Hinc patriam sustinet

Instituto Superior de Agronomia
Universidade de Lisboa

RELATÓRIO INTEGRADOR DA ACTIVIDADE PROFISSIONAL

João Gomes de Matos Duarte

Relatório para obtenção do Grau de Mestre em

Engenharia Agronómica: Proteção de Plantas

Orientador: Doutora Sara Alexandra Valadas da Silva Monteiro

Co-orientador: Doutor Ricardo Manuel de Seixas Boavida Ferreira

Júri:

Presidente: Doutora Cristina Maria Moniz Simões de Oliveira, Professora Associada com Agregação do Instituto Superior de Agronomia da Universidade de Lisboa.

Vogais: Doutora Maria da Glória Calado Inglês Esquível, Professora Auxiliar do Instituto Superior de Agronomia da Universidade de Lisboa.

Doutora Sara Alexandra Valadas da Silva Monteiro.

Lisboa, 2013

U LISBOA | UNIVERSIDADE DE LISBOA

Acknowledgments

To Doctor Sara Monteiro for her enthusiasm, guidance, support and overall friendship. For engaging me in new ideas and demanding a high quality of work in all my endeavors. For all the encouragement, comments, sound advice, and good ideas for this thesis.

To Professor Ricardo Boavida Ferreira, whose expertise, understanding, patience and scientific rigor has been vital in my work.

To MsC Herma Koehorst-van Putten and Doctor José Vidal whose teachings, advices and insights were invaluable to me.

To my colleagues for sharing their enthusiasm and for comments on my work, Alexandra Carreira, Iliana Pereira, Ricardo Chagas, Ana Lima and Alexandre Borges.

To Regina, Ana Cristina, Luis, Damiano, Carlos, Sílvia, Isabel, Jorge, Sérgio, Ana, Mário, André, and all those I forgot to mention, for all the help, sympathy and good moments in the laboratory.

To all my friends throughout these years that I did not mentioned but always accompanied me during this adventure.

To my parents and brother for their love, support and understanding, during the long years of my education.

And finally, to Catarina, my dearest friend, cheerleader, drill sergeant, voice of reason, counselor and life raft. Thank you for everything.

Abstract

This thesis constitutes a summary of my work for the past eight years. Given the different subjects that I have been involved, I selected three major lines of work to expose in this thesis, which are subdivided in four Chapters: Chapter I: "Descriptive *Curriculum vitae*" containing a brief presentation of the work developed, constituting a more descriptive *cv,* with a clear description of the projects I have been involved, techniques that I have become acquainted and a brief explanation of my responsibilities during this period. In Chapter II, entitled "Preliminary analysis of the antimicrobial protein content of the Spurge flax plant", the antimicrobial content from the aqueous extract of the Spurge flax plant proved to be proteinaceous in nature, peptidic in size (<3 kDa) and resistant to denaturant treatments. In Chapter III, "*Agrobacterium* sp. mediated transformation of four *Rosa hybrida* cultivars with antifungal polypeptide BLAD", the attempt towards the genetic modification of the rose plant with a BLAD potent antifungal polypeptide , is described. The final chapter, Chapter IV, "Optimization of the heterologous expression in *E.coli* of the antifungal polypeptide BLAD," describes the different approaches applied towards the soluble over-expression in *E.coli* of the antifungal polypeptide BLAD.

Keywords: Plant peptides, Rose transformation, Heterologous expression.

Resumo

Esta tese constitui um resumo do meu trabalho nestes últimos oito anos. Dado que me debrucei sobre diferentes áreas, selecionei três linhas de trabalho principais para expor nesta tese, tendo estas sido subdivididas em quatro capítulos: Capítulo I: "*Curriculum vitae* descritivo", contém um breve resumo do meu trabalho, podendo ser encarado como um *curriculum vitae* detalhado, onde apresento os projetos em que estive envolvido, as técnicas que aprendi e/ou implementei no laboratório e as responsabilidades que me foram atribuídas. Capítulo II: "Estudo preliminar da fração proteica com atividade antimicrobiana das folhas de Trovisco". Neste capítulo a elevada atividade antimicrobiana presente no extrato aquoso das folhas de trovisco foi caracterizada, demonstrando-se pertencer à fração proteica de massa molecular reduzida (< 3 kDa) e com elevada resistência a diferentes tratamentos desnaturantes. Capítulo III: "Transformação genética por via da *Agrobaterium* sp. de quatro cultivares de *Rosa Hybrida* com a proteína antifúngica BLAD." Este capítulo descreve toda a metodologia seguida para a transformação genética de quatro cultivares de *Rosa hybrida*. Capítulo IV: " Expressão heteróloga em *E.coli* do polipéptido BLAD", este capítulo debruça-se sobre os diferentes métodos usados na tentativa de expressar em *E.coli* de forma solúvel o polipéptido BLAD.

Palavras-chave: Péptidos de plantas, Transformação rosa, Expressão heteróloga.

Resumo alargado

Esta tese teve como objetivo compilar o trabalho mais relevante que efetuei nos últimos oito anos e, assim, obter uma equivalência ao mestrado em Proteção de Plantas. Dada a diversidade de áreas em que estive envolvido optei por descrever mais exaustivamente apenas três trabalhos desenvolvidos. Sendo assim, esta tese está subdividida em quatro grandes capítulos:

Capítulo I: "*Curriculum vitae* descritivo", neste capítulo apresento uma descrição mais alargada do meu *curriculum vitae*, contendo os projetos em que estive envolvido, as técnicas que utilizei e que fui desenvolvendo ao longo do tempo e mais relevante as minhas responsabilidades em cada trabalho;

Capítulo II: "Estudo preliminar da fração proteica com atividade antimicrobiana das folhas de Trovisco (*Daphne gnidium*)". Este capítulo descreve a forte atividade antimicrobiana presente no extrato polipeptídico da folha da planta do trovisco. Os compostos responsáveis por esta atividade foram parcialmente caracterizados, revelando pertencerem à fração com 3 kDa, constituída por péptidos resistentes à fervura, acetona e à ação da tripsina;

Capítulo III: "Transformação genética por via da *Agrobaterium* sp. de quatro cultivares de *Rosa Hybrida* com o gene do polipéptido antifúngica BLAD". Este capítulo debruça-se sobre a tentativa de transformação de três cultivares de rosa com a fração do gene que codifica o polipéptido BLAD, com forte atividade antifúngica. Neste trabalho apesar de se ter conseguido obter culturas, raízes, callus e callus pré-embriogénicos para todas a cultivares, nenhuma das linhas de roseira regeneradas se encontrava transformada, provando que é essencial a total eliminação da *Agrobacterium* sp. para um eficiente protocolo de transformação;

Capítulo IV: "Otimização da expressão heteróloga em *E.coli* da proteína BLAD", este capítulo descreve as diferentes metodologias e tentativas utilizadas na tentativa de expressar em *E.coli*, de forma solúvel, o polipéptido BLAD. Comprovou-se que o polipéptido BLAD tem uma elevada propensão para uma expressão insolúvel, na forma de corpos de inclusão. Esta expressão insolúvel é independente da temperatura de crescimento da *E.coli* e do plasmídeo usado. Demonstrou-se, que, para ocorrer um certo grau de "refold" do polipéptido BLAD após expressão insolúvel, é necessário utilizar versões mais curtas do polipéptido e diluir a sua concentração, mantendo uma concentração elevada de ureia, e só depois fazer uma diálise "Step-Wise", para retirar gradualmente a ureia.

Index

List of Tables

List of Figures

List of Abbreviations

2,4-D- 2,4-Dichlorophenoxyacetic acid
2x35SCaMV- cauliflower mosaic virus double 35S promoter
als- acetolactate synthase
AMP- antimicrobial plant peptides
BAP- N-Benzyl-9-(2-tetrahydropyranyl) adenine
bar- phosphinothricin-N- acetyltransferase
BCIP- 5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt
BLAD- Banda lupinus albus doce
BSA- bovine serum albumin
CBB R-250- Coomasie Brilliant Blue R 250
CI- callus induction solid media
CLSI- Clinical and Laboratory Standards Institute
CM-cellulose or CM column- Carboxymethyl-cellulose
DEAE-cellulose column- Diethylaminoethyl-cellulose
EDTA- Ethylene diamine tetracetic Acid
FPLC- Fast Protein Liquid Chromatography
GM- germination medium
gor- glutathione reductase
GST- glutathione-s-transferase
GUS- β – glucuronidase
GYP- Glucose, Yeast, Peptone
Hpt- hygromycin phosphotransferase
IBA- Indole-3-butyric acid
IPTG- isopropyl-β-D-1- thiolgalactopyranoside
ISA- Instituto Superior de Agronomia
LB- Luria-Bertoni medium
LBA- Luria-Bertoni solid medium
LMW- Low molecular Weight Marker
LUC- luciferase
MBP- maltose binding protein
ME- 2- β-mercaptoethanol
MGM- maturation/germination medium
MH- Mueller Hinton
MIC- minimal inhibitory concentrations
MS- Murashige and Skoog medium
MWCO- Molecular weight Cut-Off
NBT- nitro-blue tetrazolium chloride
NCBI- National Centre for Biotechnology Information
nptII- neomycin phosphotransferase II

OD- optical density
PAGE- Poliacrylamide gel electrophoresis
PCA- Plate Count Agar
PCR- Polymerase Chain Reaction
PDA- Potato Dextrose Agar
PDB- Potato Dextrose Broth
PEMs- pre-embryogenic masses
Pfx- proofreading
PR proteins- Pathogenesis-Related proteins
PSA- Ammonium persulfate
PVPP- Polyvinylpyrrolidone
RACE- Rapid Amplification of CDNA Ends
SBN- native sample buffer
SDS- Sodium dodecyl sulfate
SUMO- ubiquitin-like modifier
Taq- Thermus aquaticus
TCA- Trichloroacetic acid
T-DNA- Transfer DNA
TEMED- Tetramethylethylenediamine
trxB- thioredoxin reductase
Ub- ubiquitin
X-Gluc- 5-bromo-4-chloro-3-indolyl glucuronide

Chapter I

Descriptive Curriculum vitae

I obtained my bachelor degree in Agronomical engineering, specialized in "Melhoramento de Plantas" now called Plant Biotechnology and Breeding, from Instituto Superior de Agronomia (ISA) in early 2005. My final dissertation title was "Purificação de proteínas extraídas de *Vitis vinifera* sp. com actividade antifúngica sobre fungos e leveduras patogénicas para o Homem". The experimental work of this dissertation took place in the "Laboratório de Fisiologia Vegetal" under the guidance of Professor Ricardo Boavida Ferreira and Drª Sara Monteiro. As the title entails, this thesis focused on the purification of antifungal proteins and other compounds present in *Vitis vinifera* grapes. To accomplish the proposed objectives, I began by learning rudimentary protein handling and purification techniques, specifically how to operate an FPLC (Fast Protein Liquid Chromatography) device, in order to purify sufficient amounts of grape antifungal proteins, and the SDS-PAGE technique to visualize protein separation. In order to test the purified grape proteins I had to develop suitable methods to correctly determine the antifungal activity of the different samples. These new methods were based on the guidelines published by the CLSI organization, specifically the M-27-A2 guideline for yeast susceptibility tests and the M-38-P guideline for fungal susceptibility testing. At the same time I had to learn proper techniques and procedures to safely work with human pathogenic yeasts and fungus, such as the highly infectious *Aspergillus fumigatus.* In the end I was able to determine MIC (minimal inhibitory concentrations) values for the purified grape proteins and for other grape compounds such as isolated grape condensed tannins and tannic acid.

After the end of my dissertation I received a proposal to continue to work in the same research group, under the same guidance, but in a different subject. The "ROSAOÍDIO" Project, supported by "Agência de Inovação" (AdI), had at that time been approved, with the objective of genetically transform *Rosa Hybrida* plants with three antifungal proteins: BLAD (Banda *L.albus* Doce) polypeptide, a Thaumatin like protein from *Vitis vinifera* grapes and Stylbene sinthase also from *Vitis vinifera* grapes. I received a grant under this project and I was responsible for the actual rose plant genetic transformation. Considering our lack of knowledge on plant transformation, I was sent to work and learn with MsC Herma Koehorst-van Putten in the Laboratory of Plant Breeding, Plant Research International Wageningen University, Holland. I went to Holland, in alternating months, for a period of two years, there I learned all the tissue cultures steps towards genetically transform rose plants. Specifically, tissue culture techniques, procedures on how to handle micropropagated plant material, how to correctly identify the different stages of embryogenic development, such as pre-embryogenic calli, embryogenic calli and plantlet development and finally how to transform and regenerate genetically transformed rose plants. During this two year period, in the months that I was back in ISA, I had to construct plant expression plasmids to be used for transformation, containing our different target protein genes. To construct the expression plasmids I had to learn all the molecular biology methods, essential for this objective. I became familiar with primer design, PCR amplification, "Realtime" PCR, DNA sequence analysis, plant RNA and DNA extraction, RACE (Rapid Amplification of CDNA Ends)

technique, different DNA cloning methodologies (e.g. Gateway technologie), *E.coli* transformation, *A. tumefaciens* transformation, DNA ligation, DNA restriction and Southern Blot hybridization (DIG labeling).

At the same time that the "ROSAOIDIO" project took place, in the periods I was back to ISA, and after I finished the construction of the expression plasmids, I was responsible for determining the antimicrobial potential of the proteinaceous content of the flax-leafed (or trovisco) plant (*Daphne gnidium*). The Flax leafed plant is endemic to Southern Europe, widely used as antifungal treatment in traditional medicine around those regions. Following my previous work on the grape antifungal protein content, I began by analyzing the antimicrobial activity of the total protein content of the leaf and bark of this plant. I applied the antifungal susceptibility testing methodologies previously developed for my undergraduate thesis, and developed a new method, also based on the CLSI guidelines, adapted to antibacterial susceptibility testing. Using an optimized extraction procedure, I was able to demonstrate that the total protein content of this plant possessed a wide spectrum of antimicrobial activity, and most likely, the biomolecule responsible was a cationic plant peptide. To achieve this end result, I had to deeply improve my knowledge on protein theory, learn new techniques on protein handling and protein chromatography, specifically size exclusion chromatography and ultrafiltration separation, and adapt to try and purify a peptide.

After 2008, I began working in the recently formed CEV ("Consumo em Verde") company. A start-up company founded by Professor Ricardo Boavida Ferreira, Drª Sara Monteiro and Professor Virgílio Loureiro, for commercializing the extremely potent antifungal polypeptide BLAD, from *Lupinus albus* seedlings. Although the main focus was on BLAD production, through the isolation of this protein from *Lupinus albus* seedlings as a side work the company was interested in obtaining genetically modified plants expressing the BLAD protein. Therefore, I entered the company and the collaboration with Wageningen University was maintained. I was at that point responsible for the bench work in Holland and to report back to CEV on how the project was progressing.

During the periods that I was back in Lisbon and ISA, I was put under the guidance of Drª Alexandra Carreira in the Quality Department. At the time this department was involved in developing new techniques, to be used in the future, to accurately determine the quality of a given batch of PROBLAD fungicide. Therefore new methods had to be developed to determine this new fungicide *in vitro* antifungal activity and its shelf life. During this work I learned how to conduct industry grade antifungal susceptibility testing, how to perform solid media antifungal testing (antibiograms) and how to determine a specific product level of contamination, influential in its shelf life, by quantifying the degree of contamination, either aerobic/anaerobic bacterial or fungal.

In 2010, due to the lack of results in rose transformation, CEV focused on developing genetically modified grapevine plants constitutively expressing the BLAD polypeptide. A collaboration was set between CEV and Dr. José Ramon Vidal from the Department of Plant Physiology, University of Santiago de Compostela, Spain. I was responsible for working directly under José Vidal, learning this new transformation technique, constructing the expression plasmids and report back to CEV on the

progress of the project. I stayed in Spain first for 2 months for *calli* induction and for another period of 6 months for actual Biolistic plant transformation and then for alternating months, to follow transformed plant development. During my time in Lisbon I was responsible for expression plasmid construction. This work initiated with induction of *calli* development for 6 *V. vinifera* cultivars at different developmental stages and from four type of explants (leafs petioles, ovaries and anthers) while testing different induction media. During this stage of the work I deeply improved my tissue cultures techniques, and learned how to identify properly formed *V. vinifera* calli. During my time back in Lisbon I was responsible for developing small plant expression plasmids containing the BLAD gene and GUS gene, for Biolistic transformation. For plasmid construction I chose to construct new sets of plasmids, to reduced overall size, therefore I used different parts of two distinct plasmids. To succeed in this work I had to improve my molecular biology techniques and knowledge, specifically on primer design, sequence analysis and DNA restriction and ligation. From this work I also became proficient in the plasmid visualization vector NtI software (Invitrogen). Upon plasmid construction I went back to Spain to begin Biolistic plant transformation and regeneration. At this stage I learned how to correctly operate the Biolistic device PDS-1000/He, to handle the *V. vinifera* cells to be in optimal conditions for Biolistic transformation and how to select and regenerate properly transformed plan cells.

In 2012 due to problems in the *V.vinifera* transformation project, I was put in charge of the heterologous expression of the BLAD polypeptide in *E.coli*. For CEV the expression of pure BLAD is of enormous importance since it will allow the use of Bioreactors, and an entry into the market of human applications. But although apparently simple, BLAD *E.coli* expression proved to be a highly technical project. BLAD upon over-expression in *E.coli* is a deeply insoluble polypeptide. To improve the soluble expression of BLAD in *E.coli* I had to apply different approaches. I tested different refolding procedures, and diverse techniques for maximizing soluble expression. In the end I focused on the influence of BLAD´s primary and tertiary structure and possible structure modifications to improve solubility. Although *E.coli* expression was a completely new technique for me, the analysis of BLAD's tertiary structure was a completely different world. For structure analysis I had to learn different software analysis to predict BLAD solubility (hydrophobicity plots, REFOLD, aggregation prediction, PASTA) and to generate and compare tertiary structures (RasMol, Chimera, RASH). This work allowed to learn new techniques in *E.coli* protein expression, and more important a glimpse of proteins structure analysis and rational design of proteins to improve solubility.

Chapter II

Preliminary analysis of the antimicrobial protein content of the Spurge flax plant

The potential discovery of a new antimicrobial peptide

Introduction

Through ought time an invisible chemical warfare has been taking place between plants and their various enemies. Although in this ongoing war plants seem an unmovable and defenseless target, in reality due to millions of years of Darwinian adaptation, plants have developed a set of very effective physical and chemical tools for defense, such as: the ability to regrow damaged or eaten parts, mechanical protection, the presence of hydrophobic cuticular layers, the production of latex or resins which deter chewing insects, a cell wall which presents a physical barrier to inhibit microbial infections (De Lucca *et al.*, 2005) and more appealing for researchers is the production of a vast array of antimicrobial secondary plant metabolites (Cutler and Cutler, 1999)**.** These antimicrobial phytochemical compounds are present in all plants and plant tissues in a myriad of forms constituting a multitude of families of compounds, for instance: phenols, quinones, flavonoids, tannins, coumarins, terpenoids, alkaloids, and relevant to this work, plant proteins and peptides (Broekaert *et al*., 1997; García-Olmedo *et al.*, 1998; Cowan, 1999; Selitrennikoff, 2001).

On the chase for new antimicrobial compounds, research, led by the agrochemical and pharmaceutical companies is focusing on substances with new modes of action, high specificity, low toxicity to humans and wildlife, benign environmental profile, high stability and if possible low production costs. Plant antimicrobial proteins and peptides possess several of the above characteristics and are considered an obvious group of compounds to find the next big molecule (Cutler and Cutler, 1999). Taking into consideration that in the world today floristic analyses indicate that approximately 250 000 – 500 000 plant species exist, and that only about 6% have been screened for biologic activity, and a reported 15% have been evaluated phytochemically (Verpoorte, 2000). It is safe to say that tens of thousands of active antimicrobial proteins and peptides are there to be discovered (Tivy, 1993). Obviously we do not possess the ability to analyze the entire known flora for active biological molecules therefore screening techniques must be applied. One good starting point and pre-selection technique is through the use of ethnobotanical and ethnopharmacological works to make educated guesses on witch plants to start analyzing for novel antimicrobial biomolecules.

In this work we chose an endemic Portuguese plant, Spurge flax (*Daphne gnidium*), identified in ethnopharmacology works has possessing strong antibacterial and antifungal properties, to determine the antimicrobial activity of the proteinaceous content of the leafs. Afterwards having determined the presence of biological activity, this work switched its focus and attempted to characterize, fractionate and isolate the proteinaceous molecule responsible for the observed antimicrobial activity.

1. Ethnobotany, Ethnopharmacology and the focus on *Daphne gnidium*

Ethnobotany as a discipline was first introduced in 1895 by American taxonomic botanist John W. Harshberger and defined as the study of the utilitarian relationship between human beings and vegetation in their environment, including medicinal uses (Harshberger, 1896; Cotton, 1996). Ethnopharmacology, one the other hand, focuses specifically on drug discovery by the observation, description, and experimental investigation of indigenous drugs and their biologic activities. It is based on botany, chemistry, biochemistry, pharmacology, and many other disciplines (anthropology, archaeology, history, and linguistics) and contributes directly to the discovery of natural products with biologic activity (Rivier and Bruhn, 1979). Ethnopharmacological information can be acquired from various sources such as books on medical botany, review articles (usually involving surveys of medicinal plants by geographic region or ethnic culture), field work and computer databases, like NAPRALERT database (Fabricant and Farnsworth, 2001) or the USDA–Duke database (USDA, NRCS, 2013). Considering the number of higher plant species present in nature, these works constitute a good starting point to any work focusing on the search for active biomolecules in plants (Fabricant and Farnsworth, 2001).

The plant *Daphne gnidium (Thymelaeaceae),* commonly named Spurge Flax, Flax-leaved daphne, (Marques *et al*., 2009) or in Portuguese Trovisco, is an evergreen shrub widely diffused in the Mediterranean area, with slender, straight, brown branches, with subcoriaceous, linear to ovate-oblong acute leaves and creamy-white flowers in small, terminal panicles (Roccotiello *et al.*, 2009). Species of the genus Daphne are used in natural medicine as a diuretic, laxative, antifungal, antimicrobial, and anticoagulant, in the treatment against skin diseases, toothache, and malaria. Studies have shown that plants from the genus Daphne contain a large number of classes of secondary metabolites, predominantly coumarins, flavonoids, lignans, diterpenes and steroids (Manojlovi *et al*., 2012). The *Daphne gnidium* plant, specifically, is portrayed in several ethnopharmacological works has being regularly used to treat bacteria, fungal and/or yeast infections. (Cottiglia *et al*., 2002; Ferreira, 2007; Pereira, 2008). Works focused on this plants methanolic extract found that it also possesses a strong insecticidal (Maistrello *et al*., 2005) and antioxidant activity (Deiana *et al*., 2003). A more in depth analyzes of the Spurge flax plant antimicrobial activity has been carried out, focused on its methanolic extract content (Iauk *et al.,* 1996; Iauk *et al.,* 1997; Cottiglia *et al*., 2001). The crude methanolic extract of leafs and bark of this plant has been investigated and found to possess some degree of antibacterial activity (Cabrera and García-Granados, 1981) namely against *Bacillus lentus* and *Escherichia coli* and at higher concentrations on *Staphilococcus aureus*, and *Pseudomonas aeruginosa*. But no activity was observed against the yeasts *Candida albicans* and *Candida tropicalis* (Cottiglia *et al*., 2001), incongruent with its traditional use as a treatment for fungal infections. The methanolic extracts were isolated and found to be constituted by two main groups of compounds, flavonoids and coumarins, with differences in-between plant tissue. Specifically, the bark contains mainly: daphnetin, daphnin, acetylumbelliferon, daphnoretin, apigenin, luteolin, quercetin, orientin, isoorientin, apigenin 7-O-glucoside, genkwanin, and 5-O-β-D-primeverosyl genkwanine. Whereas, the leaves contain, daphnetin, daphnin, acetylumbelliferon, daphnoretin, luteolin,

genkwanine, luteolin 7-O-glucoside, and α-tocopherol (Deiana *et al*., 2003). Of these different compounds the ones found to possess some degree of antimicrobial activity were the compounds daphnetin and genkwanin, present in both the plant tissues (Cottiglia *et al*., 2001). Regarding the proteinaceous content of the Spurge flax plant and its biological activity, to this date no work has been published on this subject. A cursory look at the NCBI (National Centre for Biotechnology Information) database (Geer *et al*., 2010), only presents us with one protein entry from the Spurge flax plant, consisting of a NADH dehydrogenase protein subunit belonging to the chloroplasts (Gene Bank: ACU12875.2). Therefore considering its widespread use in traditional medicine, and the lack of published works on its proteinaceous content, the Spurge flax plant is a valid target for the search of new antimicrobial proteins or peptides.

