three azurin derived peptides in the pWH844 plasmid, which encodes a poly-His tag in the N-terminal of the target protein, *E. coli* expression strains SURE and BL21(DE3) were transformed. Because *E. coli* SURE did not demonstrate to have substantially better expression levels and because *E. coli* BL21(DE3) already showed to be a good expression strain for azurin derived peptides (Yamada *et al.*, 2005), *E. coli* BL21(DE3) was chosen as expression strain. For Azu 80-128, no expression could be achieved neither using *E. coli* BL21(DE3) neither using *E. coli* SURE (data not shown) possibly because it may

be structurally unstable. In a first approach, in order to obtain Azu 80-128 expression, more expressing assays should be performed to evaluate other expression strains and other growth conditions (temperature of induction and medium culture) (James R, 2001; Peti and Page, 2007; Stevens, 2000). In the case of Azu 1-77, expression is observed in Coomassie blue stained SDS-PAGE gels, albeit the expression bands are not as strong as it is expected in superexpression assays. For Azu 1-50, Western blotting analysis were necessary to conclude that there is expression of the peptide. At this point, it became obvious that Azu 1-50 was being expressed in much lower levels than Azu 1-77, albeit this last peptide is also being expressed at low levels. Like it was said for Azu 80-128, more expression assays should be performed for both these peptides in order to achieve higher levels of expression. The lower expression of Azu 1-50 comparing to Azu 1-77 may have to do with the increment of the p28 extended hydrophobic α-helix that may stabilize Azu 1-77.

Having in consideration the observed aberrant mobility of the Azu 1-50 in 16% SDS-PAGE gels, it was noted that other authors previously had the same observation when producing GST-Azu 36-50 (Yamada *et al.*, 2005). In fact, GST-Azu 36-50 (27.6 kDa) band was slightly higher than GST-Azu 50-77 (28.9 kDa) band. From all seven azurin derived peptides produced in this study, GST-Azu 36-50 was the only peptide that showed aberrant mobility and was also the only peptide that did not include the PTD or part of the PTD (Yamada *et al.*, 2005). These observations made us induce that the absence of the α-helix may influence its migration through the gel matrix.

In both Azu 1-50 and Azu 1-77 dimers formation was detected. The formation of these dimers may be a consequence of β-sheets interaction between near peptides. In fact it has been published that some proteins form dimers and agglomerates due to β-sheets interactions that form stable complexes evolving two or more proteins (Maitra and Nowik, 2000). As azurin forms a tight β-barrel with its eight β-strands, the β-strands present in Azu 1-50 and Azu 1-77 (A, B1, B2, C and D) may interact with their homologous on the near peptides to form a stable structure. The formation of these dimers may be related with the low quantity of final purified peptides obtained, as they can reduce peptides solubility and therefore reduce the peptide recovery and purification yield. In order to increase the purification yield, denaturant conditions can be used as the poly-His tags can be purified under such conditions (Waugh, 2005). However, a refolding step would be needed.

Analysis of the purified fractions of Azu 1-50 obtained from the Ni-NTA column show that at relatively high concentrations of imidazole, there are still present some contaminants. This may be due to the main disadvantage of using poly-His tag purification methods. Poly-His tag purification systems use IMAC (immobilized metal affinity chromatography) columns. Because many proteins can have some affinity to metals, these IMAC techniques are not so specific as other techniques that use specific ligands (Waugh, 2005). Curiously, Western blotting analysis of Ni-NTA collected fractions narrowly detects few Azu 1-50. The presence of Azu 1-50 in the collected fractions is confirmed through the Western blotting of the desalted fractions obtained

from the ÄKTA apparatus, *i. e.*, Azu 1-50 should be detected in the Ni-NTA fractions and yet, only traces are detected. The explanation to this observation is that, prior or during the column purification, Azu 1-50 forms dimers or agglomerates that shield the poly-His tags from the involving environment, therefore poly-His tags become almost inaccessible to the primary antibody used in immunoblotting. This hypothesis also explains the great loss of peptide during the purification step. If poly-His tags are shielded with peptide structures, it has no access to the column affinity resin and therefore cannot bind to it and is washed out of the column in the first washing steps. Because there is only a small amount of poly-His tags competing to bind to the immobilized metals, other proteins with some affinity to the resin remain in the column and contaminate the collected peptide.