2. Plant antimicrobial proteins

Plant antimicrobial proteins can be found constitutively expressed in the plants or induced as a result of pathogen infection. Plant antimicrobial proteins constitute a broad group of proteins with different and sometimes unique modes of activity that have been proved active against a wide variety of plant and human pathogens, with low propensity for resistance development, and in some cases possess synergistic action with classical antibiotics (Marr *et al*., 2006; Costa *et al.*, 2011). The induced fraction of plant antimicrobial proteins constitute a diverse group of proteins named Pathogenesis-Related proteins (PR proteins), that generally possesses the highest biological activity. A unifying nomenclature was introduced and in the present time fourteen sub-families have been proposed for this group, encompassing proteins with various modes of action for instance: chitinases (PR- 3), beta 1-3 glucanase (PR-2) peroxidase (PR-9) and thaumatin (PR-5) and also some plant peptides, such as families PR-12, 13 and 14 (Van Loon and Van Strien, 1999). The non-induced fraction of plant antimicrobial proteins is not has diverse and can be subdivided in proteins with lectin activity, Trypsin inhibitor, 2S albumins, Ribosome-inactivating proteins, several antimicrobial plant peptides (AMP) and some ungrouped proteins.. However plant proteins possess some shortcomings that have been delaying their wider use in the field and in the hospital. Plant proteins in general are highly instable substances easily degraded (proteases, pH shifts, and even U.V light), with some toxicity towards other organisms (e.g. mammals) and possible allergenic effects. Finally antimicrobial proteins tend to appear in low concentrations on the plant making their production and isolation tedious and costly (Cottiglia *et al.,* 2005). Because of these characteristics very few natural antimicrobial proteins have hit the market, one rare example is the highly active antifungal BLAD (Banda *Lupinus albus* Doce) protein.

2.1. Plant antimicrobial peptides (AMP)

Plants as a form of imunity response to biotic and abiotic stresses produce an enormous variety of chemical compounds, some show strong antimicrobial activity and relevant to this thesis a few are of are of small size (<10 kDa) and proteinaceous in nature and are commonly referred as antimicrobial peptides or AMP. Peptides are generally defined as proteins of no more than 100 amino acid residues in length whose molecular weight is not greater than 10 kDa. In contrast, protein is defined as proteins

greater than 100 amino acids residues with molecular weights greater than 10 kDa (De Lucca *et al.*, 2005). AMPs have been found in all plant organs, their expression depending on the plant development phase and the different stresses (some AMP are included in PR-proteins classes) but as a rule of thumb the seed is normally a rich source of AMPs (Rogozhin *et al.*, 2012). Peptides with inhibitory activity towards microorganisms were first reported in 1942 by Balls and his collaborators, which isolated and identified a peptide named purothionin from wheat seeds, that showed antimicrobial activity towards bacteria and yeasts (Balls *et al*., 1942; Hernandez-Lucas *et al*., 1978)**.** Almost 70 years after hundreds of plant antimicrobial peptides have been identified as having broad spectrum of activities against viruses, bacteria, fungi, and parasites (Montesinos and Bardaj, 2008). Natural AMPs show an extraordinary molecular diversity, and have been naturally design to target molecules present in one of the most ancient and well established cell systems: the membrane (Montesinos and Bardaj, 2008). Despite the multitude of different AMP they all show some similar structural characteristics such as: a comparative small size (molecular weight less than 10 kDa); presence of more cationic (arginine and lysine) and hydrophobic (tryptophan, phenylalanine, leucine, and isoleucine) aminoacids; amphipathicity conferred by their aminoacid content and arrangement; a total positive charge at physiological pH values and the existence of disulfide bonds in their structure (Marcos *et al*., 2008). Although they share the above mentioned characteristics they have a remarkable diversity of structures and conformations, including α-helices, β-sheets, non- conventional structures, or even extended conformations. Being the first two the more abundant in nature (Powers and Hancock, 2003). To date there is not absolute consensus on the classification the plant AMP and so far depending on the literature up to 10 families have been described in plants these include thionins lipid, defensins, transfer peptides (LTPs) hevein and knottin like proteins, as well as antimicrobial proteins isolated from *Macadamia integrifolia* (MBP-1) and *Impatiens balsamina* (Ib-AMP) (Pelegrini, 2011). Of all these families the more diverse and better analyzed peptide families are the thionins, defensins, Hevein-like, Knottines and cyclotides the other are smaller families sometimes with only one representative peptide (Loeza-Ángeles *et al.*, 2011).

2.1.1. Plant Thionins

The plant thionins were the first family of AMP to be isolated and therefore one of the better characterized. Around 100 individual thionin sequences have been identified in more than 15 different plant species (Stec, 2006). Thionins occur in seeds, stems, roots, and leaves of several plant species and all share common structural features with a high degree of sequence homology (Chandrashekhara *et al*., 2010). Thionins type peptides show a low molecular weight (approximately 5 kDa), monomeric structure, rich in arginine, lysine and cysteine residues and include three or four conserved disulfide linkages and due to the high content in basic aminoacid residues thinonis are normally positively charged at neutral pH (García-Olmedo *et al.*, 1998). The presence of disulphide linkages imposes a amphipathic rigid structure and a higher stability of the molecule, depending on the number of disulphide linkages thionins can be subdivided on three groups. In spite of their extreme divergence in disulphide bonds and also in aminoacid content all thionins have essentially the same three-dimensional shape representing a unique, three-dimensional structure: a vertical stem

consisings of a pair of an-tiparallel alpha-helices, and a horizontal arm consisting of a short antiparallel Beta-sheet resembling the Greek capital letter gamma (Thevissen *et al.*, 1996). The hypothesis that thionins might play a role in the protection of plants against pathogens was first proposed by the research group of *Fernandez de Caleya et al.* who investigated the susceptibility to wheat endosperm thionins of phytopathogenic bacteria in the genera *Pseudomonas*, *Xanthomonas*, *Agrobacterium*, *Erwinia*, and *Corynebacterium.* Later on several peptides from the thionin family proved to have biological activity (Kragh *et al*., 1995; Thevissen *et al.*, 1996; Hammerschlag, 2004; Loeza-Ángeles *et al.*, 2008). Due to their antimicrobial activity and their pathogen induced behavior thionins are included in the pathogenesis related proteins group PR-13 (Van Loon and Van Strien, 1999).

2.1.2. Plant Defensins

The defensin family is one of the bigger and better characterized families of plant AMP (Sagaram *et al.*, 2011). Peptides belonging to this family have been found in a wide variety of plants monocot and dicot such as the edible root vegetable radish (Rs-AFP2), the flower dahlia (DMAPI) or the horse-chestnut tree (AhAMP1) (De Lucca *et al.*, 2005). Defensins have been mainly found in the seeds and leaves of plants and due to their inducible behavior when the plant is subject to stresses they are considered a PR protein belonging to the PR 12 group (Van Loon and Van Strien, 1999). To date, sequences of more than 80 different plant defensins genes from different plant species are available (Thomma *et al.*, 2002). Members of this family are quite diverse regarding amino acid composition, but normally constitute a basic cysteine-rich peptide with 45 to 54 amino acids in length with a similar scaffold, possessing one α-helix and two or three β-sheets stabilizing the structure are eight disulfide-linked cysteines, forming a typical cysteine stabilized αβ structural motif (CSαβ) (Cândido *et al.*, 2011) with an amphipathic structure characteristic of AMP (Epand and Vogel, 1999). The eight structure-stabilizing cysteines appear to be conserved among all plant defensins (Portieles *et al.*, 2006). In terms of activity although there are plant defensins that are active against bacteria, the majority of them are active against fungi (Cândido *et al.*, 2011). A good example is the plant defensins Hs-AFP1 and Rs-AFP2 that are lethal to the germinating conidia *Aspergillus flavus* and *Fusarium moniliforme* (De Lucca *et al.*, 2005). MsDef1, a 45-amino acid protein from the seed of *Medicago sativa*, inhibits the growth of a filamentous fungus, *Fusarium graminearum*, at micromolar concentrations (Sagaram *et al.*, 2011).

2.1.3. Plant Hevein-like peptides

Hevein is a small chitin-binding peptide consisting of 43 amino acid residues, which has been isolated originally from rubber latex and contains eight disulfide bridge-linked cysteine residues homologous to the chitin-binding domain of different types of multidomain proteins from plants (Kanokwiroon *et al.*, 2008). Hevein as the hability to bind to chitin inhibiting hyphal growth of fungi, especially yeast, not possessing any bactericide activity (Parijs *et al.*, 1991). This peptide served as the model for a new family of AMP named hevein-like peptides (Kanokwiroon *et al.*, 2008). Hevein-type peptides are cystein rich highly basic with pI values higher than 10, consisting of three folded Beta-strands, the second and third strands being linked by an alpha-helix. Hevein-like peptides are found in a wide

variety of plants and in different tissues such as: seeds of *Amaranthus caudatus* leaves of *Beta vulgaris*, and fruits of *Sambucus nigra*. While hevein is a rather weak antifungal, the hevein-like peptides have been shown to inhibit the growth of yeasts and fungi such as *Candida albicans, Candida tropicalis* (Koo *et al.*, 2002), *Alternaria brassicicola*, *Ascochyta pisi*, and *Fusarium culmorum* (De Lucca *et al.*, 2005). One good example of this family of peptides are the Ac-AMP1 and Ac-AMP2, with 29 and 30 amino acid residues respectively (Broekaert *et al.*, 1992)**,** they share sequence homology to the cysteine and (or) glycine-rich domain of chitin-binding proteins, and they both reversibly bind to chitin. Ac-AMP1 and Ac-AMP2 are strongly antagonized by cations, and are considered the first hevein-like peptides to show fungicidal effects similar to those of the thionins (De Lucca *et al.*, 2005). IWF4 is another hevein-like peptide present in sugar beet leafs and it has shown a high inhibition of the growth of *C. beticola*. Two hevein-like peptides, SmAMP1.1a and SmAMP2.2a, were previously isolated and proven to have inhibitory action against the fungi *Bipolaris sorokiniana* and *Thielaviopsis basicola* (Shukurov *et al.,* 2012).

2.1.4. Plant Knottins and cyclotides peptides

Knottins are a group of structurally related small molecules with around 30 aminoacids and have a common stable tertiary fold that is formed and stabilized by a characteristic pattern of disulfide bonds. Knottins fold into a triple-stranded Beta-sheet and form a knot-like structure (Kolmar, 2008). A special form of knottin peptides, constitute a family of peptides their one, are the cyclotides, the largest family of circular proteins currently known. This type of peptides, exclusively found in plants, have the unusual feature of a head-to-tail cyclized peptide, obviously maintaining the knottin-like structure. This peptide class is only found in the botanical families *Rubiaceae, Violaceae*, *Apocynaceae*, *Curcubitaceae*, and *Poaceae* (Cândido *et al.*, 2011). Both open chain and circular miniproteins containing the cystine-knot motif are remarkably stable towards extreme pH, chemical and thermal denaturation, and proteolytic attack. This is probably a result of the enormous conformational rigidity that is introduced by covalent disulfide linkage of the knotted core making them ideal targets for the pharmaceutical and agrochemical industries (Colgrave and Craik, 2004). Few non cyclic knottin-like peptides have been identified so far, those that have, are highly homologous with the knottin antimicrobial peptides Mj-AMP-1 and Mj-AMP-2 (36 and 37 residues, respectively) from *Mirabilis jalapa* L. seeds and PAFP-s (38 residues) from the seeds of *Phytolacca americana* (pokeweed). These open chained plant knottins display a limited activity against bacteria but a large antifungal activity against several fungi including *Fusarium oxysporum*, *Trichoderma viridae*, *Fusarium oxysporum*, *Fusariumgraminearum*, *Alternaria tenuis* (Shao *et al.*, 1999; Gao *et al.*, 2001).The cyclotide branch of this family has been more thoroughly searched and tested and has been found to possess mainly a bactericidal activity against Gram – and Gram + and only moderate antifungicide activity (Tam *et al.*, 1999).

Material and Methods

1. Plant material

Leaves of *Daphne gnidium* L. (termed Spurge flax hereafter) were collected from plants growing in the neighborhood of Loures (Covas de Ferro) during the spring of 2008, and stored separately at -80 ºC.

2. Protein extraction

Due to the presence on this plant of antimicrobial compounds soluble in organic solvents, we devised a specific two step protein extraction protocol to remove these plant's bioactive contaminants from our protein extract.

2.1. Organic solvent extraction (1 st step)

The organic solvent extraction comprised on adding cold pure acetone (Merck) to the leafs in a reason of 0.5 g of leafs to 10 mL of acetone, and grinding this mixture into fine particles with a mortar and pestle. After this step the emulsion was allowed to settle, and when there was a clear separation between phases, the emulsion was decanted and the liquid phase discarded. The mixture of plant powder and acetone was put on ice and left to rest for 2 hours in a fume hood, in order to remove the remaining acetone through evaporation. Upon evaporation the remaining plant powder was weighed and carefully labeled and stored a -80 ºC, ready for a second extraction with an aqueous buffer.

2.2. Aqueous buffer extraction (2 nd step)

Following the acetone extraction the plant powder obtained underwent a further extraction protocol using an aqueous buffer for the solubilization of the protein content. The plant powder was grinded with a mortar and pestle in the presence of 250 mM Trizma base (Merck), 10 mM of Sodium diethyldithiocarbamate (Merck) and Polyvinylpyrrolidone or PVPP, (Sigma-Adrich) at a rate of 0.5 g PVPP per gram of leaf plant powder. This slurry was sieved using a Miracloth paper (Calbiochem) and then centrifuged (Beckman Ultracentrifuge) at 18 500 *g* for 10 minutes at 4 ºC. The supernatant from the centrifuge step was further purified by dialysis using a membrane with a MWCO (Molecular weight Cut-Off) of 2 kDa (Thermo-Scientific) using as exchange buffer, or dialysate, the buffer best suited for the subsequent usage, at volumes 500 times larger than the sample. The dialysis took place overnight at 4 ºC with mild agitation.

3. Protein quantification

The chosen procedure for this work was a modification of the Lowry method, as first proposed by Bensadoun and Weinstein (1976). In this method the protein reacts with cupric sulphate and tartrate in alkaline solution, which results in the formation of a tetradentate copper-protein complex, reducing the Folin-Ciocalteau reagent. The blue colored, water-soluble product can be quantified at 750 nm and its

intensity is protein dependent. The calibration curve was constructed with bovine serum albumin (BSA), with a protein gradient between 1 and 25 µg, in a final volume of 250 µL (0.004 – 0.1 g/L). Five different concentrations of BSA were prepared, in triplicate, starting from a BSA 0.5 g/L solution (stored at -20 ºC). To this preparation, 50 µL of a 1% (w/v) solution of Sodium desoxycholate and 1 mL of Trichloroacetic acid 10% (w/v) were added, leaving to incubate for 10 min. After incubation, the samples were centrifuged at 10 000 *g* for 5 min, rejecting the supernatant and dissolving the pellet in 1 mL of solution C (combining 1 part of solution B to 100 parts of solution A). The solutions A and B are composed by: Solution A: sodium carbonate deca-hydrate 2% (w/v), sodium hydroxide 0.4% (w/v), sodium tartrate dihydrate 0.1% (w/v) and Sodium Dodecyl Sulphate (SDS) 1% (w/v). Solution B: copper sulphate pentahydrate 4% (w/v). Finally, 100 µL of Folin-Ciocalteau reagent (Sigma), diluted in a ratio of 1:1 with Milli-Q water, was added. The final mixture was vortexed and placed in the dark, at room temperature, for 45 min. The absorbance values were measured at 750 nm in a spectrophotometer (Shimadzu UV-2100) in plastic cuvette of 1 cm path length, against a blank, where the volume of the protein solution is substituted by Milli-Q water. Determination of the protein content of the samples was performed analogously, diluting the sample with Milli-Q water until a final volume of 250 µL. This dilution is critical since the absorbance values measured have to be within the calibration curve limit values. The absorbance is a linear function of the protein concentration between 5 and 50 µg of BSA. Under these conditions, the extrapolation of the protein content was performed based on the tendency line, acquired by the least squares method, from the average value of absorbance relative to the each BSA concentration (Bensadoun and Weinstein, 1976).

4. Protein charge fractioning

4.1.Ion exchange Chromatography

We began the protein fractioning by exploring the ionic interactions of the crude protein extract with two chosen charged resins, in different buffers and pH's, in a technique known as ion exchange chromatography. The two most commonly used resins for ion-exchange chromatography of proteins are Carboxymethyl-cellulose (CM-cellulose) and Diethylaminoethyl-cellulose (DEAE-cellulose). These are granular celluloses which have been chemically modified so that CM-cellulose at neutral pH is negatively charged, i.e. it is a cation exchanger and DEAE-cellulose at neutral pH is positively charged, i.e. it is an anion exchanger. After experiments with these columns we selected a more sensitive approach using an FPLC ("Fast Protein Liquid Chromatography") system (GE Healthcare, Pharmacia) coupled with a anion exchange Mono Q HR 16/10 column (GE Healthcare, Pharmacia).

4.1.1. CM columns

A set of four different pH's values was chosen to test the plant extract fractionation, specifically: 20 mM Maleic acid (Merck) at pH's 1.5; 2; 2.5 and 20 mM of Sodium acetate (Merck) at pH 4. For each chosen pH and buffer, a protein sample was set by dialyzing (MWCO 2 kDa) against the specific buffer. The CM column was set up by dissolving 1 g of CM-52 powder (Whatman) in each chosen buffer and placing this slurry in an empty column holder. For each 1 g of column, 2 mL of protein mixture (from step 2.2.) was placed and allowed to pass through by gravity. Immediately after, the column was washed with 8 mL of the specific buffer at the chosen pH (Wash - W). Finally the sample was eluted by 5 mL of the particular buffer supplemented with 1 M NaCl (Merck) (Elute - E). Later on, all of these samples were dialyzed against 20 mM Tris-HCl (Merck) at pH 7.5 for subsequent testing.

4.1.2. DEAE columns

A set of four different pH's values was chosen to test our plant extract, specifically: 20 mM Tris-HCl (Merk) at pH's 7.5; 8; 8.5 and 20 mM of Piperazine (Merck) at pH 9; 9.5; 10. For each chosen pH and buffer, a protein sample was set by dialyzing against the specific buffer. The DEAE column was set up by dissolving 1 g of DEAE powder (Whatman) in each chosen buffer and placing this slurry in an empty column holder. For each 1 g of column 2 mL of protein mixture (from step 2.2.) was placed and allowed to completely pass through by gravity. Immediately after, the column was washed with 8 mL of the specific buffer at the chosen pH (Wash - W). Finally the sample was eluted by 5 mL of the specific buffer supplemented with 1 M NaCl (Merck) (Elute - E). Later on all of these samples were dialyzed against 20 mM Tris-HCl (Merck) at pH 7.5 for subsequent testing or against 20 mM Piperazine, pH 9.5 for posterior anion exchange chromatography in an FPLC with a Mono Q system in order to remove the salt.

4.1.3. Ion exchange Chromatography – FPLC system

The sample, termed elute, from the DEAE columns at pH 9.5 (see 4.1.2.), was dialyzed against 20 mM Piperazine at pH 9.5 to remove the salt and injected, at different volumes, in a Mono-Q HR 16/10 column using an FPLC system previously equilibrated in the same buffer. The charged group of Q-Sepharose is a quaternary amine which carries a non-titratable positive charge (GE healthcare brochure). The flow rate applied was 1.5 mL/min and the bound proteins were eluted with 1 M of NaCl in 20 mM piperazine buffer, pH 9.5. The sample was collected 1 mL at a time, until no further peak was observed in the chromatogram, and dialyzed against buffer 20 mM Tris-HCl pH 7.5,

5. Protein size fractioning

5.1. Size exclusion chromatography

The eluted sample from the DEAE column, at pH 9.5 (see 4.1.2.), was dialyzed against 20 mM Tris-HCl pH 7.5 and injected (600 µL), in a Superose 12 HR 10/30 column (GE Healthcare – Pharmacia) previously equilibrated in the same buffer. The flow rate applied was 0.5 mL/min and the sample was allowed to pass through the column until 30 mL of running buffer (20 mM Tris-HCl pH 7.5) had passed. Samples of each peak were collected for SDS-PAGE analysis and antimicrobial activity and its size inferred based on the standard curve for this specific column.

5.2. Membrane ultrafiltration

For a more crude size fractioning of the protein mixture following aqueous extraction (see 2.2.), the sample was separated through membrane ultrafiltration using two distinct MWCO, 3 kDa or 10 kDa, using the Centriplus (YM-3, Millipore) and the Amicon Ultra (Millipore) respectively. Independent of pore size, the protein sample was placed, 15 mL at a time, in an ultrafiltration device and centrifuged for at 2500 *g* for 1 h (Beckman Coulter Allegra 25R Centrifuge). Samples were carefully collected and labeled: the fraction that passed through the filter was named, < respective MWCO (e.g. < 3 kDa), and the other, > respective MWCO (e.g. > 3 kDa).

6. Antimicrobial susceptibility testing methodologies

To determine the antifungal activity of different protein fractions, it is necessary a robust, replicable and easy to use method. To this end we selected our method on the guidelines issued by the CLSI (Clinical and Laboratory Standards Institute, formerly NCCLS or National Committee on Clinical Laboratory Standards), with certain alterations (Cantón *et al.*, 2009). CLSI has released three standard methods for liquid antibacterial and antifungal susceptibility testing namely, M31-A2 (CLSI, 2002) for microtiter bacterial testing, M27-A3 (CLSI, 2008a) for macrobroth and microtiter yeast testing and the M38-A2 (CLSI, 2008b) for microtiter mould testing (Hall, 2012).

6.1 Antibacterial susceptibility testing, based on CLSI guideline M31-A2

Based on the CLSI guideline M31-A2 a simplified method was developed to determine the activity of the plant extract against several strains of bacteria, namely: *Bacillus subtilis* (ISA nº 3970), *Listeria monocytogenes* (ISA nº 3952), *Escherichia coli* (ISA nº 3967), *Pseudomonas aeroginosa* (ISA nº 4076) and *Staphilococcus aureus* (ISA nº 3015). For a single test all of these bacteria were first grown overnight in a petri dish containing PCA (Plate Count Agar) at 34 ºC (Annex I.1). After 24 hours, and with fresh cultures at hand, an inoculum was prepared by placing a portion of grown bacteria in 3 mL of sterile salt solution (0.85 % w/v NaCl). This solution was sturdily vortexed for 15 s, its optical density measured in a spectrophotometer (Sprectronic 21, Bauch & Lomb) and adjusted to a transmittance value of 85%, corresponding to $1x10^8$ cells/mL. This inoculum was then diluted 100 fold, with 2x

concentrated MH (Mueller Hinton) liquid media (Annex I) to start the susceptibility test with $1x10^6$ cells/ mL. 100 µL volume of this bacterial dilution was then applied to a well in an ELISA plate, and 100 µL of the intended protein extract was filter sterilized (0.22 µM Milipores seringe filters) and mixed with the bacterial inoculum. The ELISA plate was left overnight at 35 ºC, and on the next day, the results were recorded by reading the transmittance value in a Biotech, Synergy HT microplate reader at 640 nm. Transmittance values below 70 % were registered as (+) sign for the presence of bacterial growth and values above 70 % were recorded as (-) sign for the absence of bacterial growth. All antibacterial tests were conducted in sterile conditions and with the appropriate control wells.

6.2.Antifungal susceptibility testing, based on CLSI guideline M-27 A3

Based on the CLSI guideline M-27-A3, a simplified method was developed to determine the activity of the plant extract against several strains of yeasts, namely: *Candida albicans* (ISA nº 1862), *Candida dubliniensis* (ISA nº 2228), *Candida glabrata* (ISA nº 2163), *Candida parapsilosis* (ISA nº 1811) and *Candida lusitaneae* (ISA nº 2213). For a single test all of these yeasts were first grown over-night in a petri dish containing GYP (Glucose, Yeast, Peptone) at 37 ºC (Annex I.1). After 24 hours, and with fresh cultures at hand, an inoculum was prepared by placing a portion of grown yeasts in 5 mL of sterile salt solution (0.85 % w/v NaCl). This solution was sturdily vortexed for 15 s, its optical density measured in a spectrophotometer (Sprectronic 21, Bauch & Lomb) and adjusted to match the optical density of a 0.5 MacFarland solution (1% Barium chloride, 1% Sulfuric acid). When both solutions registered the same optical density, according with the CLSI guidelines, the bacteria solution has $1x10^6$ – $5x10^6$ cells/ mL. This inoculum was then sequentially diluted two times (1:100 and 1:20), with 2x concentrated PDB pH 7.5 (Potato Dextrose Broth) (Annex I.1), to initiate the test with $5x10^2$ - $2x10^3$ cells/ mL. Then, 100 µL of this final dilution was applied to a well in an ELISA plate and 100 µL of the intended protein extract filter sterilized (0.22 µM Milipores seringe filters) was mixed with the yeast inoculum. The ELISA plate was left 48 hours at 37 ºC and, on the next day, the results were recorded by reading the transmittance value in a Biotech, Synergy HT microplate reader at 640 nm. Transmittance values below 70 % were registered as (+) sign for the presence of bacterial growth and values above 70 % were recorded as (-) sign for the absence of bacterial growth. All antibacterial tests were conducted in sterile conditions and with the appropriate control wells.

6.3. Antifungal susceptibility testing, based on CLSI guideline M-38-P

Based on the CLSI guideline M-38-P a simplified method was developed to determine the activity of the plant extract against a strain of *Aspergilus fumigatus* (ISA nº 3333) and *Aspergilus niger* (ISA nº3214). For a single test these fungus were first grown for 7 days in a petri dish containing PDA (Potato Dextrose Agar) at 35 ºC (Annex I.1). After one week, and with fresh culture at hand, an inoculum was prepared by placing 3 mL of sterile salt solution (0.85 % w/v NaCl, 0.05 % Tween 80) inside the petri dish, with the grown fungus. This solution was removed with a pipette and sieved using a specific filter to separate the hiphae from the fungus spores. The spore solution was sturdily vortexed for 15 s and 10 µL was collected and placed on a Neubauer improved haemocytometer chamber (Marienfeld). By using this specially designed chamber we were able to determine the spore

concentration in the inoculum solution, by counting the number of visible spores using a microscope. With this value, the spore solution was adjusted to $1x10^6$ cells/ mL. Afterward, the fungal solution was again diluted 50 fold with 2x concentrated PDB pH 7.5 (Annex I.1), to obtain a t starting spore solution of $2x10^4$ cells/ mL. Then, 100 µL of this final dilution was then applied to a well in an ELISA plate, and 100 µL of the intended protein extract filter sterilized (0.22 µM Milipores seringe filters)mixed with the fungal inoculum. The ELISA plate was left 48 hours at 37 ºC and, on the next day, the results were recorded by reading the transmittance value in Biotech, Synergy HT microplate reader at 640 nm. Transmittance values below 70 % were registered as a (+) sign for the presence of fungal growth and values above 70 % were recorded as a (-) sign for the absence of fungal growth. All antifungal tests were conducted in sterile conditions and with the appropriate control wells.

7. Poliacrylamide gel electrophoresis (PAGE)

Several forms of PAGE exist and can provide different types of information about the target protein. Nondenaturing PAGE, also called native PAGE, separates proteins according to the net charge, size and shape of their native structure. Denaturing and reducing SDS-PAGE separates proteins primarily by mass because the ionic detergent Sodium dodecyl sulfate (SDS) denatures and binds to proteins to make them evenly negatively charged. Thus, when a current is applied, all SDS-bound proteins in a sample will migrate toward the positively charged electrode (Hames, 1998). SDS-PAGE technique can be further subdivided in Glycine–SDS-PAGE (also known as Laemmli–SDS-PAGE) and Tricine–SDS-PAGE, based on Glycine-Tris and Tricine-Tris buffers used (Haider et al., 2011). Tricine–SDS-PAGE is used preferentially for the optimal separation of proteins lower than 30 kDa, and Glycine–SDS-PAGE for proteins higher than 30 kDa. The different separation characteristics of the two techniques are directly related to the strongly differing electrophoretic mobilities of the trailing ions (Glycine and Tricine) relative to the electrophoretic mobilities of proteins (Schägger, 2006). Once separated by electrophoresis, proteins can be detected in a gel with various stains, such as Coomassie dyes, Coomassie G-250 and R-250 that bind to basic and hydrophobic residues of proteins, or a more sensitive method, for example the silver staining, which deposits metallic silver onto the surface of a gel at the location of the protein band.

7.1. SDS-PAGE

7.1.1. Glycine–SDS-PAGE

Polypeptides were separated in 17 % acrylamide gels and the composition of the different buffers used was as follows:

- Stacking gel- 5 % (w/v) acrylamide (Sigma-Aldrich), 0.13 % (w/v) bis-acrylamide (Sigma-Aldrich), 125 mM Tris-HCl pH 6.8, 0.1 % (w/v) SDS, 0.1 % (w/v) PSA (Ammonium persulfate) (Sigma-Aldrich) and 0.05 % (v/v) TEMED (Tetramethylethylenediamine)(Sigma-Aldrich);
- Resolving gel- 17 % (w/v) acrylamide, 0.1 % (w/v) bis-acrylamide, 375 mM Tris-HCl pH 8.8; 0.01 % (w/v) SDS, 0.03 % (w/v) PSA and 0.05 % (v/v) TEMED;
- Cathode buffer- 25 mM Tris-HCl solution, 192 mM Glycine (Merck);
- Anode buffer- 25 mM Tris-HCl solution, 192 mM Glycine and 0.1 M sodium acetate (Merck);

The first step in preparing the samples for electrophoresis was the protein precipitation in 80% (v/v) cold (- 20 ºC) acetone (Merck) for 30 min at -20 ºC, in order to remove unwanted contaminants. When this time elapsed, the samples were centrifuged at 10.000 g for 10 min, the supernatant discarded, and the precipitate dissolved in a reducing buffer solution termed SBM, which comprised of: 0.08 M Tris-HCl pH 6.8; 0.1 M 2-mercaptoethanol (ME) (Sigma-Aldrich), 2 % (w/v) SDS, 15 % (v/v) glycerol (Merck) and 0.006 (w/v) m-cresol purple (Sigma-Aldrich). After completely dissolved, the protein sample was denatured, by boiling for 3 min in SBM. To sort out our complex polypeptide mixture by size, protein standards with known molecular masses were used. The chosen marker was the Low protein molecular markers (LPM) from Sigma-Aldrich with a protein size range from 14.2 kDa to 66 kDa. The electrophoresis ran at constant current of 45 mA, using an EPS 500/400 power source (Pharmacia/LKB), until the m-cresol purple from the SBM present in the samples reached the end of the gel, then the gel was promptly removed and treated according with the chosen staining protocol.

7.1.2. Tricine–SDS-PAGE

The procedure for Tricine-SDS-PAGE, from acrylamide gel manufacture to sample preparation, is similar to the already mentioned Glycine-SDS-PAGE (see 7.1.1.). The only differences are the running buffers and sample buffer constitution, namely:

- Cathode buffer- 100 mM Tris-HCl solution pH 8.8; 100 mM Tricine (Merck), 0,03 % (w/v) SDS;
- Anode buffer- 200 mM Tris-HCl solution pH 8.9;
- Sample buffer (SB)- 0.01 M Tris-HCl pH 6.8;, 1 % (w/v) SDS, 15 % (v/v) glycerol (Merck), and 0.01 (w/v) “Coomasie Briliant blue R-250”(CBB R-250)(Sigma-Aldrich), 8 M Urea (Panreac) and 1 % (v/v) 2-β-mercaptoethanol (ME).

The chosen marker for this type of electrophoresis was the “Low molecular Weight Marker” (LMW) from GE-Healthcare with a protein size range from 17 kDa to 2.5 kDa. As with the Glycine-SDS-PAGE the electrophoresis ran at constant current of 45 mA, until the Coomasie brilliant blue from the

samples reached the end of the gel. Then, the gel was promptly removed and treated according with the chosen staining protocol.

7.1.3. Native gel electrophoresis

The polypetides were separated in 5 % acrylamide gels and the composition of the different buffers used was as follows:

- Stacking gel- 3 % (w/v) acrylamide; 2.7 % (w/v) bis-acrylamide; 125 mM Tris-HCl pH 6.8; 0.1 % (w/v) PSA and 0.05 % (v/v) TEMED;
- Resolving gel- 5 % (w/v) acrylamide, 2.7% (w/v) bis-acrylamide, 375 mM Tris-HCl pH 8.8; 0.03 % (w/v) PSA and 0.05 % (v/v) TEMED;
- Cathode buffer- 25 mM Tris-HCl solution, 192 mM Glycine (Merck);
- Anode buffer- 25 mM Tris-HCl solution, 192 mM Glycine and 0.1 M sodium acetate (Merck);

To maintain the polypeptide native state, all the samples used in this type of analysis were lyophilized (Edwards, Modulyo), and ressuspended in “native sample buffer” (SBN), which comprised of: 0.08 M Tris-HCl pH 6.8, 15 % (v/v) glycerol and 0.006 (w/v) m-cresol purple. The electrophoresis ran at constant current of 20 mA, at 6 ºC, until the m-cresol from the sample reached the end of the gel. Then the gel was promptly removed and treated according with the chosen staining protocol.

7.2. Staining of PAGE gels

In this work three distinct types of gel staining procedures were employed with different detection limits: Coomasie Brilliant Blue R 250 (CBB R-250) (*50-100 ng)* > Coomassie G-250 (colloidal Coomasie) (*10-20 ng)* > Silver staining (1-10 ng).

The most common method for protein detection is staining with the coomassie dye. Several coomassie stain procedures exist in the literature that use either the G-250 (“colloidal”) or R-250 form of the dye. Despite of the recipe, all rely on the capability of the coomasie dye to bind to basic and hydrophobic residues of proteins, and in the process, changing its color from reddish-brown to intense blue (Hames, 1998). In the Coomassie R-250 staining procedure the polypeptides were fixed to the gel by placing it in a 10 % (w/v) TCA (Trichloroacetic acid) (Sigma-Aldrich) solution, for 10 min. After removing the TCA solution and further washing the gel with water, the staining solution was added (0.25 % (w/v) CBB-R250; 25 % (v/v) 2-propanol (Merck) and 10% (v/v) acetic acid (Merck) and left to act with a gentle swirl, for 2 hours. Finally, the gel had to be discolored to visualize the polypeptide bands with a solution containing: 25 % (v/v) 2-propanol and 10 % (v/v) acetic acid.

In the Coomasie G250 procedure, the polypeptides were fixed to the gel by adding a solution of 34 % (v/v) Methanol and 2 % Phosphoric acid (w/v) and allowing it to act for 2 hours. After a repetitive wash with water the gel was stained for 48 hours in the staining solution, containing: 1.1 % Coomasie G-250, 34 % (v/v) Methanol.

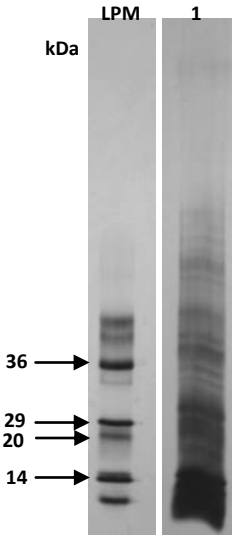
The silver staining procedure is a highly sensitive procedure ideal for samples with low polypeptide content but prone to false positives such as artifacts and “ghost” bands (Hames, 1998). In this method the polypeptides were fixed to the gel by dipping it for 20 min in a solution of 50 % (v/v) Methanol, 12 % (v/v) Acetic acid and 0.05 % (v/v) Formaldehyde. After this step the gel was washed three times with 50 % (v/v) Ethanol, one time with a 0.02 % (w/v) solution of Na2S2O3.5H2O and again three times with water. The staining itself was accomplished by immersing the gel in a solution of 0.2 % (w/v) Silver nitrate and 0.075 % (v/v) of Formaldehyde, followed by two washes in a solution of 6 % (w/v) Sodium carbonate, 0.05 % (v/v) Formaldehyde and 0.4 % (v/v) Na2S2O3.5H2O and left in until the polypeptide bands were clearly visible. The staining reaction is stopped by immersing the gel in a 50 % (v/v) Methanol and 12 % (v/v) Acetic acid solution (Chevallet et al., 2006).

Results

1. Preliminary analysis of the aqueous extract from the leafs of the Spurge flax plant

We began our work by analyzing the antimicrobial activity and the total protein content of the Spurge flax leaf, extracted through the two step extraction protocol (see Material and Methods 2.2.). The proteinaceous content in the sample was quantified by the Lowry method, subjected to a Glycine-SDS-PAGE and a Native-PAGE to visualize its polypeptide profile and protein complexity. Finally, the antimicrobial potential of this aqueous extract was evaluated through a series of antimicrobial susceptibility tests based on the CLSI methodologies.

(A) SDS-PAGE



(B) Native - PAGE

(C) Antimicrobial evaluation

Antimicrobial activity		
		Total Protein Content (TPC)
Bacteria	***Bacillus subtilis***	+
	Listeria monocytogenes	+
	Escherichia coli	-
	Pseudomonas aeroginosa	-
	Staphilococcus aureus	+
Yeasts	***Candida albicans***	+
	Candida dubleniensis	+
	Candida glabrata	+
	Candida parapsilosis	+
	Candida lusitaneae	+
Fungus	***Aspergillus niger***	+
	Aspergillus fumigatus	+

Figure 1: Glycine SDS-PAGE (A) and Native-PAGE (B) of the total protein content present in the aqueous fraction of the leaf of the Spurge flax plant, and its corresponding antimicrobial activity in selected microorganisms (C). Figure 1 A) Glycine-SDS-PAGE (silver stained) of the total protein content extracted from the leaf of the Spurge flax plant, lane 1 containing 15 µg of total protein, LPM refers to the low protein marker. Figure 1 B) Native PAGE (coomasie G-250 stained) of the total protein content in the leafs of the Spurge flax plant, lane 2 containing 150 µg of total protein; Figure 1 C) table depicting the results of the antimicrobial susceptibility tests conducted with 100 µl of the aqueous extract of the leafs of the Spurge flax plant at a protein concentration, as determined by Lowry modified method, of 0.1758 µg/µl. (+) Value of transmittance below 70 %, considered lack of growth, therefore active compound; (-) Value of transmittance over 70 %, considered as presence of growth and therefore non-active compound.

Observing the antimicrobial results alone (Figure 1 C) it is clear that we are in the presence of a aqueous plant extract with a high biological activity against Gram (+) bacteria (*B. subtilis*, *L. monocytogenes*, *S. aureus*), yeasts and fungus and with no activity towards Gram (-) bacteria (*E.coli* and *P. aeroginosa*). Regarding the protein profile, it possesses a somewhat complex polypeptide composition (Figure 1 A), but a rather simple Native PAGE protein profile (Figure 1 B), with in-between 6 to 10 possible distinct proteins or isoforms of the same protein.

Obviously, even with the selective extraction protocol, it is not possible to conclude that the antimicrobial activity of the aqueous extract is of the entire responsibility of the proteinaceous content. It is a known fact that the Spurge flax plant contains several non-proteinaceous antimicrobial compounds (Deiana *et al*., 2003). Therefore an experiment to determine the influence of the Spurge flax's peptide/protein content in the antimicrobial activity needed to be devised. Considering this, the plant extract was subjected to different protein denaturing treatments and its antimicrobial activity determined, the treatments were as follows: **a)** aqueous extract of the Spurge flax plant was incubated 2 hours at 37 ºC with a 10 mg/mL solution of Trypsin, a protease that cleaves peptides on the C-terminal side of lysine and Arginine amino acid residues (Keil, 1971); **b)** aqueous extract of the Spurge flax plant was incubated 2 hours at 37 ºC with 16 mg/ mL of Pronase, a mixture of several nonspecific endo and exo proteases that generally digest proteins down to single amino acids (Trop and Birk,1970); **c)** aqueous extract of the Spurge flax plant was boiled,100 ºC, for 10 minutes to denature its protein fraction.

		Antimicrobial activity			
		Total Protein Content (TPC)	**TPC Trypsin**	**TPC Pronase**	**TPC 100 ºC**
Bacteria	***Bacillus subtilis***	+	+	-	+
	Listeria monocytogenes	+	+	-	+
	Escherichia coli	-	-	-	-
	Pseudomonas aeroginosa	-	-	-	-
	Staphilococcus aureus	+	+	-	+
Yeasts	***Candida albicans***	+	+	-	+
	Candida dubleniensis	+	+	-	+
	Candida glabrata	+	+	-	+
	Candida parapsilosis	+	+	-	+
	Candida lusitaneae	+	+	-	+
Fungus	***Aspergillus niger***	+	+	-	+
	Aspergillus fumigatus	+	+	-	+

Table 1: Results of the antimicrobial susceptibility test for the Spurge flax plant aqueous extract subjected to three different protein denaturing treatments. A volume of 100 µL of the Spurge flax plant aqueous extract was tested at a concentration of 0.13 µg/µL (as determined by the Lowry method). The denaturing treatments were as follows: 2 hours at 37 ºC with 10 mg/mL Trypsin; 2 hours at 37 ºC with 16 mg/mL of Pronase, 10 min at 100 ºC. Adequate controls were set up, where the organisms were grown in the presence of 10 mg/mL Trypsin and 16 mg/mL Pronase solution. (+) Value of transmittance below 70 %, considered lack of growth, therefore active compound; (-) Value of transmittance over 70 %, considered as presence of growth and therefore non-active compound..

Following this experiment (see Table 1, above), the aqueous plant extract, maintained its spectrum of activity when incubated with Trypsin or submitted to 100 ºC for 10 minutes. The mixture of proteases, e.g. Pronase, on the other hand, had a deleterious effect on the extracts biological activity, clearly indicative of a proteinaceous nature, most likely the result of protein/peptide digestion from the different enzymes present in the Pronase mixture. To try and visualize the effect of the different protein denaturing treatments on the polypeptide profile, the different samples were subjected to an SDS-PAGE analysis and compared to a non-treated sample (lane TPC).

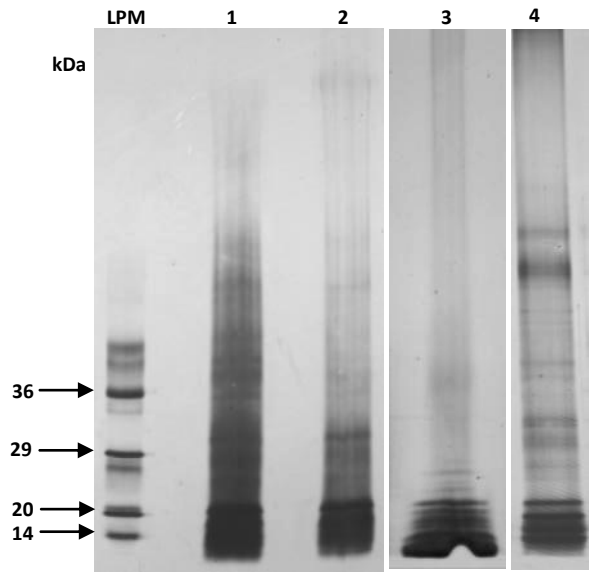


Figure 2: Glycine SDS-PAGE (Silver staining) portraying the effect of different protein denaturing treatments on the polypeptide profile of the aqueous extract of the Spurge flax plant. Lane 1: Total protein in the aqueous extract of the Spurge flax plant. Lane 2: Total protein in the aqueous extract of the Spurge flax plant subjected to 100 ºC/ 10 min. Lane 3: Total protein in the aqueous extract of the Spurge flax plant incubated with 16 mg/mL of Pronase protease mixture. Lane 4: Total protein in the aqueous extract of the Spurge flax plant incubated with 10 mg/mL trypsin.

Analyzing the SDS-PAGE above (Figure 2), it is clear that boiling (Lane 2) and Trypsin digestion (Lane 4) treatments had no significant effect on the polypeptide profile, supporting the previous biological activity results. In the case of the Pronase incubated sample (lane 3), there is a significant difference in the polypeptide profile, when compared with TPC control (Lane 1). It is visible an accumulation of low molecular weight polypeptides, characteristic of protein degradation.

The conjuction of these results, namelly, loss of antimicrobial activity and changes in the polypeptide profile due to pronase, leads us to hypothezised, that the active compound on the aqueous extract of the Spurge flax plant is proteinaceous in nature. This proteinaceous molecule is: i) extremelly resistant to high temperatures, indicative of either, an highly stable protein, or a simpler hydrophilic polypeptide or peptide with a crude tertiary structure, a good example of this is the knottin family of antimicrobial plant peptides (Doon, 2002; Colgrave and Craik, 2004); ii) and resistant to trypsin digestion, indicative of an absence or low number of the positively charged residues Lysine and Arginine on its primary structure (Keil, 1971).

2. Charge fractioning the total protein content present in the aqueous extract of the Spurge flax leafs, using CM and DEAE columns

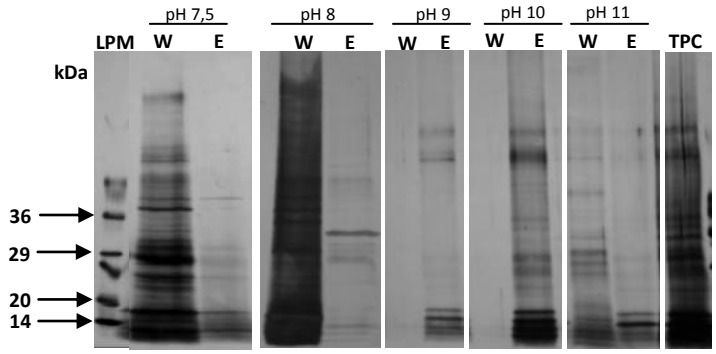
After coming to the conclusion that we were in the presence of a antimicrobial protein/peptide, the next step was the purification of our target biomolecule. Therefore, to remove unwanted protein and peptide contaminants, we began fractioning our plant extract through by ion exchange chromatography. To this effect, 2 mL samples of plant extract were dialyzed against a variety of buffers at different pH values, and separately applied to CM (cation exchange) and DEAE (anion exchange) chromatography columns, with the final purpose of finding a pH range at which our protein interacted with the column, and in this way allowing a separation process. A low range pH's were applied to the CM cellulose and conversely the high range pH values to the DEAE cellulose. Ideally, samples that do not interact with the chromatography columns will be removed by the washing buffer (termed samples W) and those that do interact will be eluted with a buffer with 1 M of NaCl salt (termed samples E). All samples were subjected to an SDS-PAGE and a simplified version of the antimicrobial susceptibility test, where only the *Candida albicans* and *Aspergillus niger* microorganisms were tested.

Regarding the interaction with the CM-columns (see Figure 3 C), independent of the pH value, only in the wash step samples (W) it was possible to detect some degree of antimicrobial activity. The SDS-PAGE results (Figure 3 B) for these samples, shows that there is no clear difference in polypeptide profile between the wash samples and the eluted ones, even though only W samples possess biological activity. Clearly at the tested pH values our protein/peptide is unable to bind with the negatively charged column. Probably, by lowering the pH values even further (pH 1.5 or 1), our biomolecule might interact with the CM column matrix, but those are extreme pH values, possibly also affecting the peptide bond (Vlasak and Ionescu, 2011), and hence too low for subsequent steps. This inability to interact with the CM columns it is in accordance with our past result with Trypsin. Resistance to Trypsin comes from the absence of negatively charged arginine and lysine residues, probably conferring a highly negative charge to our protein/peptide.