After purifying the peptides, both were desalted through size exclusion chromatography using the ÄKTA apparatus. Peptides were successfully desalted and collected in phosphate buffer. This desalting step is important for imidazole removal. Because the production of these peptides intends to obtain them in a suitable medium for cell testing, imidazole needed to be removed, so that it does not influence the obtained results. On the other hand, because spectrophotometric measures were made, buffer exchange from PBS to phosphate buffer was needed to remove the NaCl salt that affects the CD results. Imidazole removal is also very important in CD technique because it also affects the results (Kelly *et al.*, 2005).

In the first attempt of Azu 1-50 purification, Azu 1-50 extracted from 1L of cell culture was purified and desalted. In the desalting step low concentration of protein was detected (15 mAu peak) and the obtained peak was well resolved and thin. In azurin purification, usually 1L of cell culture purification results in a 200-250 mAu peak that forms a platform at this absorbance value. Because size exclusion columns intended for desalting only separate macromolecules (proteins) from small salts, having a thick and high peak indicates a good quantity of protein. For a starting 1L of culture cell, the obtained peak was low and thin, therefore the initial volume of cell culture was increased in order to obtain a higher and thicker peak. To the maximum volume of initial cell culture tested, 2L, the resulting peak was not much higher nor much broader than the one obtained for 1L. Having in consideration the previous observations, Azu 1-50 dimerization hiding the poly-His tails may really pose a great obstacle in its recovery and difficult the process scale up.

For Azu 1-77 desalting and buffer exchange step, 1L of initial cell culture resulted in a thin peak with higher absorbance (20 mAu) of the obtained for Azu 1-50. The higher absorbance of the peak shows that Azu 1-77 is produced in higher quantity than Azu 1-50. The upscale of cell culture volume to 2L resulted in a thicker but lower absorbance peak (similar to Azu 1-50 peak). Because no great differences were observed between the results using 1L or 2L of initial cell culture, like for Azu 1-50, this process seems difficult to upscale.

Immunobloting by itself show that very low quantities of produced peptides are achieved. In order to obtain enough quantity of peptides for cell assays, more than one batch of production

will be needed, which is more cost and time consuming than what should be desired. Therefore it is important to test different expression and purification strategies. For instance, the use of a fusion tag that enhances peptide solubility is a good hypothesis. The SUMO (small ubiquitin modifying protein) tag has been demonstrated to enhance protein expression and correct folding of proteins therefore increasing their solubility. This tag can be fused together with a poly-His tag for purification purposes. Besides, SUMO can be easily cut from the target protein using very efficient SUMO proteases that do not leave extraneous residues in the target protein (Butt *et al.*, 2005).

At the end of the purification spectrophotometric analysis of the obtained peptides were made.

UV-Vis spectrum of Azu 1-50 is very similar to the azurin spectrum except that Trp-48 characteristic well resolved peak at 292 nm (Mei *et al.*, 1996) is reduced to a slight elbow in Azu 1-50 spectrum. The lowering of the Trp-48 peak is certainly due to changes in the involving environment of the aminoacid. While in azurin the Trp is strategically located in the hydrophobic core, in Azu 1-50 Trp-48 is more exposed to the environment. On the other hand, Azu 1-77 UV-Vis spectrum is very different from the azurin and Azu 1-50 spectra. Its maximum absorption is located at 258 nm, which is far distant from the maximum absorption obtained for azurin and Azu 1-50 (280 nm). Typically, protein have their maximum absorption peak located near 280 nm, therefore typical detection systems like the one used by the ÄKTA apparatus measure the samples absorption near this wavelength. However, there are documented cases of protein which have absorption maximum distant from the typical 280 nm. One example is the superoxide dismutase (SOD) that, like Azu 1-77, has its maximum absorption peak at 258 nm (Chen *et al.*, 2009). Chen *et al.* attribute this uncommon absorption peak and absence of 280 nm peak to the low content of tryptophan and presence of only one tyrosine residue (Chen *et al.*, 2009). Other protein with this characteristic absorbance is the poly(ADP-ribose) (PAR) (Shah *et al.*, 1995). Like in the case of SOD, the unusual absorbance maximum of Azu 1-77 may have to do with a possible interaction between Trp-48 and Tyr-71, since the Trp-48 is more exposed.