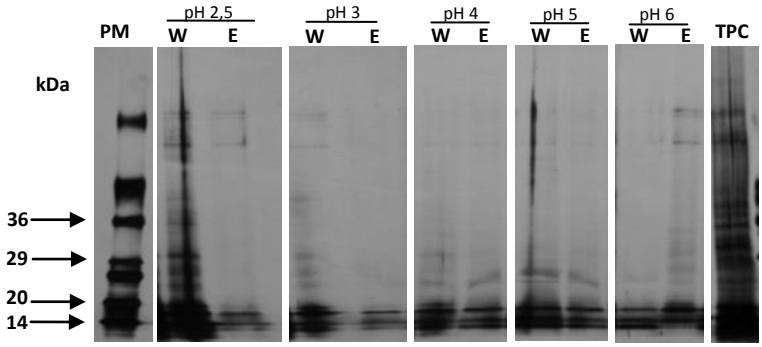
Focusing on higher pH values, and on the DEAE column samples, we observe that at any pH higher than 8, the eluted (E) samples possess antimicrobial activity (see Figure 3 C). Thus, our biologically active protein/peptide interacts with the DEAE column (anion exchanger). Consequently our target protein is negatively charged in this pH range. The polypeptide profile from these samples (see Figure 3 B), the eluted samples from the DEAE column appears to have a simpler profile, when compared with the total protein (lane "total"), therefore constituting a valid purification step.

Apparently this molecule is highly resistant to extreme pH values, maintaining its biological activity from pH 2 through pH 10, probably the result of a hydrophilic primary structure and a simple tertiary structure (Doon, 2002). In general, these results reveal that our target molecule is most likely a hydrophilic and highly charged anionic molecule, only able to interact with the positively charged DEAE column.

(A) DEAE column



(B) CM column



(C) Antifungal tests

	Antimicrobial activity - DEAE column									
	pH 7.5		pH 8		pH 9		pH 10		pH 11	
	W	E	W	E	W	E	W	E	W	E
C. albicans	+	-	-	+	-	+	-	+	-	+
A. niger	+	-	-	+	-	+	-	+	-	+
	pH 2.5		pH 3		pH 4		pH 5		pH 6	
	W	E	W	E	W	E	W	E	W	E
C. albicans	+	-	+	-	+	-	+	-	+	-
A. niger	+	-	+	-	+	-	+	-	+	-

Figure 3: Glycine SDS-PAGE' containing the charge fractioning at different pH through CM and DEAE columns of the total protein extracted from the leafs of the Spurge flax plant and their corresponding antimicrobial activity in selected microorganisms: **A)** Glycine-SDS-PAGE (silver staining) containing the fractions of the total protein content in DEAE-columns at different pH values: W – 15 µg sample of the washing step; E – 15 µg sample of the elution step. **B)** Glycine-SDS-PAGE (silver staining) containing the fractions of the total protein content in CM-columns at different pH values: W – 15 µg sample of the washing step; E – 15 µg sample of the elution step; TPC refers to the total protein samples; LPM refers to the low protein marker; **C)** table depicting the results of the antifungal susceptibility tests conducted with 100 µL of the different fractions with their pH adjust to 7.5. (+) Value of transmittance below 70%, considered lack of growth, therefore active compound; (-) Value of transmittance over 70%, considered as presence of growth and therefore non-active compound.

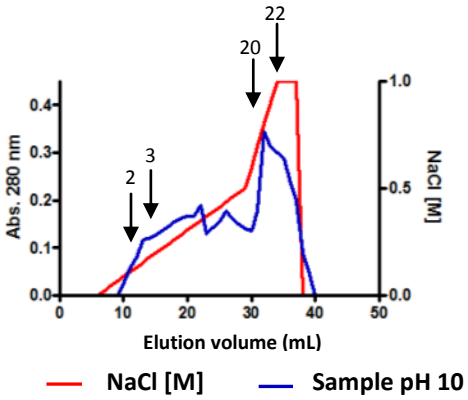
3. Anion exchange chromatography in an FPLC system of the aqueous extract of the Spurge flax leafs

Since the target molecule interacted with the DEAE column, we set out to further purify our protein extract through anion exchange chromatography using a Mono Q column (anion exchange) in an FPLC system. The sample, eluted from the DEAE purification step at pH 10, after dialyzes to remove the salt while maintaining the same pH, was injected into the column. By reading the absorbance at 280 nm we were able to register when the elution began and collect samples of the different peaks until the end of elution. All samples were subjected to an SDS-PAGE and to a simplified version of the antimicrobial susceptibility tests, where only the *Candida albicans* and *Aspergillus niger* microorganisms were tested.

By combining the results of the FPLC chromatogram and the antimicrobial activity it is possible to observe (Figure 4 A and B) that the eluted samples found to possess biological activity are sample numbers 2, 3, 20, 21 and 22 eluted at two salt concentrations intervals, [0.1- 0.4] M and [0.7 – 1] M respectively (represented by the arrows in the chromatogram). Therefore there is no clear peak in the chromatogram with biological activity. What is clear is that the active biomolecule, somehow, attached weakly, and at the same time strongly, to the charged column. Looking at the SDS-PAGE from these same samples (Figure 4 C), we notice some differences in polypeptide profile throughout the chromatogram, but with no clear polypeptide band or pattern that one might correlate with the presence of biological activity. In all the eluted samples the predominant bands in the SDS-PAGE gels are small sized, below the 20 kDa marker.

The Mono-Q column was re-tested using higher sample volumes and different salt gradients (results not showed), but with similar results, making the purification through anion exchange chromatography in these conditions impossible.

(A) Salt gradient in Mono Q at pH 10



(B) Antimicrobial tests

Antifungal activity - DEAE column																							
	Fractions from the Mono – Q column																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
C. albicans	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-
A. niger	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-

(C) SDS-PAGE

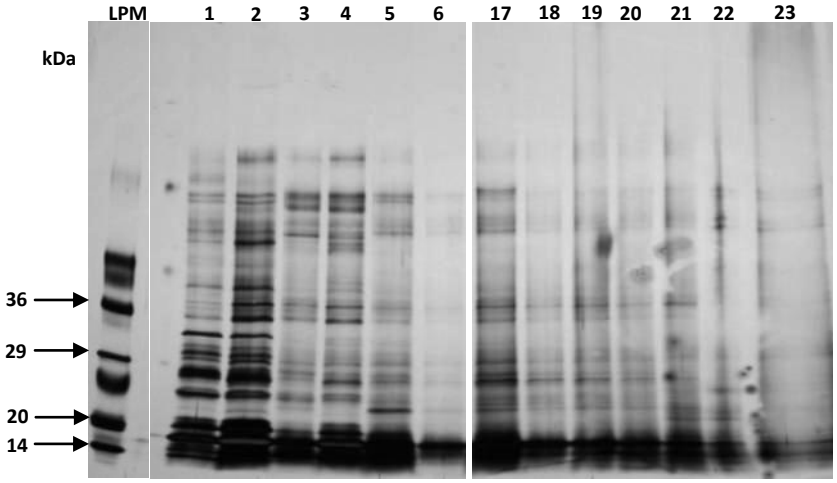


Figure 4: Chromatogram and corresponding antimicrobial test and SDS-PAGE of the different peaks generated by anion affinity chromatography (Mono-Q column) at pH 10 in an FPLC system, of the DEAE fractioned extract from the leafs of the Spurge flax plant: Figure 4 **A)** Chromatogram of the elution at different NaCl concentrations of the plant fractioned extract, through anion exchange chromatography using a Mono-Q column in an FPLC system. Working buffer: 20 mM piperazine pH 10; Elution buffer: 20 mM piperazine pH 10 with 1 M NaCl. Injection of 50 mL of sample in a total of 1.5 mg of total protein extract. Flow of 1 ml/ min and samples were constantly measure by absorbance readings at 280 nm.Arrow (⟶) illustrates where the antimicrobial activity was found in the chromatogram. Figure 4 **B)** Table depicting the results of the antifungal susceptibility tests conducted with 100 µL of the different fractions with their pH adjust to 7.5. (+) Value of transmittance below 70 %, considered lack of growth, therefore active compound; (-) Value of transmittance over 70 %, considered as presence of growth and therefore non-active compound. Figure 4 **C)** Glycine-SDS-PAGE (silver staining) containing the different fractions from the Mono-Q column of the total protein content. 500 µL of each sample was applied in the gel. LPM refers to the low protein marker.

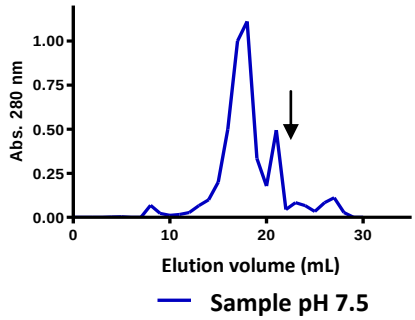
4. Protein size fractioning through Gel Filtration chromatography in an FPLC system of the Spurge leaves extract

Having proven that our target protein interacted with the DEAE column, but unable to further purify it through ion exchange chromatography, we set out to size fractionate our protein extract, through gel filtration chromatography, using an Superose 12 HR 10/30 column in an FPLC system. A dialyzed sample from the DEAE purification step at pH 10, in 100 mM Tris-HCl pH 7.5. was injected into the Superose 12 column. By reading the absorbance we were able to register when the sample passed and began to collect 1 mL at a time until the elution ended. All samples were subjected to an SDS-PAGE and to a simplified version of the antimicrobial susceptibility tests, where only the *Candida albicans* and *Aspergillus niger* microorganisms were tested.

Looking at the results in Figure 5 A and B, it is clear that our target biomolecule elutes from the Superose 12 column at the end of the run, at elution volumes [21 – 23] mL. To estimate protein size we need to compare our elution volume to the elution volumes of standard proteins (Annex I.2). Considering that the smallest protein in the standard curve (Cytochrome C – 12.4 kDa) elutes at 15.7 mL, and the late elution of our protein, two possibilities arise: either our target molecule interacts non-specifically with the column matrix, with the consequential delay of elution volume, or, our active biomolecule is smaller than the smallest protein in the standard curve, i.e., the Cytochrome C, 12.4 kDa. An analysis of the polypeptide profile from the SDS-PAGE gels (Figure 5 C) supports the latter hypothesis. The active fraction eluted from the column between [21-23] mL, only possesses polypeptide bands below the 20 kDa marker (especially fraction 21). Whereas, the common polypeptide profile associated with a sample of leaf extract from the DEAE columns elutes at around 10 to 17 mL, with no antimicrobial activity.

Previous results also support the presence of an antimicrobial polypeptide or peptide. Resistance to Trypsin is more likely in a small anionic polypeptide or peptide (Keil, 1971). Resistance to boiling temperatures although common in proteins found in hyperthermophilic organisms (Kim *et al.*, 2000), is more commonly found in smaller hydrophilic polypeptides and peptides with a very simple or with a complete absence of tertiary structure (Bischof an He, 2005; Oliveira *et al.*, 2007). This low size, low complexity, also fits its extreme resistance to pH values, since the effect of very high or low pH values on a protein molecule, depends primarily on the complexity of its tertiary structure, and consequently on the molecules size and primary structure diversity (Doon, 2002).

(A) Size exclusion chromatography Superose 12 HR 10/30



(B) Antimicrobial tests

Antifungal activity - DEAE column																						
	Fractions from the Mono – Q column																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
C. albicans	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
A. niger	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+

(C) SDS-PAGE

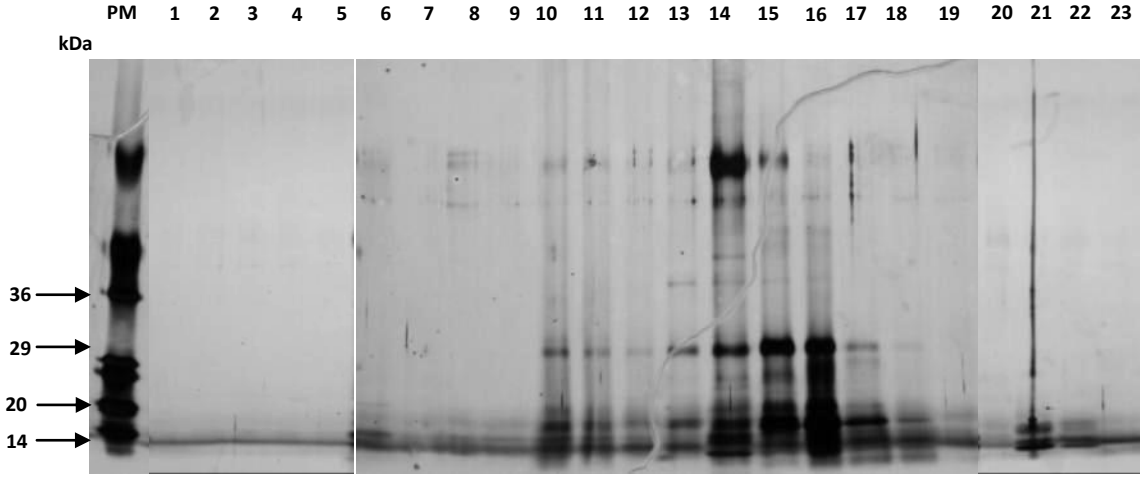


Figure 5: Chromatogram and corresponding antimicrobial test and SDS-PAGE of the different peaks generated by size exclusion chromatography in a Superose 12 HR 10/30 column in an FPLC system, of the DEAE fractioned extract from the leafs of the Spurge flax plant: Figure 5 **A)** Chromatogram of the elution at different volumes of the plant fractioned extract, through size exclusion chromatography using a Superose 12 HR 10/30 column in an FPLC system. Working buffer: 100 mM Tris-HCl at pH 7.5. Injection of 600 µL of sample corresponding to 100 µg of total protein extract. Flow of 0.5 ml/ min and samples were constantly measure by absorbance readings at 280 nm. Arrow (⟶) illustrates where the antimicrobial activity was found in the chromatogram. Figure 5 **B)** Table depicting the results of the antifungal susceptibility tests conducted with 100 µL of the different fractions with their pH adjust to 7.5. (+) Value of transmittance below 70 %, considered lack of growth, therefore active compound; (-) Value of transmittance over 70 %, considered as presence of growth and therefore non-active compound. Figure 5 **C)** Glycine-SDS-PAGE (silver staining) containing the different fractions from the Superose 12 column. 500 µL of each sample was applied in the gel. LPM refers to the low protein marker.

5. Size fractioning the protein content from the aqueous extract of the leafs of Spurge flax plant using ultrafiltration membranes with MWCO of 10 kDa and 3 kDa

Given the expected size of the active protein/peptide responsible for the antimicrobial activity, an attempt to fractionate them by ultrafiltration was made using membranes with a MWCO of 10 kDa. A sample from the DEAE purification step at pH 10, corrected to pH 7, was used and the fractionated samples were termed "< 10 kDa" and "> 10 kDa". For each sample we performed a full scale antimicrobial test, and, as in previous experiments, we tested whether our active biomolecule was proteinaceous in nature by submitting it to three distinct protein denaturing treatments: Trypsin and Pronase incubation and boiling at 100 ºC/10 min.

		Antimicrobial activity							
		Non treated		Trypsin (10 mg/mL)		Pronase (16mg/mL)		100 ºC/10 min	
		< 10 kDa	>10 kDa	< 10 kDa	>10 kDa	< 10 kDa	>10kDa	< 10kDa	>10kDa
Bacteria	*Bacillus subtilis*	+	-	+	-	-	-	+	-
	Listeria monocytogenes	+	-	+	-	-	-	+	-
	Escherichia coli	-	-	-	-	-	-	-	-
	Pseudomonas aeroginosa	-	-	-	-	-	-	-	-
	Staphilococcus aureus	+	-	+	-	-	-	+	-
Yeasts	*Candida albicans*	+	-	+	-	-	-	+	-
	Candida dubleniensis	+	-	+	-	-	-	+	-
	Candida glabrata	+	-	+	-	-	-	+	-
	Candida parapsilosis	+	-	+	-	-	-	+	-
	Candida lusitaneae	+	-	+	-	-	-	+	-
Fungus	*Aspergillus niger*	+	-	+	-	-	-	+	-
	Aspergillus fumigatus	+	-	+	-	-	-	+	-

Table 2: Results of the antimicrobial susceptibility test for the Spurge flax plant DEAE purified protein, separated through ultrafiltration in a Centriplus apparatus with a 10 kDa MWCO. Fraction lower than MWCO termed < 10 kDa and fraction higher MWCO termed >10 kDa. Both fractions were subjected to three different protein denaturing treatments, as follows: 2 hours at 37 ºC with 10 mg/mL Trypsin; 2 hours at 37 ºC with 16 mg/mL of Pronase, 10 min at 100 ºC. For the antimicrobial tests a volume of 100 µl of each sample was tested at a concentration of 0.11 µg/µl (as determined by the Lowry method). Controls were set up, where the organisms were grown in the presence of 10 mg/mL Trypsin and 16 mg/mL Pronase solution. (+) Value of transmittance below 70 %, considered lack of growth, therefore active compound; (-) Value of transmittance over 70 %, considered as presence of growth and therefore non-active compound.

In the ultrafiltration size separation, only the fraction under 10 kDa ("<10 kDa") displayed a level of antimicrobial activity (see Table 2, above) similar to the non-treated sample. At the same time this faction portrayed the same resistance to different protein denaturing treatments, except Pronase incubation. These results supported our view, that our active biomolecule is a small weight protein or peptide, with a molecular weight lower than ultrafiltration apparatus MWCO, i.e., 10 kDa.

Considering our protein possible molecular mass, a change in the method of polypeptide separation in an attempt to separate the "< 10 kDa" protein fraction, from the ultrafiltration step, was made using a Trycine SDS-PAGE gel.

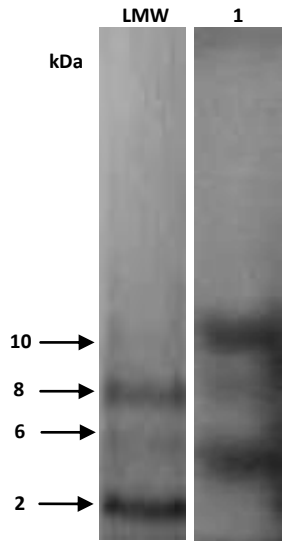


Figure 6: Image of a Trycine SDS-PAGE (silver stained) of the "< 10 kDa" protein fraction from the ultrafiltration step of the DEAE eluted fraction from the leafs of the Spurge flax plant. Lane 1: containing 15 µg of the "<10 kDa" fraction; Lane LMW: refers to the low molecular weight protein marker.

As expected, by means of the Trycine SDS-PAGE (Figure 6), we were able to properly separate and visualize the polypeptide complexity of our "< 10 kDa" fraction. This fraction displays a simple content, with only three distinct polypeptide bands, with a molecular weight between [2- 10] kDa. But although a simple polypeptide profile, this result is still insufficient to be considered a full purification protocol. Any combination of the three bands is possibly larger than the membrane MWCO therefore we are probably facing three distinct peptides, and further purification must ensue.

To determine which of these three polypeptide bands might be our active molecule we submitted our "<10 kDa" sample to a subsequent ultracentrifugation step using an MWCO of 3 kDa. Using the under 10 kDa sample from the first ultrafiltration step, we performed another size separation using an ultrafiltration apparatus with an MWCO of 3 kDa. The new samples were, termed "< 3 kDa" and "> 3 kDa". For each of these samples we performed a full scale antimicrobial test, and, as in previous experiments, we tested whether our active biomolecule was proteinaceous in nature by submitting it to three distinct protein denaturing treatments: Trypsin and Pronase incubation and boiling at 100 ºC/10 min.

		Antimicrobial activity							
		Non treated		Trypsin (10 mg/mL)		Pronase (16mg/mL)		100 ºC/10 min	
		< 3 kDa	> 3 kDa	< 3 kDa	>3 kDa	< 3 kDa	> 3kDa	< 3 kDa	>3 kDa
Bacteria	*Bacillus subtilis*	+	-	+	-	-	-	+	-
	Listeria monocytogenes	+	-	+	-	-	-	+	-
	Escherichia coli	-	-	-	-	-	-	-	-
	Pseudomonas aeroginosa	-	-	-	-	-	-	-	-
	Staphilococcus aureus	+	-	+	-	-	-	+	-
Yeasts	*Candida albicans*	+	-	+	-	-	-	+	-
	Candida dubleniensis	+	-	+	-	-	-	+	-
	Candida glabrata	+	-	+	-	-	-	+	-
	Candida parapsilosis	+	-	+	-	-	-	+	-
	Candida lusitaneae	+	-	+	-	-	-	+	-
Fungus	*Aspergillus niger*	+	-	+	-	-	-	+	-
	Aspergillus fumigatus	+	-	+	-	-	-	+	-

Table 3: Results of the antimicrobial susceptibility test for the Spurge flax plant DEAE purified protein, separated through ultrafiltration in a Centriplus apparatus with a 10 kDa MWCO and afterword separated through ultrafiltration in a Centriplus apparatus with a 3 kDa MWCO. Fraction lower than MWCO termed "< 3kDa" and fraction higher MWCO termed ">3 kDa". Both fractions were subjected to three different protein denaturing treatments, as follows: 2 hours at 37 ºC with 10 mg/mL Trypsin; 2 hours at 37 ºC with 16 mg/mL of Pronase, 10 min at 100 ºC. For the antimicrobial tests a volume of 100 µl of each sample was tested at a concentration of 0.10 ug/ul (as determined by the Lowry method). Controls were set up, where the organisms were grown in the presence of 10 mg/mL Trypsin and 16 mg/mL Pronase solution. For the antimicrobial tests a volume of 100 µl of each sample was tested at a concentration of 0.1 µg/µl (as determined by the Lowry method). (+) Value of transmittance below 70 %, considered lack of growth, therefore active compound; (-) Value of transmittance over 70 %, considered as presence of growth and therefore non-active compound.

By analyzing the above Table 3, it is clear that our biologically active antimicrobial compound is under the 3 kDa range, displaying the already common susceptibility to Pronase incubation and resistance to the other protein denaturing treatments. By comparing this final result, with the previous Tricine SDS-PAGE analysis (see Figure 6), one can assume that our target molecule is most likely the smallest of the three discernible peptide bands in the SDS-PAGE.

We have to take into account that some degree of error could be present in the ultrafiltration size separation and the in the Tricine SDS-PAGE separation, therefore an additional Tricine SDS-PAGE of the "< 3 kDa sample" is mandatory to purify our target peptide and also determine if this three step protocol is successful in the complete isolation of our target peptide.

In summary we are in the presence of a new, highly active, anionic plant antimicrobial peptide, present in leafs of the Spurge flax plant. This highly anionic peptide possesses an extreme resistance to acetone, temperature and extreme pH values, probably with a low number, or complete absence of lysine and arginine residues in its primary structure. Assuming all these characteristics, this new AMP is impossible to group in any of the more relevant families of AMP therefore constituting a possible new type of plant AMP.

Conclusions

This work had two clearly defined and dependent objectives: the first was to determine the antimicrobial activity of the proteinaceous content of Spurge flax plant leafs; the second was an attempt to fractionate, isolate and identify the proteinaceous molecule responsible for the observed antimicrobial activity. From the beginning we were able to determine that the target molecule is (i) highly active, (ii) with a wide antimicrobial spectrum, and (iii) proteinaceous in nature. (i) Highly active - protein concentrations, as low as 0.1 µg/µL, possessed antimicrobial activity. (ii) Wide spectrum - since it is active against microorganisms belonging to groups as diverse as Gram (+) bacteria (*B.subtilis*, *L.monocytogenes*, *S. aureus*), yeasts (*C. dubleniensis, C. glabrata, C. parapsilosism C. lusitaneae*) and fungus (*A. niger, A. fumigatus*). (iii) Proteinaceous - is easily degraded by Pronase incubation.