For far-UV CD spectrum analysis different predictions were performed using different programs and different reference databases. For Azu 1-50 the best fitting prediction was obtained using the K2d algorithm. This prediction gave 30% of α-helix, 13% of β-sheet and 57% of unordered structure. However, the NRMSD value is too high to consider this an accurate prediction (0.182). This high value is possibly due to the low concentration of the peptide.

For Azu 1-77 CONTIN algorithm using the SP175 reference data gave the best fitting prediction with an NRMSD of 0.203. Again, this value is too high to consider this prediction accurate, but it indicates the predominance of the secondary structures. The estimated percentage of α-helix is 0.122, β-sheet 0.345, turns 0.132 and unordered structure 0.401. In this case there is a predominance of β-sheet over α-helix structure and this predominance is present in all

prediction results (Table 17). Like for Azu 1-50, there is a predominance of unordered structure over the other defined structures.

Comparing to azurin structure, it would be expected that for both peptides, most of their structure would be in β-sheet conformation and that Azu 1-77 would have a higher percentage of α-helix structure over Azu 1-50 due to the presence of the p28. Although, the best fitting predictions were obtained by different programs for each peptide, a brief comparison show that our predictions, higher percentage of α-helix in Azu 1-77 than in Azu 1-50, were not observed in CD analysis and peptides are predicted to be mainly in unordered conformation. Because it lacks the rest of azurin protein and the common sequence of the peptides (1-50) is part of a rigid β-core, the absence of surrounding stabilizing β-sheets may dictate the structural inability to form the predicted β-sheets. On the other hand, the p28 in Azu 1-77 is located in a peripheral region of the peptide. Because of this position, the predicted α-helix may not be able to stabilize and does not form a defined structure. In sum, for both peptides, the majority of the peptide structure is unordered. However, this may not dictate the absence of activity because there are known antimicrobial peptides like melittin that in solution are unstructured but in contact with membranes are able to acquire an active conformation (Mihajlovic and Lazaridis, 2010). This can be the case of these peptides, although other studies have to be performed in order to access this hypothesis.

6 Conclusions and future perspectives

Heterologous protein and peptide expression in *E. coli* is well documented and is often effective for producing sufficient quantities of protein for further works. Many factors can be optimized, like the host strain, the chosen vector, growth medium and temperature (Stevens, 2000). In this work three azurin derived peptides, Azu 1-50, Azu 1-77 and Azu 80-128 were cloned in a pWH844 vector that codes a histidine tag in the N-terminal of the target protein and is regulated by a T5 promoter. Expression was performed in *E. coli* BL21(DE3) cultured at 37°C in a rich culture medium designed for protein superexpression (SB medium). At the end of the work, no expression was achieved for Azu 80-128 and low levels of expression of Azu 1-50 and Azu 1-77 were achieved. Poly-His tag purification had a low yield and the final quantity of peptides obtained was far beyond what is needed for further work. Therefore, the strategy for peptide production adopted in this work demonstrated to inadequate for the work objectives. Studies concerning optimization of previously referred factors must be done in order to produce adequate quantities of peptide per batch. A great factor that must be tested is the change and conjugation of affinity tags. Poly-His tag did not show to be effective and, because Yamada *et al.* (Yamada *et al.*, 2005) have already successfully produce other azurin derived peptides using a GST-tag, a GST-tag is an alternative worth to test. Another worth trying technique would be the use of a SUMO protein fused with poly-His tag. This construct have already shown to be effective in expression and solubility enhancement of proteins and peptides (Butt *et al.*, 2005).

7 References

Apiyo, D., and Wittung-Stafshede, P. (2005). Unique complex between bacterial azurin and tumor-suppressor protein p53. Biochemical and Biophysical Research Communications *332*, 965-968.

Bernardes, N. (unpublished). Bacterial protein azurin as a new candidate drug to treat P-cadherin overexpressing breast cancer.

Bernardes, N., Seruca, R., Chakrabarty, A.M., and Fialho, A.M. (2010). Microbial-based therapy of cancer: Current progress and future prospects. Bioengineered bugs *1*, 178-190.

Bostrom, P., Soderstrom, M., Vahlberg, T., Soderstrom, K.-O., Roberts, P., Carpen, O., and Hirsimaki, P. (2011). MMP-1 expression has an independent prognostic value in breast cancer. BMC Cancer *11*, 348.

Butt, T.R., Edavettal, S.C., Hall, J.P., and Mattern, M.R. (2005). SUMO fusion technology for difficult-to-express proteins. Protein Expression and Purification *43*, 1-9.

CDG (2011). CDGTI Webpage.

Chaudhari, A., Fialho, A.M., Ratner, D., Gupta, P., Hong, C.S., Kahali, S., Yamada, T., Haldar, K., Murphy, S., Cho, W.*, et al.* (2006). Azurin, Plasmodium falciparum malaria and HIV/AIDS: inhibition of parasitic and viral growth by Azurin. Cell Cycle *5*, 1642-1648.

Chaudhari, A., Mahfouz, M., Fialho, A.M., Yamada, T., Granja, A.T., Zhu, Y., Hashimoto, W., Schlarb-Ridley, B., Cho, W., Das Gupta, T.K.*, et al.* (2007). Cupredoxin-cancer interrelationship: azurin binding with EphB2, interference in EphB2 tyrosine phosphorylation, and inhibition of cancer growth. Biochemistry *46*, 1799-1810.

Chen, S., Li, X., and Ma, H. (2009). New Approach for Local Structure Analysis of the Tyrosine Domain in Proteins by Using a Site-Specific and Polarity-Sensitive Fluorescent Probe. ChemBioChem *10*, 1200-1207.

De Grandis, V., Bizzarri, A.R., and Cannistraro, S. (2007). Docking study and free energy simulation of the complex between p53 DNA-binding domain and azurin. Journal of Molecular Recognition *20*, 215-226.

Duncan Patrick, M. (2008). Discovering and improving novel peptide therapeutics. Current Opinion in Pharmacology *8*, 616-619.

FDA (2011). FDA approved drug products.

Fialho, A.M. (2009). Innovative biotechnological approaches for cancer therapy: bacterial protein azurin as a new anti-cancer drug candidate. In Biological Sciences Research Group anual report 2009, pp. 24-26.

Fialho, A.M., and Chakrabarty, A.M. (2010). Promiscuous Anticancer Drugs from Pathogenic Bacteria: Rational Versus Intelligent Drug Design. In Emerging Cancer Therapy (John Wiley & Sons, Inc.), pp. 179-198.

Fialho, A.M., Das Gupta, T.K., and Chakrabarty, A.M. (2007a). Designing Promiscuous Drugs? Look at What Nature Made. Letters in Drug Design & Discovery *4*, 40-43.

Fialho, A.M., Das Gupta, T.K., and Chakrabarty, A.M. (2008). Promiscuous drugs from pathogenic bacteria in the post-antibiotics era. In Patho-Biotechnology, R. Sleator, and C. Hill, eds. (Landes Bioscience), pp. 145-162.

Fialho, A.M., Stevens, F.J., Das Gupta, T.K., and Chakrabarty, A.M. (2007b). Beyond host-pathogen interactions: microbial defense strategy in the host environment. Current Opinion in Biotechnology *18*, 279-286.

Gauci, C., Jenkins, D., and Lightowlers, M. (2011). Strategies for Optimal Expression of Vaccine Antigens from Taeniid Cestode Parasites in <i>Escherichia coli</i&gt. Molecular Biotechnology *48*, 277-289.

Graumann, K., and Premstaller, A. (2006). Manufacturing of recombinant therapeutic proteins in microbial systems. Biotechnology Journal *1*, 164-186.

informa-healthcare (2007). Scrip: Biosimilars, biogenerics and follow-on biologics

James R, S. (2001). Advances in Escherichia coli production of therapeutic proteins. Current Opinion in Biotechnology *12*, 195-201.

Jia, L., Gorman, G.S., Coward, L.U., Noker, P.E., McCormick, D., Horn, T.L., Harder, J.B., Muzzio, M., Prabhakar, B., Ganesh, B.*, et al.* (2010). Preclinical pharmacokinetics, metabolism, and toxicity of azurin-p28 (NSC745104) a peptide inhibitor of p53 ubiquitination. Cancer chemotherapy and pharmacology.