Having determined its proteinaceous nature we engaged on our second sub-objective, to fractionate, isolate and identify our antimicrobial protein and/or peptide. To this end, after, total protein purification and fractionating, we determined that our active biomolecule is a small sized hydrophilic antimicrobial plant peptide with a possible molecular weight in-between [2 - 6] kDa, with a negative global net charge (anionic), extremely resistant to acetone, temperature denaturation, and trypsin digestion, and probably with an absence of the positively charged Lysine and Arginine residues. Given these characteristics it was impossible, without further studies, to group this new AMP in one of the known families of cationic plant AMP's.

In our attempt to fractionate and isolate by charge and size this new plant AMP we developed a simple three step purification protocol, comprising a selective extraction protocol, followed by a bulk anion exchange chromatography step in DEAE columns and a final ultrafiltration step with MWCO of 3 kDa.

In the future, the "<3 kDa" fraction from the ultrafiltration step, must be re-tested and analyzed through a Tricine SDS-PAGE to determine if the purification protocol devised is correct. If proven properly purified, our new AMP primary structure must be sequenced through mass spectrometry (MS). In addition, further tests must be conducted regarding the spectrum of antimicrobial activity of this new AMP and its level of toxicity towards human cells should also be analyzed. These final steps in the work are currently being performed by Drª Ana Lima, belonging to the Disease and Stress Biology Lab in Instituto Superior de Agronomia.

Chapter III

Agrobacterium sp. Mediated Transformation of four *Rosa hybrida* cultivars with the antifungal polypeptide BLAD

Introduction

Classical plant breeding techniques came into existence when man learned that crop plants could be artificially mated or cross-pollinated to improve the characters of the plant. Desirable characteristics from different parent plants, such as, faster growth, higher yields, pest and disease resistance, larger seeds, or sweeter fruits, could then be combined in the offspring (Ceasar and Ignacimuthu, 2012). On the onset of the 21st century scientists have developed a set of very different improvement techniques such as, mutational breeding, molecular diagnostic tools, tissue culture/micropropagation, and plant genetic, that pick up on the classical plant breeding and improve on it, making the process faster, and achievable for more recalcitrant species (Yawson and Yawson, 2008). Modern roses, especially *Rosa hybrid*a varieties, have been developed through traditional breeding techniques, among thousands of existing cultivars, but they tend to be difficult subjects for plant breeding techniques due to their high heterozygosity, different ploidy levels, and difficulties in sexual reproduction (Office of the Gene and Technology Regulator, 2009).

Plant genetic transformation, or engineering, such as the commonly used, *Agrobacterium* .sp mediated, overcomes the classical breeding constraints and permits the introduction of foreign genes into the target plant genome, thus allowing for a quick and extremely guided plant development (Brown and Thorpe, 1995)**.** On top of that, considering that the cut-flower business possesses an high marginal profit and less legislation constraints for genetically modified crops (GM) (Zuker *et al*., 1998)**,** rose is an ideal target for applying plant transformation techniques for the introduction of foreign genes conferring commercially interesting characteristics, specifically disease resistance (Bao *et al.*, 2012).

This work focuses on the *Agrobacterium* mediated genetic transformation of four *Rosa hybrida* cultivars, “Tineke”, “Only love”, “Glad tidings” and “Hflor”, with the antifungal polypetide BLAD, (“Banda de *Lupinus albus* Doce”).

1, Genetic transformation of *Rosa x hybrida*: overview

Rose (*Rosa hybrida*) is a member of the Rosaceae family and represents an important commercial flower crop grown throughout the world. *Rosa hybrida* is not a species in the botanical sense, but is a description used for most cultivated rose cultivars of the Hybrid tea or Floribunda types or classes. These cultivated roses have been developed over centuries through complex crosses involving a number of species of the genus *Rosa*. Therefore Rosa species are very variable and hybridize freely, making species delimitation difficult. The chromosome number of rose varies from 2n=2x=14 to 2n=8x=56, with most species being diploid or tetraploid, commercial rose cultivars (*Rosa x hybrida*) tend to be either triploid or tetraploid. The difficulties in taxonomic classification of this species have resulted in reports of anywhere from 100 to 250 *Rosa* species, with innumerable cultivars (Office of the Gene and Technology Regulator, 2009). The traditional breeding goals of new rose varieties include improvements in color, fragrance, flower form, flowering habit, vase life of the cut flower, plant type, and obviously disease resistance, specifically against the two most common infections Powdery mildew caused by *Sphaerotheca pannosa* (Wallr.: Fr.) Lev. var. *rosae* and Black spot caused by *Diplocarpon rosea* (Li *et al*., 2003). The extent of results by traditional plant breeding has been limited, by the narrow gene pool, and by the incompatibility of varying ploidy level amongst species. Furthermore factors determining plant growth and development are commonly altered by sexual crossing and selection procedures, thereby giving rise to unpredictable progeny types (Bao *et al.*, 2012). These constraints on traditional breeding have early on sparked interest in the genetic modification of rose plant (Zuker *et al*., 1998). However, to this day, genetic transformation of rose is, extremely time-consuming with very low shoot-regeneration efficiencies, highly cultivar-specific and not easily reproducible. Therefore an efficient, reproducible, cultivar-independent transformation system is critical to generate the requisite number of elite transgenic rose lines (Tanaka *et al*., 2005).

1.1. Somatic embryogenesis of the Rose plant

The first step in developing transgenic plants is to have a procedure to regenerate an entire plant from individual transformed cells. Somatic cells of plant tissues have the capacity to undergo cellular dedifferentiation into a mass of unorganized cells, or callus, as well as the ability to generate differentiated cells, or somatic embryos, similar to zygotic embryos (Zimmerman, 1993). These somatic embryos have the competence to differentiate and regenerate into whole plants, and it is this ability, known as totipotency, to produce morphologically and developmentally normal organs, from somatic plant cells that present an intriguing and unique phenomenon in plants. In recent years, this observed phenomenon has become critical for successful asexual propagation of plants. Moreover, it serves as a limiting step in the area of transgenic plant development (Korban, 2006). Somatic embryogenesis may be induced via a direct or indirect pathway. For direct somatic embryogenesis, or primary, embryos develop directly on the surface of organized tissue. Alternatively, indirect somatic embryogenesis, or secondary, may occur via an intermediate step involving callus formation or a cell suspension culture (Bao *et al.*, 2012). In a typical tissue culture protocol for embryo formation, plant cells or organs are cultured under sterile conditions, on defined nutrient media, supplemented with specific concentrations of different plant growth regulators. Induction of in vitro embryogenesis from

somatic plant tissues occurs in response to high concentrations of auxin added to the culture media under controlled environmental conditions. In the presence of the auxin analog in the medium, plant cells form pre-embryogenic masses (PEMs) and upon removal of auxin from the culture medium, these PEMs then undergo differentiation from the globular stage to the heart/torpedo stage, and into plantlets (Li *et al*., 2002). Successful development of regeneration systems for a number of rose species and cultivars has already been reported. Embryogenic callus have been initiated from in vitro-derived leaf or stem segments of *Rosa hybrida* cv. Carl Red and *R. canina* (Visessuwan *et al*., 1997), *R. hybrida* cv. Carefree Beauty, and *R. chinensis* minima cv. Baby Katie (Hsia and Korban, 1996). Embryogenic callus has also been induced in leaves of *R. hybrida* cvs. Domingo and Vicky Brown (DeWit *et al*., 1990), petioles and roots of Tineke (Kim *et al*., 2004), *R. hybrida* cvs. Trumpeter and Glad Tidings (Marchant *et al*., 1996) and R. hybrida cv. Arizona, root explants of both *R. hybrida* cv.Moneyway (van der Salm *et al*., 1996) and *R. Heritage* × Alista Stella Gray (Sarasan *et al*., 2001), petals of R. hybrida cv. Arizona, and immature seeds of *R. rugosa* (Kunitake *et al*., 1993). This has also been achieved using immature leaf or stem segments of *R. hybrida* cv. Landora (Rout *et al*., 1991), in vivo mature leaves of *R. hybrida* cv. Soraya (Kintzios *et al*., 1999), anther filaments of *R. hybrida* cv. Royalty (Noriega and Söndahl 1991), as well as anthers, petals, receptacles, and leaves of *R. hybrida* cv. Meirutal (Arene *et al*., 1993). Secondary somatic embryogenesis, however, was observed in only a few of rose species and cultivars, namely: *Rosa hybrida* cvs. Carefree Beauty (Hsia and Korban, 1996; Li *et al*., 2002), *Rosa hybrida* L. cv. Landora (Rout *et al*., 1991) and *Rosa chinensis* cv. Old Blush (Vergne *et al*., 2010).The wide range of explants and experimental approaches that have been employed with different rose species and cultivars strongly suggest that it is difficult to develop a universal genotype-independent method for the production of embryogenic callus in rose, and optimum conditions must be defined for each cultivar and explant tissue type (Korban, 2006).

1.2. *Agrobacterium tumefaciens* transformation of Rose for improved fungal resistance

Plant transformation by the soil plant pathogenic bacterium, *Agrobacterium tumefaciens*, has become one of widely used methods for the introduction of foreign genes into plant cells, specifically rose. The bacteria *A. tumefaciens* naturally infects the wound sites in dicotyledonous plants causing crown gall disease, characterized by the formation of tumors at plant wound sites following infection by the bacteria (De La Riva *et al.,* 1998)**.** During the infection process, the *Agrobacterium* transfers a defined DNA fragment (the T-DNA, transfected DNA), carried by a large-size Ti (tumor-inducing) plasmid to the plant cells (Zuker *et al.*, 1998). A typical T-DNA vector plasmid contains five distinct regions: promoter, multiple cloning site, plant resistance (or marker gene), bacteria resistance, and two origins of replication, for *Agrobacterium* and *E.coli*. Recently developed *Agrobacterium* plant expression vectors typically contain a high-copy-number origin of replication for *E. coli*, a combination of either two expression promoters, the cauliflower mosaic virus (CaMV) 35S promoter , or the *Agrobacterium tumefaciens* nopaline synthase (*nos*) promoter (Horstmann *et al*., 2004), and a multitude of selectable marker genes either, for plant or bacteria selection. The most common plant selection markers are:

nptII (neomycin phosphotransferase II), hpt (hygromycin phosphotransferase), als (acetolactate synthase) and bar (phosphinothricin-N- acetyltransferase) (“Basta”) (Weigel and Glazebrook, 2006a).

In general plant genetic transformation for increase fungal resistance follows five common paths (Punja, 2001):

i) The expression of gene products that are directly toxic to pathogens or that reduce their growth. These include pathogenesis-related proteins (PR proteins) (chitinases, glucanases),), antimicrobial peptides (thionins, defensins, lectin), ribosome inactivating proteins (RIP), and phytoalexins.

ii) The expression of gene products that destroy or neutralize a component of the pathogen arsenal such as polygalacturonase, oxalic acid, and lipase.

iii) The expression of gene products that can potentially enhance the structural defenses in the plant. These include elevated levels of peroxidase and lignin.

iv) The expression of gene products releasing signals that can regulate plant defenses. This includes the production of specific elicitors, hydrogen peroxide (H_2O_2), salicylic acid (SA), and ethylene (C_2H_4).

v) The expression of resistance gene (R) products involved in the hypersensitive response (HR) and in in-teractions with avirulence (Avr) factors.

In the years following the successful establishment of rose plant *Agrobacterium tumefaciens* transformation protocols, only three works have been published pertaining rose’s genetic improvement for fungal resistance. In all these works, the main aim has been the expression of foreign antifungal protein or peptides, mainly PR proteins such as chitinases and glucanases. The first recorded rose transformation with an antifungal protein was by Marchant and colleagues *et al***.** (1998) that successfully introduced a chitinase gene into *R. hybrida* cv. Glad Tiding, and found that expression of the chitinase transgene reduced the severity of black spot (*Diplocarpon rosae* Wolf.) development by 13–43%. Subsequently, Garden rose cultivars Heckenzauber and Pariser Charme were transformed with genes encoding, Class II chitinase, a Class II ß- 1,3-glucanase and a Type I ribosome inhibiting protein, however only the secretion of the ribosome inhibiting protein into the extracellular space reduced the susceptibility against blackspot to 60 % on average (Dohm *et al*., 2001). Finally an antimicrobial protein, Ace-AMP1, was introduced into *Rosa hybrida* cv. Carefree Beauty via Agrobacterium-mediated transformation, and conferred enhanced resistance to powdery mildew *Sphaerotheca pannosa* (Wallr.: Fr.) Lev. var. rosae (Li *et al*., 2003). Although limited, these works clearly state the potential for the transformation of rose plant with antifungal proteins towards an increase fungal resistance.

2. Antifungal polypeptide BLAD a target for plant transformation

White lupines *(Lupinus albus* L.), as the other *Lupinus* species, possess in their seeds three major storage proteins: the globulins α-conglutin, β-conglutin and γ-conglutin. β-Conglutin is the main *Lupinus* globulin, and a lupine member of the vicilin-like or 7S family of storage proteins (Melo *et al*., 1994). The transcript encoding its precursor, pro-β-conglutin, a 64 kDa polypeptide, is specifically translated during the period corresponding to seed formation and development. Before the storage in the dry seeds takes place, pro-β-conglutin is subjected to very intense processing in a route which ends up with many tens, possibly hundreds of slightly different β-conglutin sub-units. Since native, mature β-conglutin is a trimer, the protein in dry *Lupinus* seeds exhibits a very high degree of micro–heterogeneity (Monteiro *et al*., 2010). In a subsequent cycle of growth, the seeds germinate during approximately one day and seedling growth ensues. Between days 3 and 5 after the onset of germination, β-conglutin suffers a dramatic change in its structure and concentration, involving the appearance of a new set of polypeptides, including a higher molecular mass group, whose concentration steadily declines until complete disappearance after 11 to 12 days, and a lighter molecular mass group, whose concentration surprisingly increases from day 5 to day 11 to 12. Particularly evident is the accumulation of a 20 kDa polypeptide, which is maintained in the cotyledons in high amounts between days 4 and 14 after the onset of germination, termed BLAD ("Banda lupinus albus doce") (Monteiro *et al*., 2010). BLAD is a very promising fungicide, and an excellent target protein for plant transformation, due its many unique properties: (i) it is a naturally-occurring food and feed item exhibiting a non-toxic mode of action; (ii) it exhibits a very effective fungicidal activity towards many fungal species tested; (iii) its unique mode of action reduces the risk of fungal resistance; (v) its small size (20 kDa) augments probability of successful plant transformation; (vi) it does not confer any particular colour, odour or flavour to the treated plants.

Material and Methods

This work was subdivided in two distinct stages. All the molecular biology work from BLAD coding gene amplification, plasmid construction to *Agrobacterium tumefaciens* strain transformation was performed at Instituto Superior de Agronomia (ISA). From this point onward, calli induction, pre-embriogenic calli transformation, plant regeneration and analysis of positive transformants, was performed at the Laboratory of Plant Breeding, Plant Research International Wageningen University, Holand, and supervised by MsC Herma Koehorst-van Putten. All tissue culture work had previously been optimized and the work published by the same research group (Condliffe *et al*., 2003).

1. Plasmid constructs

Three different plasmids were used in this work, all from the pMDC gateway vector series and containing the same resistance markers, namely kanamycin for *E.coli* cloning and Hygromicin for plant selection (Curtis and Grossniklaus, 2003)**.** Two of these plasmids were gently provided by Herma Koehorst-van Putten and the one containing the BLAD gene was constructed by us. These working plasmid constructs were termed accordingly (see Annex II.3. and III.1, for BLAD gene sequence and plasmid maps):

i) pEXPBLAD173- originally the plant expression pMDC32, BLAD cDNA encoding the 173 amino acid BLAD polypeptide was inserted (full BLAD gene, NCBI database accession number DQ142920.1), and driven by the strong plant constitutive promoter 2 x 35S CamV (Cauliflower mosaic virus);

ii) pMDC107- plasmid belonging from the pMDC gateway vector series containing the GUS (β – glucuronidase) reporter gene driven by the strong plant constitutive promoter 2 x 35S CamV;

iii) pEXPLUC- originally the plant expression plasmid pMDC 32, but containing the reporter gene LUC (luciferase) and driven by the strong plant constitutive promoter 2 x 35S CamV.

1.1. Assembly of the pEXPBLAD173 plasmid

1.1.1. PCR and cDNA purification

Lupinus albus cDNA, stored at -80 ºC, was PCR ("Polymerase Chain Reaction") amplified using specific primers containing appropriate 15 bp flanking regions termed att B essential for the subsequent cloning steps (Annex II.2). The PCR reaction was performed in a Bio-Rad device (C1000PCR) using the taq (*Thermus aquaticus*) pfx (proofreading) polimerase (Inivitrogen), with the following PCR program: 94 ºC, 2min, followed by 30 cycles of 94°C for 15 s, 55°C for 30 s and 68°C for 1 min. The buffers used for the PCR reaction were in following working concentration: 1 X Pfx Amplification Buffer, 0.3 mM dNTP mixture, 1 mM $MgSO_4$, 1 µM Primer mix, 1 uL of *Lupin* sp. cDNA, 0.4 µL of taq Pfx DNA Polymerase, all adjusted to 50 µL of autoclaved Mili-Q water. PCR products were separated in horizontal agarose gel electrophoresis containing 5 µg/mL of gel red dye (Biotium),

to stain nucleic acids. When finished the electrophoresis was visualized in a UV transiluminator (Geldoc, Bio-Rad) using the "Quantity One" software (Bio-rad). After visualizing the gel, the selected DNA bands were cut and purified using the "Wizard purification" kit (Promega). Following purification the DNA samples were quantified in a Synergy HT spectrophotometer (Biotech), and the results registered for further use.

1.1.2. Cloning BLAD gene into the pMDC 32 plasmid

This cloning reaction was performed using a specific cloning technology developed by Invitrogen Company, termed Gateway®, using the "Gateway® Technology cloning kit" (Invitrogen). The Gateway technology is based on the bacteriophage lambda site-specific recombination system which facilitates the integration of lambda into the *E. coli* chromosome and the switch between the lytic and lysogenic pathways. Simply, any target gene added with a specific 15 bp att B sequence, for example, will, in the presence of BP clonase enzyme (Invitrogen), recombine with an attP sequence, via the lysogenic pathway, generating attL sites. Furthermore, any attL sequences, in the presence of LR clonase enzyme (Invitrogen), will recombine with any attR sequences generating attB sequences, via the lytic pathway. This simple DNA recombination allows an easy cloning of our BLAD gene containing attB sequences into the entry plasmid pDONR221 (Invitrogen) containing attP sites, generating the entry plasmid pENTRBLAD173 with attL sites flanking the BLAD gene. This plasmid can then be cloned into the expression plasmid pMDC32 containing corresponding attR sites, creating the plant expression plasmid pEXPBLAD173. To begin this procedure 300 ng each, of BLAD cDNA flanked by attB regions from the previous step, and pDONR221 plasmid were placed in an eppendorf tube. To this DNA mixture 4 µL of BP clonase enzyme was added followed by 4 µL of BP clonase buffer (Invitrogen), TE buffer (10 mM Tris-HCl, 1 mM EDTA (Ethylene diamine tetracetic Acid), pH 8.0) was added to a final volume of 16 µL, and the reaction was incubated for 1 h at 25 ºC. Afterwards the reaction was stopped by adding 2 µL of Proteinase K solution (Invitrogen), and left incubating 10 min at 37 ºC. When the BP reaction was completed, 1 µL of reaction was removed and put into one vial of "Oneshot TOP10 Competent Cells TOPO" (Invitrogen). This vial was incubated on ice for 30 minutes. After this time the cells were heat shocked for 30 seconds at 42 ºC in a water bath. Immediately after, the vial was transferred back to ice and left to rest for 2 minutes. When the two minutes passed 250 µL of Luria-Bertoni medium (LB) (Annex II.1) was added to the mixture and was incubated in a orbital horizontal shaker (ES 20-60, Biosan) for 1 hour at 200 r.p.m. at 37 ºC. Following the incubation step, three volumes, 50, 100 and 200 µL, were spread on a pre-warmed Luria-Bertoni solid medium (LBA) (Annex II.1) petri dish supplemented with, 50 µg/µL of kanamycin (Sigma-Aldrich), and incubated overnight at 37 ºC. On the following day, five isolated colonies were selected and tested for the correct insertion of the gene through a digestion reaction with the specific restriction enzymes Xba I and Hind III, followed by a PCR reaction with the gene's specific primers, as described previously. When proven positive and correctly inserted, a single colony containing the pENTRBLAD173 plasmid was grown in 50 mL of LB medium with 50 µg/mL of kanamycin and DNA was purified using the "Jetstar Plasmid Maxiprep Kit" (Genomed). The purified plasmid DNA was quantified using the Synergy HT spectrophotometer (Biotech), and the results registered for further use. In order to start the lytic reaction 300 ng of both

these plasmids, pENTRBLAD173 and pMDC32, were placed in an eppendorf tube. To this tube 4 µL of LR clonase enzyme (Invitrogen) was added followed by 4 µL of LR clonase buffer, TE buffer was added to a final volume of 16 µL. This mixture was incubated at 25 ºC for 1 h, and the reaction was stopped by adding 2 µL of Proteinase K solution for 10 min at 37 ºC. When the LR reaction was completed, 1 µL of reaction was removed and put into one vial of "DB3.1" (Invitrogen), and the cells transformed as described previously. On the day following *E. coli* transformation, five isolated colonies were selected and tested for the correct insertion of the gene through a digestion reaction with the specific restriction enzymes Xba I and Hind III, followed by a PCR reaction with the genes specific primers, as described previously. A selected positive colony containing the plasmid pEXPBLAD173 was grown in 50 mL of LB medium with 50 µg/mL of kanamycin and the plasmid DNA was purified using the "Jetstar Plasmid Midiprep Kit" (Genomed). The purified plasmid DNA was quantified using the Synergy HT spectrophotometer (Biotech), and sent to be sequenced by the Stavida Company, according with its internal specifications. The sequence results were analyzed using the MULTIALIGN software (Corpet, 1988), by comparison with BLAD published nucleotide sequence (NCBI database accession number DQ142920.1) (Annex II.2). The pEXPBLAD173 plasmid obtained was then ready to be transformed into the selected LBA4404 *Agrobacterium tumefaciens* strains.

2. Agrobacterium transformation

In this work we chose the *A. tumefaciens* strain LBA4404 (NCCB nº 2760) carrying the chromosomal antibiotic resistance marker for rifampicin, and the binary plasmid pAL4404 with the antibiotic resistance marker for streptomycin (Hellens *et al*., 2000). Competent cells of LBA4404 were thawed on ice and 250 µL of this suspension was added to a solution containing 5 µg of the different working plasmids: pEXPBLAD173, pMDC107 and pEXPLUC. This mixture was kept on ice for 5 minutes and an additional 5 min at 37 ºC. After this heat shock treatment 1 mL of LB medium was added, and the inoculum was allowed to grow at ambient temperature for 4 hours with 250 r.p.m. agitation. The *A. tumefaciens* cells were collected by a mild centrifugation and spread on solid LB medium containing 50 µg/mL kanamycin, 50 µg /mL streptomycin and 10 µg /mL rifampicin, for 48 h at 28 ºC. After two days 10 single colonies of transformed *A. tumefaciens* were selected and restreaked on fresh solid LB medium with antibiotics, and placed in 10 mL cultures of liquid LB with the aforementioned antibiotics and allowed to grow overnight at 28 ºC with shaking. The cultures in the solid medium were used as stock colonies, and the entire plate sealed with parafilm and placed at 4 ºC. The plasmid DNA from the liquid cultures was extracted and purified using the "Jetstar Plasmid Maxiprep Kit" (Genomed), and quantified using the Synergy HT spectrophotometer (Biotech), and the results registered for further use (Weigel and Glazebrook, 2006b). The correct insertion of the BLAD gene was tested for each colony through a digestion reaction with the specific restriction enzymes Xba I and Hind III, followed by a PCR reaction with the gene's specific primers, as described previously.