Kelly, S.M., Jess, T.J., and Price, N.C. (2005). How to study proteins by circular dichroism. Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics *1751*, 119-139.

Khandare, J.J., and Minko, T. (2006). Antibodies and peptides in cancer therapy. Critical Reviews in Therapeutic Drug Carrier Systems *23*, 401-435.

Lees, J.G., Miles, A.J., Wien, F., and Wallace, B.A. (2006). A reference database for circular dichroism spectroscopy covering fold and secondary structure space. Bioinformatics *22*, 1955-1962.

Lien, S., and Lowman, H.B. (2003). Therapeutic peptides. Trends in Biotechnology *21*, 556-562.

Lu, Y., Yang, J., and Sega, E. (2006). Issues related to targeted delivery of proteins and peptides. The AAPS Journal *8*, 466-478.

Maitra, S., and Nowik, J.S. (2000). Beta sheet interactions between proteins. In The amide linkage: structural significance in chemistry, biochemistry, and materials science, A. Greenberg, C.M. Breneman, and J.F. Liebman, eds. (John Wiley & Sons, Inc), pp. 495-518.

Mehta, R., Yamada, T., Taylor, B., Christov, K., King, M., Majumdar, D., Lekmine, F., Tiruppathi, C., Shilkaitis, A., Bratescu, L.*, et al.* (2011). A cell penetrating peptide derived from azurin inhibits angiogenesis and tumor growth by inhibiting phosphorylation of VEGFR-2, FAK and Akt. Angiogenesis *14*, 355-369.

Mei, G., Agró, A.F., Rosato, N., Gilardi, G., Venanzi, M., and Canters, G.W. (1996). Probing the structure and mobility of Pseudomonas aeruginosa azurin by circular dichroism and dynamic fluorescence anisotropy. Protein Science *5*, 2248-2254.

Micewicz, E., Jung, C.-L., Schaue, D., Luong, H., McBride, W., and Ruchala, P. (2011). Small Azurin Derived Peptide Targets Ephrin Receptors for Radiotherapy. International Journal of Peptide Research and Therapeutics *17*, 247-257.

Mihajlovic, M., and Lazaridis, T. (2010). Antimicrobial peptides in toroidal and cylindrical pores. Biochimica et Biophysica Acta (BBA) - Biomembranes *1798*, 1485-1493.

Paredes, J., Albergaria, A., Oliveira, J.T., Jerónimo, C., Milanezi, F., and Schmitt, F.C. (2005). P-Cadherin Overexpression Is an Indicator of Clinical Outcome in Invasive Breast Carcinomas and Is Associated with CDH3 Promoter Hypomethylation. Clinical Cancer Research *11*, 5869-5877.

Paredes, J., Correia, A., Ribeiro, A., Albergaria, A., Milanezi, F., and Schmitt, F. (2007). P-cadherin expression in breast cancer: a review. Breast Cancer Research *9*, 214.

Pasquale, E.B. (2010). Eph receptors and ephrins in cancer: bidirectional signalling and beyond. Nature reviews Cancer *10*, 165-180.

Peti, W., and Page, R. (2007). Strategies to maximize heterologous protein expression in Escherichia coli with minimal cost. Protein Expression and Purification *51*, 1-10.

Pichereau, C., and Allary, C. (2005). Therapeutic peptides under the spotlight. Eur Biopharm Rev, 88-91.

Punj, V., Bhattacharyya, S., Saint-Dic, D., Vasu, C., Graves, J., Yamada, T., Christov, K., White, B., Li, G., Majumdar, D.*, et al.* (2004). Bacterial cupredoxin azurin as an inducer of apoptosis and regression in human breast cancer. Oncogene *23*, 2367-2378.

Ribeiro, A.S., Albergaria, A., Sousa, B., Correia, A.L., Bracke, M., Seruca, R., Schmitt, F.C., and Paredes, J. (2010). Extracellular cleavage and shedding of P-cadherin: a mechanism underlying the invasive behaviour of breast cancer cells. Oncogene *29*, 392-402.

Sambrook, J., and Russel, D.W. (2001). Molecular cloning: A laboratory manual, Vol 3, Third edn (Cold Spring Harbor Laboratory Press).