3. Tissue culture

3.1.Plant material and establishment of in vitro micropropagated plantlets

Micropropagated plantlets from *Rosa hybrida* cvs. Only Love, Tineke and Glad Tidings were gently provided by Herma Koehorst-vanPutten. Potted plants from the *Rosa hybrida* cv. HFlor were gently provided by Professor António Monteiro (ISA). After the leaves and shoot tips had been removed, selected robust stems were washed under running water for 30 min and cut into single nodal sections of approximately 4 cm in length. Working on a horizontal laminar air flow chamber, and being careful to maintain sterile conditions, the 4 cm nodal sections were immersed 70% (v/v) ethanol (Merck) for 30 seconds followed by a 10 min soak in 2.5% (v/v) sodium hypochlorite solution with a few drops of Tween-20 (Sigma-Aldrich) as wetting agent. The samples were then rinsed three times with sterile mili-Q water and the explants, ~4 cm long, were placed for 14 days in Van der Salms medium (VS) (Duchefa) supplemented with the hormones 4 µM BAP (N-Benzyl-9-(2-tetrahydropyranyl)adenine) (Duchefa) and 9 µM Zeatin (Duchefa) and solidified with 2.5 g gelrite (Duchefa) separately. VS medium is a standard Murashige and Skoog (MS) medium with the substitution of the FeNaEDTA salt by FeEDDHA (Van der Salm *et al.,* 1994). The cultures were kept in culture room at. After sprouting, the shoots were maintained through a 4-weekly subculture regime on the same VS medium. VS medium was prepared by adjusting its pH to pH 5.8–6.0 with 1 M NaOH and autoclaved.

3.2. Rooting of micropropagated plantlets

Rooting was stimulated by placing plantlets from each cultivar for 1 month at 23±2ºC under 16 hour photoperiod of 1500 lux light intensity, in solid VS medium non-supplemented for the *Rosa hybrida* cvs. Only Love, Tineke and Glad Tidings or supplemented with 1.5 mg/L IBA (Indole-3-butyric acid) (Duchefa) for the cv. HFlor (Condliffe *et al*., 2003).

3.2.1. Callus induction, somatic embryogenesis induction

Different explants were used to initiate callus culture, namely, leafs, petioles and roots. For callus induction fifty of these explants were placed in callus induction solid media (CI), consisting of MS basal medium containing 4 mg/L of 2,4-D (2,4-Dichlorophenoxyacetic acid) (Duchefa), plus 30 g/L sucrose (Merck) and solidified using 2.5 g/L “Seakem” agar (Duchefa). The cultures were incubated in darkness for 8 weeks, at 23 ºC, with a subculture step onto fresh medium at the end of the first 4 weeks. The calli were maintained in this medium for at least six months, and frequently checked for the appearance of a different type of cell tissue, smaller, more differentiated, termed pre-embryogenic calli. Each time this differentiated cell mass of pre-embryogenic calli appeared they were carefully isolated and placed in new CI medium and given a number according with the cultivar and time of appearance (*e.g.* OL 1 for the first isolated pre-embryogenic cell line one from rose cultivar Only Love). Constituting a new cell line for genetic transformation. These differentiated lines of pre-embryogenic calli were kept at 23 ºC in the dark and subcultured at 4 week intervals, and the excess mass was discarded (Condliffe *et al*., 2003).

4. Pre-embryogenic calli transformation

Transformation experiments were performed using homogeneous pre-embryogenic tissues from 'Glad Tidings', 'Only Love' “Tineke” and HFlor cultivars, and using transformed colonies containing the pEXPBLAD173 plasmid. As a control for the transformation and selection methodologies, each cultivar was also transformed separately with the plasmids pEXPLUC and pMDC107, containing the reporter genes coding for luciferase and Gus. For the actual transformation, overnight liquid cultures of *A. tumefaciens* colonies, containing the different working plasmids, were diluted in order to have OD (optical density) values in between 0.8 and 1. After the OD adjustment liquid cultures of *A. tumefaciens* were placed in 50 mL of CI medium, along with 8 g of pre-embryogenic tissue, and co-cultivated for 2 hours at 28 ºC with a constant 250 r.p.m. agitation. Following co-cultivation, embryogenic material was passed through a 2 mm filter, collected and transferred to 20 ml of liquid CI medium. After 2 days cultivation in the dark at 28 ºC on a horizontal shaker (250 r.p.m.), cultures were again passed through a 2 mm mesh filter and transferred to new CI medium containing 50 µg/mL kanamycin, 100 µg/mL timentin, 200 µg/mL cefotaxim and µg/mL hygromicin. Cultures were refreshed with new CI medium containing antibiotics at 4 weeks intervals (Condliffe *et al*., 2003).

4.1. Germination of transformed pre-embriogenic calli and plantlet development

Putatively transformed pre-embryogenic calli tissues were assayed through PCR after 28 days following co-cultivation. After cultivation in CI media with antibiotics for a period of 12 weeks at 25ºC, in the dark, calli showing signs of proliferation and embryogenesis were transferred onto maturation/germination medium (MGM) containing the aforementioned antibiotic selection. Germinating embryos were transferred to germination medium (GM), a modified MS based medium containing 30 g/L glucose, 1.0 mg/L 6-benzylaminopurine (BAP) and 0.8% (w/v) agar with no selection, at 25 ºC in the dark. Germinating embryos had their medium refreshed every 4 weeks and the promising ones were placed in GM medium at 25 ºC with increasing light exposure, for stimulating plantlet development. Each new plantlet was placed in the previously described VS medium supplemented with the hormones BAP and Zeatin (Condliffe *et al*., 2003).

5. Analysis of putatively transformed plant tissue

5.1. DNA extraction and PCR analysis of putatively transformed plant tissue

Putatively transformed cell tissue was only PCR tested after the germination of embryos and the antibiotic selection had been removed. Analysis of plantlets took place only when enough plant material could be removed for DNA extraction without compromising plantlet regeneration. The different tissues (~20 mg each) were grinded with a mortar and pestle under liquid nitrogen to a fine powder, and the purification protocol for this extraction kit was applied. Following DNA purification, all samples were quantified in a Synergy HT spectrophotometer (Biotech), and the results registered for further use. Isolated DNA samples were then PCR tested using the att specific BLAD primers and the recombinant taq DNA polymerase (Invitrogen) with the following PCR program: 94ºC, 2 min, followed by 30 cycles of 94°C for 15 s, 55°C for 30 s and 68°C for 1 min. The buffers used for the PCR

reaction were in following working concentration: 1 X taq Amplification Buffer, 0.3 mM dNTP mixture, 1 mM MgSO4, 1 µM Primer mix, 1 µL of roses genomic DNA, 0.4 µL of recombinant taq DNA Polymerase, all adjusted to 50 µL of autoclaved Mili-Q water. PCR products were separated in horizontal agarose gel electrophoresis as previously described and visualized in a UV transiluminator.

6. Reporter genes

6.1. GUS expression

The GUS assay is based on the ability of the ß-glucuronidase protein to react with 5-bromo-4-chloro-3-indolyl glucuronide (X-Gluc) subtract, producing a highly colored blue dye, allowing for a rapid determination of transformation efficiency. For a simple determination of transformation efficiency a reasonable portion of a day old, transformed pre-embryogenic calli with the plasmid pMDC107, was immersed in 500 µL of X-GLUC reagent (Duchefa) in 24 well plates and incubated for one hour at 37 ºC. After staining the clear blue color appears indicative of transformed pre-embryogenic calli tissue.

6.2. Luciferase expression

The Luciferase assay is based on the ability of the Luciferase expressed protein to produce light by converting the chemical energy of luciferin oxidation, forming as a byproduct the molecule oxyluciferin. For a simple determination of transformation efficiency a reasonable portion of a day old transformed pre-embryogenic calli with the plasmid pEXPLUC, was immersed in 500 µL of Luciferin reagent (Promega) in 24 well plates and immediately visualized in a luminometer at 560 nm.

Results

1. Construction of the expression plasmid

The first step on this project was the selection of the appropriate expression plasmid and *A. tumefaciens* strain. Looking at previous published works on plant transformation, we selected the improved binary Ti vector pMDC32 carrying the strong constitutive double promoter CaMV (cauliflower mosaic virus) 35S and the compatible LBA4404 *A. tumefaciens* strain (Firoozabady *et al.*, 1994; Curtis and Grossniklaus, 2003; Kim *et al*., 2004). This family of plasmids, offer two main advantages. First a wide range of binary Ti vectors, with different versions of the same basic design, permitting the placement of our target gene into a specific vector plasmid (*e.g.* pMDC32) and the purchase of a similar plasmid containing a chosen reporter gene (*e.g.* pMDC107), minimizing plasmid backbone interference. Secondly, this family of plasmids, has been integrated with the Gateway cloning technology, allowing for any gene to be cloned into a donor vector (and thus flanked by attP sites) and be conservatively transferred in one step, in the correct orientation, at high efficiency, into any vector of this series (Grossniklaus, 2003; Kim *et al*., 2004). Having chosen the expression plasmid, specific BLAD amplification primers containing the Gateway specific attB sequences were design for the cloning reaction. Following the Gateway protocol, the now amplified, attB containing BLAD gene, had to be introduced into the pDONR221 plasmid, for posterior insertion in the expression plasmid. For determining the correct cloning of the BLAD gene, we carried out a PCR amplification reaction of five isolated *E. coli* colonies, with the same protocol and primers used for initial cDNA BLAD amplification. Bear in mind that the BLAD gene contains 516 bp plus 30 bp from the att sites.

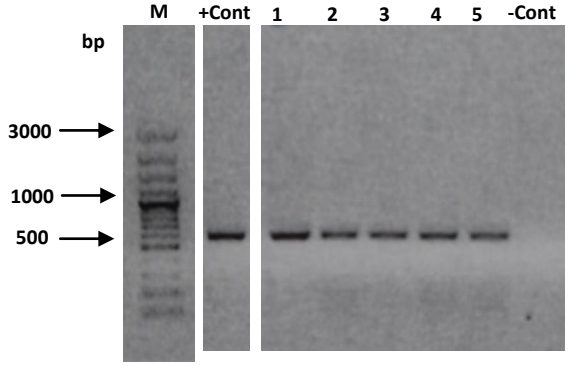


Figure 7: Agarose gel electrophoresis (1 % agarose, 5 µg/mL gel red) of the PCR amplification of the BLAD gene (~530 bp) present in the pENTRBLAD173 plasmid, using BLAD att primers and five isolated *E. coli* colonies that grew on selection LB media (50 µg/mL Kanamycin). Lane 1- 10 µL of the PCR amplification of *E.coli* colony 1. Lane 2- 10 µL of the PCR amplification of *E.coli* colony 2. Lane 3- 10 µL of the PCR amplification of *E.coli* colony 3. Lane 4- 10 µL of the PCR amplification of *E.coli* colony 4. Lane 5- 10 µL of the PCR amplification of *E.coli* colony 5. Lane +Cont - positive control, PCR amplification of BLAD cDNA. Lane -Cont – negative control, PCR amplification with no DNA, containing only water. Lane M – marker 100 bp plus (Novagen).

Examining the results obtained for the PCR reaction (Figure 7), it is possible to conclude that the BLAD gene was correctly cloned into all of the tested colonies, with a clear band between the 500 bp and 600 bp marker, therefore, the entry plasmid pENTRBLAD713 was successfully created. Considering this, we proceeded with the second cloning step, this time into the pMDC32 expression plasmid, to create the plant expression plasmid pEXPBLAD173. To test the correct insertion of our

target gene into the expression plasmid we carried out two types of tests: i) PCR amplification of five isolated *E. coli* colonies, with the same protocol and primers used for initial cDNA BLAD amplification; ii) double digestion (Sac I and Hind III) of the purified plasmid DNA from one of the PCR positive colonies.

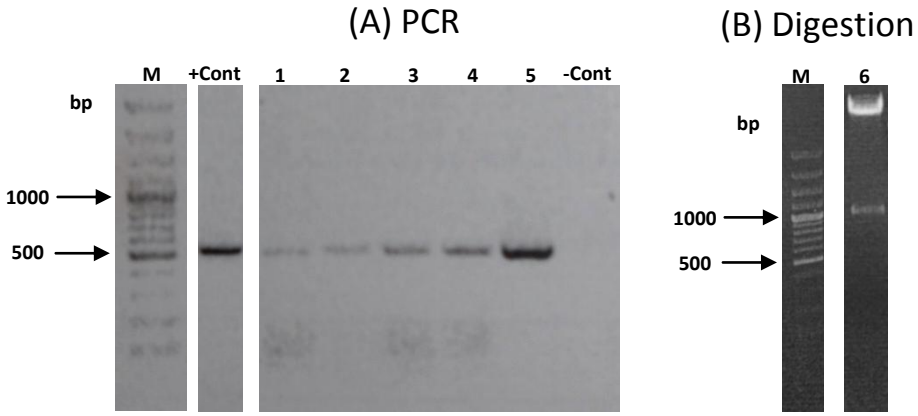


Figure 8: Agarose gel electrophoresis (1 % agarose, 5 µg/mL gel red) of the: (A) PCR amplification of the BLAD gene (~530 bp) present in pEXPBLAD173 plasmid, with the BLAD att primers, of five isolated *E. coli* colonies that grew in selection LB media (50 µg/mL Kanamycin); (B) digestion of the PCR positive *E.coli* colony 4, with Sac I and Hind III enzymes. Lane 1- 10 µL of the PCR amplification of *E. coli* colony 1. Lane 2- 10 µL of the PCR amplification of *E. coli* colony 2. Lane 3- 10 µL of the PCR amplification of *E. coli* colony 3. Lane 4- 10 µL of the PCR amplification of *E. coli* colony 4. Lane 5- 10 µL of the PCR amplification of *E. coli* colony 5. Lane 6- 10 µL of the double digestion with Sac I and Hind III enzymes of purified colony 4 DNA. Lane +Cont - positive control, PCR amplification of BLAD cDNA. Lane -Cont – negative control, PCR amplification with no DNA, containing only water. Lane M – marker 100 bp plus (Novagen).

Observing both these results (Figure 8, A and B), we can say that all of the five colonies were properly transformed with the BLAD gene. In addition, the newly formed pEXPBLAD173 plasmid had been correctly constructed since after the double digestion on colony 4 with both Hind III and Sac I, the pEXPBLAD173 plasmid generate one band with a size of around 1330 bp, as expected.. As a result, this new pEXPBLAD173 plasmid was purified and quantified for *A. tumefaciens* transformation.

Upon confirmation on the plant expression plasmid correct construction of the, we set out to place it into the appropriate *A. tumefaciens* strain. Following already described methodologies, the LBA4404 strain was transformed with the pEXPBLAD173 plasmid. The analysis for the correct insertion was carried out as before: i) PCR amplification of five isolated *A. tumefaciens* colonies, with the same protocol and primers used for initial cDNA BLAD amplification; ii) double digestion (Sac I and Hind III) of purified plasmid DNA from one of the PCR positive colonies.

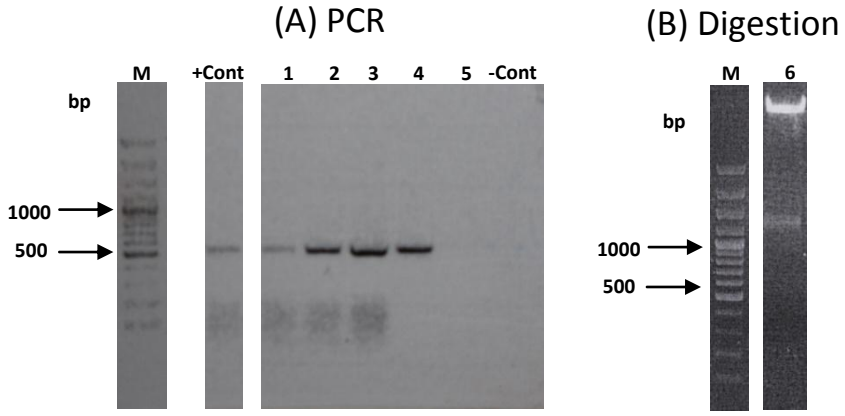


Figure 9: Agarose gel electrophoresis (1 % agarose, 5 µg/mL gel red) of the: (A) PCR amplification of the pEXPBLAD173 plasmid, with BLAD att primers, of five isolated *A. tumefaciens* colonies that grew in selection LB media (50 µg/mL Kanamycin; (B) digestion of the purified DNA from the PCR positive *A.tumefaciens* colony 3 with Sac I and Hind III enzymes. Lane 1- 10 µL of the PCR amplification of *A. tumefaciens* colony 1. Lane 2- 10 µL of the PCR amplification of *A. tumefaciens* colony 2. Lane 3- 10 µL of the PCR amplification of *A. tumefaciens* colony 3. Lane 4- 10 µL of the PCR amplification of *A. tumefaciens* colony 4. Lane 5- 10 µL of the PCR amplification of *A. tumefaciens* colony 5. Lane 6- 10 µL of the double digestion with Sac I and Hind III enzymes of purified colony 3 DNA. Lane +Cont - positive control, PCR amplification of BLAD cDNA. Lane -Cont – negative control, PCR amplification with no DNA, containing only water. Lane M – marker 100 bp plus (Novagen).

Analyzing the PCR band patterns from colonies 1 through 4 (Figure 9 A) apparently these four colonies were successfully transformed with the expression plasmid pEXPBLAD173. Using the same methodology as before, a single colony was selected, colony 3, and its plasmid DNA was purified and submitted to a double digestion with Hind III and Sac I. The double digestion of colony 3 generated the expected band profile and the pEXPBLAD173 plasmid a band with the expected base pair size. As a result, this new pEXPBLAD173 plasmid was purified and sent to be sequenced in order to confirm the exact sequence match with the published BLAD gene sequence (Annex II.4).

2. Rooting, Callus induction and pre-embryogenic callus formation on the four rose cvs

2.1. Rooting

According with the established methodology in the Dutch laboratory for *A. tumefaciens* genetic transformation of rose plant (Condliffe *et al*., 2003), our work began by inducing roots from the four target rose cultivars. We placed 50 micropropagated plantlets from each cultivar in rooting medium (non-supplemented solid VS media). And after one month of growth the rooting efficiency was determined.

Rose cv.	Rooting percentage
Only love	87 %
TIneke	89 %
Glad tidings	96 %
HFlor	2 %

Table 4: Percentage of root formation on 50 micropropagated plantlets of the four rose cultivars (Only love, Tineke, Glad tidings and HFlor), after one month on rooting induction media.

The rooting efficiency was higher for the cv. Glad tidings (96%) followed by the rose cv's. Tineke (89 %) and Only love (87 %) (Table 4). The cultivar HFlor, under these induction conditions, proved to be

a more recalcitrant cultivar with only a small percentage of root formation (2 %). Typically to improve root induction on rose, auxins, such as IBA (Indole-3-butyric acid), at different concentrations and combinations, are added, (Pratap *et al.*, 2006). Therefore, to test whether the presence of IBA improves HFlor rooting induction, two different IBA concentrations were tested (1 and 1.5 mg/L) on 10 plantlets each of cv HFLor,.

IBA (mg/mL)	Rooting percentage
0	2 %
1	34 %
1,5	75 %

Table 5: Percentage of root callus formation on 10 micropropagated plantlets each, of cv HFlor, in the presence of different IBA concentrations, 0 mg/mL, 1 mg/mL and 1,5 mg/mL, after one month on the rooting induction media.

The addition of IBA to the rooting media resulted in a drastic increase of rooting percentage (Table 5). By adding 1.5 mg/mL of IBA the cultivar HFLor is able to produce a sufficient amount of calli material for the subsequent transformations steps, and therefore constitutes a valid addition to the rooting induction media for this cultivar. Looking at previous published works, IBA has been applied in-between 0.1 to a concentration of 2 mg/mL (Pratap *et al.*, 2006). Thus this level of IBA concentration is close to the highest registered level of IBA concentration for root induction.

2.2. Calli induction and Pre-embryogenic calli cell lines

After a successful rooting procedure, 30 isolated roots, from each of the tested rose cultivars were placed in CI media in the dark, to stimulate callus formation. After two months, the root callus induction efficiency was determined.

Rose cv.	Calli percentage
Only love	70 %
Tineke	65 %
Glad tidings	83 %
HFlor	23 %

Table 6: Percentage of root callus formation on 30 roots belonging to the four rose cultivars (Only love, Tineke, Glad tidings and HFlor), after two month on rooting induction media.

Similar to the rooting efficiency, calli induction was higher on the cultivar Glad tidings (83%) followed by cultivars Only love (70 %) and Tineke (65 %) (Table 6). Again, cultivar HFlor displayed the worst result, proving a difficult cultivar to induce calli. All the obtained calli were left in CI media, for at least six months. The media was refreshed, and when a gray mass of more differentiated calli appeared, the new cell mass was collected, labeled and placed in a new CI media. As an example, the picture below (Figure 11), portrays the difference between a typical root callus of Glad tidings cv., and the more differentiated mass of gray pre-embryogenic calli, for the same cultivar.

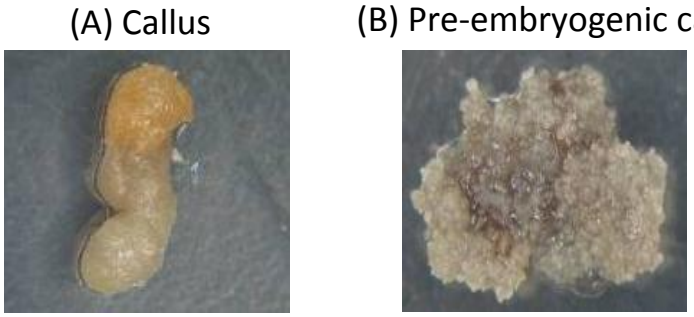


Figure 10: Image portraying the difference between two different type of calli of the Glad tidings rose cultivar: (A) mass of 8 week old callus in CI media; (B) mass of 24 week old pre-embryogenic calli in CI media.

After the six month period, every calli, resembling the one pictured in Figure 10 B, was separated and carefully labeled. Although no record was kept on the percentage of calli induction, the number of cell lines obtained was registered and is displayed below (Table 7).

Rose cv.	Nº PE cell lines
Only love	16
Tineke	8
Glad tidings	19
HFlor	3

Table 7: Number of pre-embryogenic (PE) cell lines obtained after a six month refreshing and selection period.

As in previous tissue culture steps, Glad tidings had the best results (19) with the highest number of pre-embryogenic cell lines, followed by the Only love (16). Considering the cvs Tineke (8) and HFlor (3), this method for pre-embryogenic calli induction is not well suited for these two types of genotypes. These disparate results between Rose genotypes are a common feature in Rose plant tissue culture. It is a well established fact, that production of embryogenic callus in rose must be optimized for each cultivar (Tanaka *et al*., 2005).

2.3. Reporter gene analysis

All of these selected lines of pre-embryogenic calli were then transformed with *A. tumefaciens,* carrying the expression plasmid pEXPBLAD173. As a control for the transformation efficiency, each cultivar was also co-cultivated with *A.tumefaciens* carrying the plasmids pMDC107 and pEXPLUC, containing the reporter genes Gus and Luc respectively. If following a transformation event no transient expression of the Gus and Luc genes was observed the entire transformation experiment was discarded. On the other hand, if some level of expression of the reporter genes was observed, as the example below (Figure 11), the experiment was maintained.

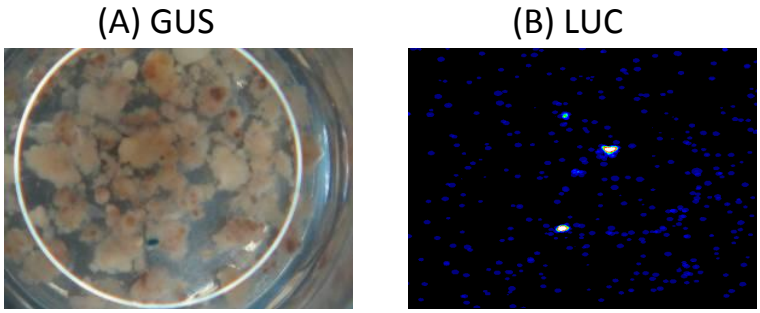


Figure 11: Image displaying a typical result for the transient expression of the reporter genes (A) GUS and (B) LUC in 24 hour old transformed cell tissue.