Sato, A.K., Viswanathan, M., Kent, R.B., and Wood, C.R. (2006). Therapeutic peptides: technological advances driving peptides into development. Current Opinion in Biotechnology *17*, 638-642.

Schagger, H. (2006). Tricine-SDS-PAGE. Nat Protocols *1*, 16-22.

Shah, G.M., Poirier, D., Duchaine, C., Brochu, G., Desnoyers, S., Lagueux, J., Verreault, A., Hoflack, J.C., Kirkland, J.B., and Poirier, G.G. (1995). Methods for Biochemical Study of Poly(ADP-Ribose) Metabolism in Vitro and in Vivo. Analytical Biochemistry *227*, 1-13.

Sharpe, S., Yau, W.-M., and Tycko, R. (2005). Expression and purification of a recombinant peptide from the Alzheimer's β-amyloid protein for solid-state NMR. Protein Expression and Purification *42*, 200-210.

Soler, A.P., Knudsen, K.A., Salazar, H., Han, A.C., and Keshgegian, A.A. (1999). P-cadherin expression in breast carcinoma indicates poor survival. Cancer *86*, 1263-1272.

Sorensen, H., and Mortensen, K. (2005). Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. Microbial Cell Factories *4*, 1.

Stevens, R.C. (2000). Design of high-throughput methods of protein production for structural biology. Structure *8*, R177-R185.

Taylor, B.N., Mehta, R.R., Yamada, T., Lekmine, F., Christov, K., Chakrabarty, A.M., Green, A., Bratescu, L., Shilkaitis, A., Beattie, C.W.*, et al.* (2009). Noncationic Peptides Obtained From Azurin Preferentially Enter Cancer Cells. Cancer Research *69*, 537-546.

Vlieghe, P., Lisowski, V., Martinez, J., and Khrestchatisky, M. (2010). Synthetic therapeutic peptides: science and market. Drug Discovery Today *15*, 40-56.

von Strandmann, E.P., Hansen, H.P., Reiners, K.S., Schnell, R., Borchmann, P., Merkert, S., Simhadri, V.R., Draube, A., Reiser, M., Purr, I.*, et al.* (2006). A novel bispecific protein (ULBP2-BB4) targeting the NKG2D receptor on natural killer (NK) cells and CD138 activates NK cells and has potent antitumor activity against human multiple myeloma in vitro and in vivo. Blood *107*, 1955-1962.

Vu, T.H., and Werb, Z. (2000). Matrix metalloproteinases: effectors of development and normal physiology. Genes & Development *14*, 2123-2133.

Waugh, D.S. (2005). Making the most of affinity tags. Trends in Biotechnology *23*, 316-320.

Whitmore, L., and Wallace, B.A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Research *32*, W668-W673.

Whitmore, L., and Wallace, B.A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. Biopolymers *89*, 392-400.

Yamada, T., Fialho, A.M., Punj, V., Bratescu, L., Das Gupta, T.K., and Chakrabarty, A.M. (2005). Internalization of bacterial redox protein azurin in mammalian cells: entry domain and specificity. Cellular Microbiology *7*, 1418-1431.

Yamada, T., Goto, M., Punj, V., Zaborina, O., Chen, M.L., Kimbara, K., Majumdar, D., Cunningham, E., Das Gupta, T.K., and Chakrabarty, A.M. (2002a). Bacterial redox protein azurin, tumor suppressor protein p53, and regression of cancer. Proceedings of the National Academy of Sciences of the United States of America *99*, 14098-14103.

Yamada, T., Goto, M., Punj, V., Zaborina, O., Kimbara, K., Das Gupta, T.K., and Chakrabarty, A.M. (2002b). The Bacterial Redox Protein Azurin Induces Apoptosis in J774 Macrophages through Complex Formation and Stabilization of the Tumor Suppressor Protein p53. Infection and Immunity *70*, 7054-7062.

Yamada, T., Mehta, R.R., Lekmine, F., Christov, K., King, M.L., Majumdar, D., Shilkaitis, A., Green, A., Bratescu, L., Beattie, C.W.*, et al.* (2009). A peptide fragment of azurin induces a p53-mediated cell cycle arrest in human breast cancer cells. Molecular cancer therapeutics *8*, 2947-2958.