3. PCR analyses of the putatively transformed cell lines

Following a successful plant transformation, putatively transformed calli were kept in the dark for a period of 12 weeks with antibiotic selection. After this period of tissue selection, calli showing signs of proliferation and embryogenesis were transferred onto maturation/germination medium (MGM). At this stage, although we had transformed 46 cell lines belonging to the four rose genotypes, only ten cell lines from three cultivars survived the selection procedure. Specifically cell lines: GT 1, 12, 13 18 (Glad tidings), OL 3, 4, 8 14 19 (Only love) and cell line T 3 (Tineke). All the cell lines placed in MGM medium were PCR screened to test for the presence of the BLAD gene, and the efficiency of the selection process.

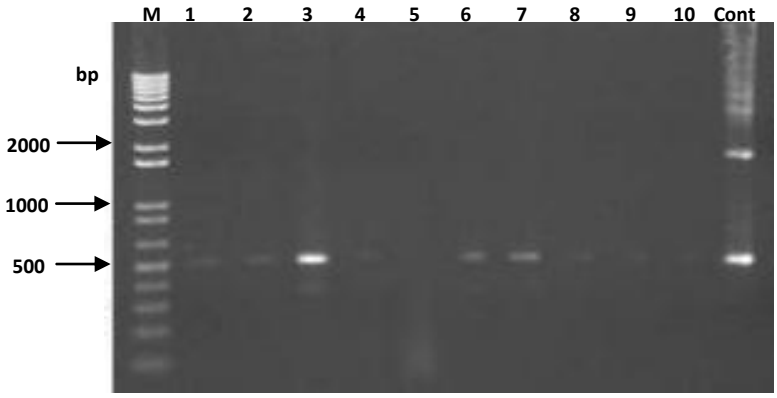


Figure 12: Agarose gel electrophoresis (1 % agarose, 5 µg/mL gel red) of the PCR amplification of the BLAD gene (~530 bp) using BLAD att primers and as template DNA extracted from the putatively transformed cell lines. Lane 1 – cell line 1 Glad tidings (GT 1); Lane 2 – cell line 12 Glad tidings (GT 12); Lane 3 – cell line 13 Glad tidings (GT 13); Lane 4 – cell line 18 Glad tidings (GT 18); Lane 5 – cell line 19 Glad tidings (GT 19); Lane 6 – cell line 3 Only love (OL 3); Lane 7 – cell line 4 (OL 4); Lane 8 – cell line 8 (OL 8); Lane 9 – cell line 14 (OL 14); Lane 10 – cell line 3 Tineke (T3); Lane M – marker 1 kb ladder (Invitrogen).

From this PCR screening result, we detected the presence of the BLAD gene in nine of the selected cell lines, although, in some cases only a faint presence (Figure 12). Only in the cell line GT 19 there was no amplification of the BLAD gene, constituting a non-transformed cell line. Therefore at this point of the work the selection methodology followed appeared to be adequate.

These nine PCR positive embryogenic cell lines were left in MGM medium for an adequate time followed by the application of the plant regeneration protocol. Upon plantlet development, each cell

line was subjected to a second PCR screening analysis. The genomic DNA from each cell line was extracted and PCR amplified with BLAD att primers to test for the presence of the BLAD gene in the plantlet genome.

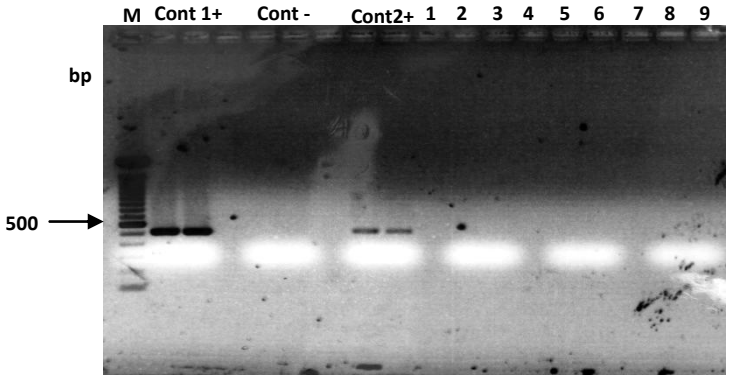


Figure 13: Agarose gel electrophoresis (1 % agarose, 5 µg/mL gel red) of the PCR amplification of the BLAD gene (~530 bp) using BLAD att primers and as template DNA extracted from plantlets belonging to the putatively transformed cell lines. Lane 1 – cell line 1 Glad tidings (GT 1); Lane 2 – cell line 12 Glad tidings (GT 12); Lane 3 – cell line 13 Glad tidings (GT 13); Lane 4 – cell line 18 Glad tidings (GT 18); Lane 5 – cell line 3 Only love (OL 3); Lane 6 – cell line 4 (OL 4); Lane 7 – cell line 8 (OL 8); Lane 8 – cell line 14 (OL 14); Lane 9 – cell line 3 Tineke (T3); Lane Cont1+- positive control, PCR amplification of pEXPBLAD173 plasmid. Lane- Cont(-) – negative control, PCR amplification with no DNA, containing only water. Lane Cont2 +- positive control, PCR amplification of BLAD cDNA; Lane M – marker 100 bp plus (Novagen).

This last PCR screening gave unequivocal results. Even though after antibiotic selection, nine out of ten cell lines appeared to be properly transformed. After plantlet development none of the cell lines were correctly transformed. Due to our incredulity, this PCR was repeated two times, with similar results. In retrospect, our selection media with antibiotic was not stringent enough. Possibly, there was some degree of contamination with the *A.tumefaciens* on the first PCR screening, resulting in false positives.

Conclusions

In this work we set out to use the *A. tumefaciens* bacteria to genetically transform four cultivars of Rose with the antifungal polypeptide BLAD. In the first part of the work, we constructed the plant expression plasmid pEXPBLAD173 containing the BLAD gene, driven by the strong constitutive promoter 2x35 CaMV, and placed it inside the LBA4404 *A. tumefaciens* strain. After this initial step, the second stage of the work was focused on all the tissue culture steps that a plant transformation protocol encompasses. Following the protocol developed in the Dutch laboratory, we began by inducing root formation in all of the tested cultivars, to serve as the plant explants for pre-embryogenic calli induction. The root induction efficiency was higher for the cv. Glad tidings (96%) followed by the rose cv's. Tineke (89 %), Only love (87 %) and HFlor (2 %). By adding 1.5 mg/mL of IBA to the induction media we were able to improve rooting efficiency of this cultivar, to values close to the other plant cultivars (72 %). Following the development for each plant genotype of an efficient rooting protocol, we set out to induce callus. Similar to the rooting efficiency calli induction was higher on the cultivar Glad tidings (83%) followed by cultivars Only love (70 %), Tineke (65 %) and HFlor (23%). From these newly formed calli we were able to isolate 46 independent lines of pre-embryogenic calli from each of the four cultivars. Each of these cell lines was transformed with LBA4404 *A.tumefaciens* containing the expression plasmid pEXPBLAD173. These transformed calli belonging to 46 cell lines were left in selection media and only those showing good signs of development were transferred to MGM media. Of the initial 46 cell lines only ten cell lines from three cultivars survived the selection procedure. All the ten cell lines on MGM media were immediately PCR screened to test for the presence of the BLAD gene, and in nine out of ten the presence of the BLAD gene was detected. These nine cell lines were allowed to develop into micropropagated plantlets. At this stage a new PCR screening was performed. Of the nine, putatively transformed plantlets, none showed any sign of BLAD gene presence. In all likelihood there was contamination with the *A.tumefaciens* used in the transformation, resulting in false positives.

Chapter IV

Optimization of the Heterologous Expression in *E.coli* of the antifungal polypeptide BLAD

Introduction

In the past century several bioactive proteins were identified and found suited for wider industrial usage, however, many of these, are large and highly complex molecules present in low amounts in its parent organism, making difficult to be mimetized through classic synthetic chemistry and costly, if not impossible, to isolate in working amounts from the parent organism (Palomares *et al.,* 2005). The development of molecular cloning technology in the early 1970s, for propagating targeted segments of DNA, created a revolution in the biological and biomedical sciences triggering the development of several new techniques (Brown, 2011)**.** One such technique relied in the introduction of an encoding gene into easily cultivable recipient cells from other species, so as to use them as biofactories to produce recombinant proteins, in a technique termed heterologous protein expression (Villaverde and Carrió, 2003). With this revolutionary technology many proteins that were very scarce and difficult to obtain could now be readily recombinantly produced at affordable prices (Palomares *et al.,* 2005). Today more than 200 recombinant proteins have been approved for commercial uses, and approximately 70 % of these are for human health concerns (Schmidt, 2004). Also, an increase in the usage of recombinant proteins is expected for agricultural and other industrial purposes (Palomares *et al.,* 2005). For obvious reasons heterologous protein production goes beyond large scale preparative purposes. It is also a much sought out tool for subsequent research on target proteins. By expressing a given protein a researcher obtains significant amounts of the pure form of the target protein, ideal for ensuing studies (Gileadi *et al.*, 2008). Among the many systems available for heterologous protein production, the Gram-negative bacterium *Escherichia coli* remains one of the most attractive, because of its ability to grow rapidly, and at high density, on inexpensive substrates. In addition, its well-characterized genetics, as well as the increasing availability of cloning vectors and mutant host strain, accounts for an improved versatility of the method (Baneyxn, 1999). However, protein production, primarily of eukaryotic origin, still remains a major challenge in *E.coli* (Peleg and Unger, 2012). The biggest obstacle lies in obtaining large amounts of a given protein in a correctly folded form. Several strategies have been devised to either increase protein solubility or refold incorrectly folded form, and therefore insoluble target protein, in the form of inclusion bodies (Sahdev *et al.*, 2008).

In this work we attempted to develop a working protocol for the heterologous protein expression in *E.coli*, of a soluble form of the antifungal polypeptide BLAD. Given the problems encountered two different expression strategies were tested, namely: i) maximizing soluble protein expression and ii) refolding of the insoluble expressed BLAD protein.

1. Heterologous protein expression in *E. coli*

In general, it is difficult to decide which host and plasmid system is the best for heterologous protein production. The choice often relies in numerous variables including: cell growth characteristics, expression levels, intracellular and extracellular expression, posttranslational modifications and biological activity of the protein of interest. If considering commercial production of the protein, one must also take into account regulatory issues affecting the production of the proteins and consequently economic impact (Makrides, 1996). As a rule of thumb, if the target protein does not rely on post-translational modifications to remain active, a bacterial expression system, such as *E. coli,* is favored (Makrides, 1996). The Gram (-) bacteria *E. coli*, is the workhorse in any molecular biology laboratory. It's easy in handling, possess a rapid growth in inexpensive media, variety of mutant strains, and cloning vectors, along with the high recombinant protein yields and it's scalable potential, make it the ideal organism for heterologous protein expression (Baneyxn, 1999). A typical *E. coli* expression system consists in a specific mutant strain of *E. coli* optimized for protein expression and a cloning vector composed of a circular sequence of double stranded DNA (plasmid) with a promoter region, a multiple cloning site for target gene insertion, a transcription terminator, a selection gene, which is normally an antibiotic resistance gene, and a fusion tag connected with the target gene to facilitate the purification process of the expressed protein. When the gene of interest is cloned into the expression vector using restriction enzymes and placed inside the *E.coli* cells the native DNA replication and transduction machinery assumes this foreign DNA as its own and protein expression takes place, normally induced by specific conditions that activate the promoter such as the presence of IPTG (T7 promoter), pH changes (PcadBA promoter) or low temperatures (PcspA promoter), depending on the promoter. Normally this expressed protein accumulates in the cytosol or periplasmatic space and then with a simple protein extraction, milligram amounts of protein might be obtained (Hannig and Makrides, 1998). The major drawbacks of *E. coli* as an expression system include the inability to perform many of the posttranslational modifications found in eukaryotic proteins (e.g. glycolisation, phosphorylation), the lack of secretion mechanism for the efficient protein release into the culture medium, the limited ability to facilitate extensive disulfide bond formation and most problematic, *E. coli* has a high propensity to produce the target protein in aggregates and therefore insoluble, termed inclusion bodies (Terpe, 2006).

2. Inclusion bodies

Inclusion bodies are electron dense protein products, frequently enriched in beta-sheet structure, that form homogeneous aggregates without cross-aggregation. The aggregation propensity is strongly affected by mutations often occurring during high-level expression of heterologous proteins in *E.coli* (Wang *et al.*, 2008). Inclusion body formation occurs by deposition of misfolded or partially folded polypeptides which, due to the exposition of hydrophobic patches, enhance the consequent intermolecular interactions that otherwise would not take place. The resulting polypeptides are, at least to a great extent, biologically inactive, although important residual enzymatic activity has been reported in a few enzyme model inclusion bodies (Tokatlidis *et al.*, 1991; Peternel *et al.*, 2008). Usually, after expression, inclusion bodies are solubilized by the use of a high concentration of

denaturants such as urea or guanidine hydrochloride that eliminates protein secondary structure and exposes the hydrophobic surface. However, the removal of these denaturants often generates incorrect refolding of the target protein and subsequent insolubilization (Idicula-Thomas and Balaji, 2005). Data collected from a sample of more than 200 non-membrane genes from one bacterium species indicate that 50% of the genes will require further optimization to obtain soluble proteins for crystallization (Roodveldt *et al.*, 2005). Several factors have been suggested to contribute to inclusion body formation in E. coli: (1) high local concentration of the over expressed protein; (2) excessive glutathione levels promoting a reducing environment in the cytoplasm and affecting disulfide bond formation; (3) lack of posttranslational modifications (*e.g.* glycosylation) potentially improving protein solubility; (4) improper interactions with chaperones and other proteins participating in *in vivo* folding; (5) increased aggregation of folding intermediates due to their limited solubility; (6) intermolecular cross-linking via disulfides bridges; (7) kinetics of protein translation in the context of rare codons (8) protein sequence; (9) presence of fusion tags or co expression with molecular chaperones (Idicula-Thomas and Balaji, 2005). Nevertheless, if a protein in a form of inclusion bodies is easily refolded to recover its bioactivity, these structures can be viewed as a major advantage for the heterologous protein production process as they are: highly expressed, easily isolated and highly homogeneous protein products (Singh and Panda, 2005). Therefore, in *E.coli* heterologous protein expression two strategies can be applied. The first involves the use of different techniques to promote inclusion body formation followed by refolding optimization. The other, is to try and minimize the formation of inclusion bodies, by maximizing the expression level of soluble target protein (Hannig and Makrides, 1998).

3. Techniques for maximizing the refolding of inclusion bodies

One way to look at the problem of inclusion bodies formation is to consider it as the first step in the purification of the target protein. Inclusion bodies facilitate straightforward purification of the protein of interest, although the loss of bioactivity of the expressed protein could be a major bottleneck. So far, a major challenge has been to convert these inactive and insoluble inclusion bodies into more efficient, soluble, and correctly folded products (*e.g.* refolding) (Villaverde and Carrió, 2003). The major drawbacks during the refolding of inclusion body proteins into bioactive components are: i) reduced recovery, ii) requirement for rigorous optimization of refolding conditions for each target protein, and iii) possibility that the re-solubilization procedure could affect the activity of the refolded protein (Sahdev *et al.*, 2008). After cell lysis inclusion bodies are normally dissolved in high concentrations of chaotropic reagents such as urea and guanidine hydrochloride. These solubilized proteins are then refolded into their native state during the removal of chaotropic reagent creating an appropriate environment where they can fold spontaneously. Typical refolding strategies use low protein concentrations to deter protein aggregation and slow buffer exchanges from denaturing buffer to aqueous buffer to stimulate correct folding (Kouhei *et al.*, 2003; Singh and Panda, 2005). Several strategies have been devised to improve refolding and reduce protein aggregation. The simpler and the most commonly used is the dilution of the solubilized protein in a renaturation buffer. In this case, a low concentration of denatured protein prevents aggregation and the dilution of the chaotropic

reagent buffer used in the extraction stimulates refold of the target protein. Dilution can be achieved by diluting in multiple, sequential steps or by using continuous fed-batch systems (Kouhei *et al.*, 2003). The problem with dilution is that it brings the unfolded sample into a rapid collapse, sometimes promoting the misfold of the target protein. Therefore a low concentration of denaturant is still required. Also, this technique has considerable limitations if scale-up is in mind. Industrial grade refolding vessels are required which generally makes this process expensive and cumbersome (Sahdev *et al.*, 2008). Another common technique used is dialysis where gradual change of denaturing to native aqueous buffer conditions takes place, converting solubilized and un-folded protein to its native structure. There are two main types of dialysis: one-step dialysis, and a step-wise dialysis, the difference being that in step-wise dialysis protein samples in concentrated denaturant solution are dialyzed against a refolding buffer with decreasing denaturant concentration until they reach zero, unlike one-step where, as the name suggests, the sample is dialyzed against a refolding buffer with no denaturant. This technique performance may be limited if, the rate of misfolding or aggregation is faster than the rate of refolding. The low denaturant concentration may not be enough to keep the unfolded or intermediate structures soluble generating sample precipitation (Kouhei *et al.*, 2003). To circumvent this limitation, additives are added to the dialysis buffer to promote the rapid and correct folding of the protein and decrease aggregation. There are a multitude of additives but, in general, they are low molecular weight compounds, easy to remove from the buffer. Taking into account the extensive literature regarding the different additives used, a database has been created, named REFOLD, which compiled the protocols used to refold more than 350 recombinant proteins. Of these massive amounts of additives, the most frequently used were: L-arginine and glycerol followed by the detergent Triton x-100 (Chow *et al.,* 2006).

4. Techniques for maximizing soluble protein production

One of the first things to consider in soluble protein expression is the choice of the *E.coli* mutant strain. The two most commonly used strains are the BL21, a protease deficient strain, and the K 12 strains, thioredoxin reductase (trxB) and glutathione reductase (gor) expressing strain, with enhanced disulfide bond formation in the cytoplasm. Several derivatives of these strains were developed, some with both their characteristics namely, the Origami B and others tackling different problems also affecting solubility such as, codon bias derived from expressing eukaryotic genes (e.g. BL21 CodonPlus), or overexpression of a toxic protein for *E.coli* (e.g. BL21 Star-pLysS and Origami B pLysS) (Hannig and Makrides, 1998; Terpe, 2006). As mentioned, inclusion bodies are also known to be formed due to high local concentration of the overexpressed protein. Therefore by decreasing the level of protein production one might obtain higher levels of soluble protein. One form of lowering protein production is to have tightly regulated promoters being able to decrease protein expression to a level that it remains soluble (Georgiou and Valax, 1996). The most commonly used promoter system is the T7 RNA polymerase system. In this system, the T7 polymerase gene, which is encoded in the chromosomal DNA, is under the control of a lac promoter that can be induced by the addition to the media of the lactose analog isopropyl-β-D-1- thiolgalactopyranoside (IPTG) thus allowing for protein expression. This system allows for the expression of high amounts of target protein in a controllable fashion by

altering the IPTG concentration (Terpe, 2006). Another way to decrease protein expression is to lower culture temperature. Low growth temperatures strongly decrease the efficiency of traditional promoters, which are used in routine vectors for recombinant protein expression in *E. coli*. Low expression temperature also stimulates correct folding patterns of the target protein, due to the fact that hydrophobic interactions determining inclusion body formation are temperature dependent (Sahdev *et al.*, 2008). The expression of soluble target proteins might also be accomplished by co-expressing our target protein with specific proteins that help the folding process, namely, molecular chaperones and foldases. Molecular chaperones are ubiquitous and highly conserved proteins that help other proteins reach their final structure. Nevertheless, they are not true folding catalysts since they do note accelerate folding rates. Normally, different sets of chaperones act cooperatively during folding attempts. The coordinated activity of the molecular chaperones DnaK, DnaJ and GrpE and of GroEL and GroES has been well described and they are commonly used in *E.coli* protein overexpression (Baneyx and Palumbo, 2003). By contrast, foldases are true folding catalysts and are responsible for increasing the assembly rhythm of correctly folded proteins. Therefore, by co-overexpressing these proteins one might increase the level of correctly folded target protein. In protein expression four types of foldases have been used so far: peptidyl prolyl cis/trans isomerases (PPI's), disulfide oxidoreductase (DsbA), disulfide isomerase (DsbC) and protein disulfide isomerase (PDI) (Villaverde and Carrió, 2003). The soluble expression of the target protein can also be achieved by fusing our target protein to a more soluble protein (tag) in bacteria, which will serve as a folding aid and later, as a purification tag. Protein fusion tags have been shown to act as solubility enhancers and chaperones but neither mechanism is well understood. Possibly, the fusion of a stable or conserved structure to an insoluble recombinant protein stabilizes and promotes proper folding. Moreover the recombinant protein or fusion tags may act as a nucleus for proper folding ("molten globule hypothesis")**.** Examples of popular fusion tags include: glutathione-s-transferase (GST), maltose binding protein (MBP), NusA, thioredoxin (TRX) (Yasukawa *et al.*, 2005), polyhistidine (HIS), small ubiquitin-like modifier (SUMO) and ubiquitin (Ub) (Hall, 2007)**.** Alternatively protein solubility can also be improved through protein engineering, either by rational design, through comparison with homologous proteins already soluble (Eijsink *et al.*, 2004), or by direct evolution through random mutations to improve the chances of obtaining a soluble protein upon expression (Roodveldt *et al.*, 2005). Independently of the approach, scientists might be able to calculate protein solubility, and change, by point mutations, the primary sequence of the target protein in order to improve solubility (Pastor *et al.*, 2005). There are now several examples of proteins that have been stabilized by the introduction of mutations with small yet cumulative stabilizing effects (Idicula-Thomas *et al.*, 2006).

5. BLAD as a target protein for heterologous production in *E.coli*

Storage proteins are defined as any seed protein that accumulates in significant quantities, with no known function during seed development, and rapidly hydrolyzed upon germination, acting as a source of nitrogen and carbon on the early stages of seedling growth (Foley *et al*., 2011). According to the currently used classification, seed storage proteins are classified based on their solubility and fall into one of the following categories: (a) globulins or salt soluble fraction, (b) albumins or water soluble fraction, (c) glutelins or dilute acid/alkali fraction, and (d) prolamins or alcohol-soluble fraction (Foley *et al*., 2011). The globulins fraction is also subdivided based on their sedimentation coefficients into 7S or vicilins and 11S or legumins (Tandang-Silvas *et al.,* 2011). The *Lupin* sp. seed storage globulin protein fraction is comprised of four distinct types of proteins, termed: 11S globulin (also known as α-conglutin, legumin, legumin-like and glycinin), 7S globulin (also known as β-conglutin, vicilin, convicilin and vicilin-type), 7S basic globulin also known as □-conglutin) and 2S sulphur-rich albumin also known as δ-conglutin) (Foley *et al*., 2011). The major globulin protein present on the lupin seeds is the β-conglutin or vicilin-type fraction with a sedimentation coefficient around 7S, and these are trimeric proteins in which the monomers consist of a number of polypeptides ranging from 16 to over 70 kDa, with no disulphide bridges linking them (Duranti *et al*., 1981). Studies on this protein catalysis, during seed germination, show that β-conglutin suffers a rapid change in its structure and concentration. With the transient accumulation of stable polypeptide intermediates from the 4 th day of germination until the 14 th. Specifically an abrupt accumulation of a 20 kDa polypeptide occurs. This peptide, BLAD ("Banda de *Lupinus albus* Doce"), displays a strong lectin and antimicrobial activity (Ramos *et al.*, 1997). BLAD is a non-glycosylated, phosphorylated, 20 kDa polypeptide that strongly interacts with a variety of glycoproteins in a lectin type fashion it possesses high thermal stability, resistance to methanol and sodium dodecyl sulphate and a strong antimicrobial activity against Gram (+) bacteria and fungi. The complete sequence of this stable intermediate of β-conglutin digestion contains an extremely high proportion of the nitrogen-rich amino acids, notably arginine (18 residues out of 173), asparagine (17 residues), glutamine (11 residues) and lysine (7 residues). Analyses of the aminoacid sequence show strong similarities (99-86 %) with the β-conglutin form from *Lupinus angustifolius,* followed by more distant similarities for β-conglycinin form soya seeds (56 %), major peanut allergen Ara H 1 (55 %) and Adzuki Bean 7s Globulin-1 (53 %) (Monteiro *et al.*, 2010). Since BLAD is non-glycosylated, and with a total absence of disulphide bridges due to the lack of cysteine residues, a heterologous bacterial expression system might be the appropriate choice for expression system (Terpe, 2006). Obviously problems may arise due to the proteins phosphorylation. However several phosphorylated proteins have already been purified from bacterial expression systems with intact biological activity. In some cases, endogenous phosphorylation in *E.coli* has been observed (Sahdev *et al.*, 2008). Therefore, BLAD, being a highly stable, non-glycosylated, antifungal polypeptide, impossible to isolate in nature, constitutes the ideal target for heterologous expression in *E.coli*. Although the heterologous expression of BLAD has never been attempted, the already mentioned, polypeptides similar to BLAD, have already been successfully expressed in *E.coli* and can serve as example protocols for BLAD. The Soybean β-Conglycinin cDNA sequence was placed into a pET-21d vector (Novagen company) with a T7 promoter and kanamycin selection, and transformed into *E. coli* strain BL21 (DE3)

(Invitrogen). These cells harboring the expression plasmid were grown at 20°C and induced through the addition of IPTG. The soluble fraction of the expressed proteins was collected through centrifugation and the target protein purified through anion-exchange chromatography in a Q-Sepharose (Maruyama *et al.*, 1998). In the case of the major peanut allergen Ara H 1 subunit, its encoding cDNA was placed into a pET-21d plasmid and transformed into *E. coli* BL21 (DE3) strain (Invitrogen). Cultures were induced by the addition of IPTG to a final concentration of 1 mM at 30 ºC. The soluble expressed Ara h1 subunit was then purified through a combination of ammonium sulfate fractionation, ion affinity and size exclusion chromatography (Cabanos *et al*., 2010). In our last example the adzuki bean β-conglycinin α, α' and β subunits were expressed in pET21d plasmids (Novagen company) and transformed into three different *E. coli* strains HMS174(DE3), BL21(DE3), AD494(DE3), and Origami(DE3). For all three constructs, and in all the three tested *E.coli* strains, there was some degree of soluble protein production. In all cases, protein purification was achieved through ammonium sulfate fractionation and ion exchange chromatograph (Fukuda *et al.*, 2008).

Material and Methods

1. Plasmid construction

Different plasmids and BLAD genes were used in this work, and the constructs obtained were termed accordingly (see Annex III.1. and III.4, for BLAD gene sequences and pET 151 plasmid map):

- pET151BLAD173his- pET151BLAD D-TOPO plasmid plus BLAD cDNA encoding the 173 amino acid BLAD polypeptide (full BLAD gene, NCBI database accession number DQ142920.1);

- pET151BLAD127his- pET151BLAD D-TOPO plus BLAD cDNA encoding a 127 amino acid shortened version of the BLAD polypeptide, missing its N and C terminal regions;

- pColdBLAD173his- pCold I plasmid plus BLAD cDNA encoding the 173 amino acid BLAD polypeptide (full BLAD gene).

1.1.PCR and DNA purification

Lupin albus cDNA, stored at -80 ºC, was PCR ("Polymerase Chain Reaction") amplified using specific primers. The primers sets varied according with the BLAD construct and plasmid desired (see Annex III.3, for primers used). Regardless of the primer set chosen, the PCR reaction was performed in a Bio-Rad device (C1000PCR) using the taq (*Thermus aquaticus*) pfx (proofreading) polimerase (Invitrogen), with the following PCR program: 94 ºC, 2min, followed by 30 cycles of 94°C for 15 s, 55°C for 30 s and 68°C for 1 min. The buffers used for the PCR reaction were in following working concentration: 1 X Pfx Amplification Buffer, 0.3 mM dNTP mixture, 1 mM MgSO4, 1 µM Primer mix, 1 µL of *Lupin* sp. cDNA, 0.4 µL of taq Pfx DNA Polymerase, all adjusted to 50 µL of autoclaved Mili-Q water. PCR products were separated in horizontal agarose gel electrophoresis containing 5 µg/mL of gel red dye (Biotium), to stain nucleic acids. When finished, the electrophoresis was visualized in a UV transiluminator (Geldoc, Bio-Rad) using the "Quantity One" software (Bio-Rad). After visualizing the gel, the selected DNA bands were cut and purified using the "Wizard purification" kit (Promega). Following purification the DNA samples were quantified in a Synergy HT spectrophotometer (Biotech), and the results registered for further use.

1.2. Cloning into pET151/D-TOPO

This specific cloning was performed following the suppliers instruction on the "Champion pET directional TOPO expression kit" (Invitrogen). The cloning into the pET151 plasmid (Invitrogen) is conditional on four over-hanged nucleotides (GTGG) present in the 3' extremity of the cloning site in the plasmid. Our target DNA molecule had to be PCR amplified with the complementary bases in its forward primer (CACC) and incubated for 5 minutes at room temperature with the linearized pET151/D-TOPO expression vector to be directionally cloned. The now circular plasmids were then transformed into "One Shot TOP10 Competent Cells" for multiplication and maintenance.

1.3. Cloning into pCold I

The pCold I plasmid (Takara), unlike the pET151, has a multiple cloning site flanked by several nucleotide sequences recognized by specific restriction enzymes (see Annex II.3. for vector map). On this work we chose the NdeI (Roche) and XbaI (Roche) enzymes, and proceeded to doubly digest the pCold I plasmid circular DNA and the BLAD DNA, already PCR amplified with the specific set of primers that add restriction sites for this enzymes to our known BLAD sequence. A digestion reaction was then set up for 2 h/ 37 ºC, with: 1 µg of DNA, 1 unit of each enzyme, 2.5 µL "Sure cut" Buffer A (Roche) in 25 µL Mili-Q water. Afterword both were purified with the "Wizard purification kit" and quantified in a Synergy HT spectrophotometer. With both the plasmid and the target gene digested and quantified a ligation reaction was set up, where a molar ratio of 1:2 between plasmid and insert was strictly kept. Therefore the ligations reaction was set up with the following reagents: 100 ng of total DNA, 4 µL of 10xLigase Buffer, 0.25 µL of T4 DNA Ligase (1 U/µL) all adjusted to a 20 µL volume with autoclaved Mili-Q water. The reaction was left at 4 ºC overnight. On the following day the plasmid was transformed into "One Shot TOP10 Competent Cells", for multiplication and maintenance.

1.4. Transforming "One Shot TOP10 Competent Cells TOPO"

When the ligation reaction was completed, 3 µL of ligation reaction were removed and put into one vial of "Oneshot TOP10 Competent Cells TOPO" (Invitrogen). This vial was incubated on ice for 30 min. After this time the cells were heat shocked for 30 seconds at 42 ºC in a water bath. Immediately after the vial was transferred back to ice and left to rest for 2 min. When the two minutes passed 250 µL of Luria-Bertoni medium (LB) (Annex II.1) was added to the mixture and was incubated for 1 hour at 200 rpm at 37 ºC. Following the incubation step, three volumes, 50, 100 and 200 µL, were spread on a pre-warmed Luria-Bertoni solid medium (LBA) petri dish supplemented with, 100 µg/µL of Ampicilin (Sigma-Aldrich), and incubated overnight at 37 ºC. On the following day, five isolated colonies were selected and tested for the correct insertion of the gene through a digestion reaction with specific enzymes (Xba I and Hind III restriction enzymes for the PET151/D-TOPO plasmid, and Nde I and XbaI for the pCold I plasmid). When proven positive, a single colony from each cloning reaction was grown in 50 mL of LB medium and consequently higher amounts of plasmid DNA were purified using the "Jetstar Plasmid Midiprep Kit" (Genomed), ideal for the subsequent steps. In the end after sufficient amounts of plasmid DNA were obtained all the new constructs were seuquenced by the Stavida company, according with its internal specifications. The sequence results were analyzed using the MULTIALIGN software (Corpet, 1988), by comparison with BLAD published nucleotide sequence (NCBI database accession number DQ142920.1). The plasmid obtained were then ready to be transformed into the "BL21 Star™(DE3) One Shot" *E.coli* cell line.

1.5. Transforming BL21(DE3) Competent Cells transformation

To express our target protein the plasmid DNA had to be introduced into a specific strain of *E.coli,* BL21 Star™(DE3) One Shot cells, following the suppliers instructions (Invitrogen). For every transformation reaction 10 ng of plasmid DNA was added to one vial of *E.coli* cells. The mixture was mixed and incubated on ice for 30 minutes. Afterwards the vial was incubated ate 42 ºC for 30 s and

immediately transferred to ice. After, 250 µl of room temperature LB medium was added and the mixture was incubated at 37°C for 30 minutes with shaking (200 r.p.m.). The entire transformation reaction was added to 10 mL of LB containing 100 µg/µL of ampicillin and grown overnight at 37°C with shaking.

2. Protein expression

2.1. Protein expression - pet151D TOPO plasmid

The protein expression procedure for this plasmid was conducted according to the "Champion pET directional TOPO expression kit" and can easily be subdivided into two types: an optimization expression or low volume expression, and a high volume expression. The low volume approach was used to optimize expression conditions, such as temperature and incubation time. The high volume approach was used, for determining the type of protein expression, soluble or insoluble, and as a means to obtain sufficient protein amounts for subsequent tests, such as the refolding procedures.

In low volume expression 500 µL of overnight culture from the previous step was placed in 10 mL of LB medium (containing the 100 µg/µL ampicillin). The culture was allowed to grow at a set temperature (normally 30 ºC). The culture growth was monitored by determining the optical density at 600 nm (OD600) in a Shimadzu UV 1800 spectrophotometer. When the OD_{600} reading reached 0.5, about mid log, the culture was split into two 5 mL cultures. To one culture, IPTG was added to a final concentration of 1 mM, the other served has a non-induced control culture. After IPTG addition, a 500 µL sample was taken from each culture at one hour interval for 24 hours, to determine the best incubation time for protein expression, through SDS-PAGE analysis. When the best conditions were determined, the *E.coli* was allowed to grow in bigger volumes to obtain larger working samples. High volume expression consisted on a 10 times larger culture volume. So a 50 mL LB medium with ampicillin was inoculated with 1 mL of the overnight culture from the previous. Afterword the culture was at the set temperature at 250 r.p.m. until the OD600 reached 0.5. When this value was met, 1mM of IPTG was added and culture was again allowed to grow for 6 hours. The grown cells were harvested by centrifugation at 3.000 *g* for 10 minutes at 4 ºC and stored at -80 ºC.

2.2. Protein expression- pCold I plasmid

This plasmid due to its unique characteristics was tested in different temperature conditions according with the suppliers instructions (Takara, 2008). A volume of 500 µL of inoculum from the overnight culture step was placed in 10 mL of LB medium (containing the 100 µg/µL ampicilin). The culture was allowed to grow at 37 ºC at 250 r.p.m. The culture growth was monitored by determining the optical density at 600 nm (OD600) in a Shimadzu UV 1800 spectrophotometer. When the OD600 was 0.5, about mid log, the culture was split into two 5 mL cultures and refrigerated until it reached 15 ºC. To one culture IPTG was added to a final concentration of 0.5-1 mM, the other serve has a non-induced control, and both cultures were grown at 15 ºC. After the IPTG addition a 500 µL sample was taken from each culture at one hour interval for 24 hours, to determine the best incubation time for protein expression, through SDS-PAGE analysis.

3. Protein extraction and purification

The protein extraction protocols can be subdivided into two distinct types, termed: native and denaturing. The main differences are based, in whether our target protein is soluble expressed in *E.coli* or not. The native protocol relies on the protein maintaining its conformation and therefore being soluble expressed in *E.coli*. In the denaturing procedure our protein is insolubly expressed in *E.coli* and the sample as to be treated with high urea concentrations to dissolve it in a solution.

3.1. Protein extraction and purification- Native protocol

In this protocol, 50 mL overnight culture was centrifuged at 2500 *g*/ 5 minutes in an Allegra 25 R centrifuge. The retrieved cells were then ressuspended in 8 mL of Native Binding buffer (250 mM NaH_2PO_4, 2.5 M NaCl pH 8.0), and disrupted through sonication (Hielscher UP50H) at 80 % amplitude, 0.5 cycles/10 seconds, with a cooling period between each burst. To recover the soluble proteins, the lysate was centrifuged at 3.000 *g*/15 minutes and the supernatant recovered and analyzed through SDS-PAGE (termed sample "Initial") and further purified in nickel columns. After extraction, an 8 mL sample of soluble protein was applied to a 2 mL Ion metal affinity chromatography column, containing agarose with bound nickel ions (His-column). Nickel ions exhibits a high affinity towards histidine residues that are present in our target proteins due to a 6 histidine fusion tag present in both our plasmids (Invitrogen). The sample was then incubated with the column for 30 minutes using gentle agitation to keep the resin suspended in the lysate solution. The agarose was allowed to settle by gravity and a sample collected, termed sample "Flow-through", followed by the column wash with 20 mL of Native Wash Buffer (250 mM NaH_2PO_4, 2.5 M NaCl, 20 mM Imidazole pH 8.0) (sample named "Wash). Following the wash step the sample was eluted from the column with 5 mL Native Elution Buffer (250 mM NaH_2PO_4, 2.5 M NaCl, 250 mM Imidazole pH 8.0). Samples were taken for each purification step and analyzed through SDS-PAGE and the protein content quantified.

3.2. Protein extraction and purification- Denaturing protocol

In this protocol a 50 mL overnight culture was centrifuged at 2500 *g*/ 5 minutes in an Allegra 25 R centrifuge. The retrieved cells were then ressuspended in 8 mL of Denaturing Binding buffer (20 mM NaH_2PO_4, 500 mM NaCl, 8 M Urea at pH 7.8) and disrupted through sonication at 80 % amplitude, 0.5 cycles/10 seconds, with a cooling period between each burst. To recover the insoluble proteins the lysate was centrifuged at 3.000 *g*/15 minutes and the supernatant recovered and analyzed through SDS-PAGE and further purified in nickel columns. After extraction, an 8 mL sample of insoluble protein content was applied to a 2 mL Ion metal affinity chromatography column, containing agarose with bound nickel ions (His-column). The sample was incubated with the column for 30 minutes using gentle agitation to keep the resin suspended in the lysate solution. The agarose was allowed to settle by gravity and a sample collected, termed "Flow through". "After the column wash with 10 mL of Denaturing Binding Buffer and a sample collected , named "Wash pH 7.8" another wash step was carried out with 10 mL of Denaturing Wash Buffer (20 mM NaH_2PO_4, 500 mM NaCl, 6 M Urea at pH 6) and a sample collected , named "Wash pH 6". After washing, the sample was eluted from the column

with 5 mL Denaturing Elution Buffer (250 mM NaH_2PO_4, 2.5 M NaCl, 250 mM Imidazole pH 4). Samples were taken for each purification step and analyzed through SDS-PAGE and the protein content quantified. After purification all the samples had to be subjected to the dialysis procedure in order to remove the high urea concentrations and in attempt to refold the protein back to its native conformation.

4. Dialysis, Dilution and combinations procedure

Independent of the refolding methods tested the dialysis procedure was similar. A pre-wet dialysis membrane (Thermo-Scientific) with a "Molecular Weight Cut-Off" of 2.000 Da was cut in appropriate sizes according to the sample volume. The dialysate was poured into the tubing, and the tubing sealed shut. The volume difference between sample and buffer was always kept at 1000 fold, and dialysis was allowed to run at 4 ºC overnight with gentle agitation, with 2 buffer replacement. With 50 mL *E.coli* cultures expressing insoluble forms of the BLAD polypeptide, four types of refolding procedures were carried out:

1) One step dialysis: where the urea is removed against a buffer without urea (20 mM NaH_2PO_4, 500 mM NaCl at pH 7.8);

2) Step-wise dialysis: the sample was dialyzed against a buffer with ever decreasing Urea concentrations (4, 3, 2, 1, 0 M) while maintaining the other constituents of the buffer (20 mM NaH_2PO_4, 500 mM NaCl at pH 7.8);

3) Dilution/step-wise dialysis: while maintaining the same buffer (20 mM NaH_2PO_4, 500 mM NaCl at pH 7.8) the urea was first diluted to a concentration of 1 M and then dialyzed against a buffer containing no urea (20 mM NaH_2PO_4, 500 mM NaCl at pH 7.8);

4) Alternate version of Dilution/step-wise dialysis: after denaturing protein extraction, each sample was diluted tenfold, using the standard denaturing elution buffer with 6 M urea, and immediately after, Step-wise dialyzed against a buffer with ever decreasing urea concentrations (4, 3, 2, 1, 0 M of urea), while maintaining the other constituents of the buffer (20 mM NaH_2PO_4, 500 mM NaCl at pH 7.8).

After dialysis the samples from the dilution/step wise protocol were concentrated (10x) to the original volume using Amicon Ultrafiltration apparatus with a MWCO of 10 kDa (Millipore).

5. Protein quantification

For protein quantification two types of the Bradford quantification method were employed, depending on the presence of the chaotropic reagent, urea.

The standard Bradford assay is based on the observation that the absorbance maximum for an acidic solution of Coomassie Brilliant Blue G-250 shifts from 465 nm to 595 nm when binding to protein occurs. Both hydrophobic and ionic interactions stabilize the anionic form of the dye, causing a visible color change. The calibration curve was constructed with bovine serum albumin (BSA), with a protein gradient between 1 and 35 µg, in a final volume of 100 µL (0.002 – 0.05 g/L). Five different concentrations of BSA were prepared, in triplicate, starting from a BSA 0.5 g/L solution (stored at -20 ºC), and put in a 1 mL plastic cuvette (PVL). To this preparation, 1 mL of Bradford reagent (Sigma-Aldrich) was added, stirred and left to incubate in the dark for 10 min. After this time elapsed the absorbance values were read at 595 nm in a spectrophotometer (Shimadzu UV-2100), against a blank, where the volume of the protein solution is substituted by Milli-Q water. Determination of the protein content of the samples was performed analogously, diluting the sample with Milli-Q water until a final volume of 100 µL. This dilution is critical since the absorbance values measured have to be within the calibration curve limit values. Under these conditions, the extrapolation of the protein content was performed based on the tendency line, acquired by the least squares method, from the average value of absorbance relative to the each BSA concentration (Bradford, 1976). The binding of the Coomassie Brilliant Blue G-250 dye to proteins is highly affected by the presence of basic reagents such as urea and basic carrier ampholytes. The modified Bradford assay corrects the pH of the sample through the addition of a diluted acid solution to neutralize the protein sample. So that in the presence of samples containing urea, the modified Bradford method allows for their direct quantification by generating two near-linear segments, one over the range < 0.5 to 5 µg total protein that permits the application of Beer's law and a second linear response encompassing 5 to 50 µg total protein (Ramagli and Rodriguez 1985). Regarding the procedure, the calibration curve for the modified Bradford was prepared in the same concentrations as the standard protocol, but the BSA samples were all dissolved in a 20 mM Sodium phosphate, 500 mM NaCl, 6 M Urea at pH 7.8 solution. Ten micro liters of each BSA dilution were put on ELISA 96 well plate and the pH adjusted with 90 µL of a 0.1 N HCl solution. To each sample well 150 µL of Bradford solution was added and the Elisa plate was left to incubate in the dark for 10 min. After this time elapsed the absorbance values were read at 595 nm in a microplate reader (Bioteck Synergy Ht), against a blank, where the volume of the protein solution is substituted by 500 mM NaCl, 6 M Urea at pH 7.8 solution. Determination of the protein content of the samples was performed analogously (Ramagli and Rodriguez, 1985).

6. SDS-PAGE Electrophoresis

Polypeptides were separated in 17 % acrylamide gels and the composition of the different buffers used was as follows:

- Stacking gel- 5 % (w/v) acrylamide (Sigma-Aldrich), 0.13 % (w/v) bis-acrylamide (Sigma-Aldrich), 125 mM Tris-HCl pH 6.8, 0.1 % (w/v) SDS, 0.1 % (w/v) PSA (Sigma-Aldrich) and 0.05 % (v/v) TEMED (Sigma-Aldrich);
- Resolving gel- 17 % (w/v) acrylamide, 0.1 % (w/v) bis-acrylamide, 375 mM Tris-HCl pH 8.8; 0.01 % (w/v) SDS, 0.03 % (w/v) PSA and 0.05 % (v/v) TEMED;
- Cathode buffer- 25 mM Tris-HCl solution, 192 mM Glycine (Merck);
- Anode buffer- 25 mM Tris-HCl solution, 192 mM Glycine and 0.1 M sodium acetate (Merck);

The first step in preparing the samples for electrophoresis was the protein precipitation in 80% (v/v) cold (- 20 ºC) acetone (Merck) for 30 min at -20 ºC, in order to remove unwanted contaminants. When this time elapsed, the samples were centrifuged at 10.000 g for 10 min, the supernatant discarded, and the precipitate dissolved in a reducing buffer solution termed SBM, which comprised of: 0.08 M Tris-HCl pH 6.8; 0.1 M 2-mercaptoethanol (ME) (Sigma-Aldrich), 2 % (w/v) SDS, 15 % (v/v) glycerol (Merck) and 0.006 (w/v) m-cresol purple (Sigma-Aldrich). After completely dissolved, the protein sample was denatured, by boiling for 3 min in SBM. To sort out our complex polypeptide mixture by size, protein standards with known molecular masses were used. The chosen marker was the Low protein molecular markers (LPM) from Sigma-Aldrich with a protein size range from 14.2 kDa to 66 kDa. The electrophoresis ran at constant current of 45 mA, using an EPS 500/400 power source (Pharmacia/LKB), until the m-cresol purple from the SBM present in the samples reached the end of the gel, then the gel was promptly removed and treated according with the chosen staining protocol.

7. Immunoblot

The electrotransfer was performed in a semi-dry system (“Transblot Semi-dry Transfer cell”, BioRad) using nitrocellulose membranes with a pore size of 0.2 µM (Millipore). After polypeptide separation by SDS-PAGE in 17 % acrylamide gels, the gel and membrane were submerged for 15 min in transfer buffer (50 mM Trizma base, 3.7 mM Glycine, 0.04 % (w/v) SDS and 20 % (v/v) Methanol). After this time elapsed the gel and membrane were placed in the transfer apparatus in between 8 sheets of 3MM paper (Millipore), previously soaked in transfer buffer. Air bubbles were carefully removed from the paper stack. The transfer took place for 45 min at a constant 15 v (Twobin et al., 1979). Following transfer, the polypeptides were permanently set to the membrane by boiling the membrane for 10 minutes. The membrane was then blocked for 1 hour with buffer PBST0.05% (137 mM NaCl, 1.5 mM KH2P04, 8.1 mM Na2HPO4, 2.7 mM KCl, 0.01 %(w/v) NaN3, 0.05 % (v/v) Tween 20, 2 % Powder milk). After blocking the membrane was incubated for 1 h at 37ºC, with the specific antibody (rat antibody, diluted 1:500) in PBST0, 0.5% buffer. Following antibody incubation the membrane was washed two times for five minutes with the buffer PBST0.1% (137 mM NaCl, 1.5 mM KH2P04, 8.1 mM Na2HPO4, 2.7 mM KCl, 0.01 %(w/v) NaN3, 0.05 % (v/v) Tween 20) and another two times for 10

minutes with the buffer solution SAIS (1 M NaCl, 10 mM Na2PO4 e 0.5 % (v/v) Tween 20). After these washing steps the membrane was again incubated for 1 h at 37 ºC with secondary anti-rat antibody (Sigma-Aldrich), diluted at 1:30.000. Past this second incubation step the membrane was washed, as previously described, with buffers PBST0.1% and SAIS, and afterword, it was equilibrated for 5 minutes in detection buffer PBS (0.1 M NaCl and 0.1 M Tris- HCl pH 9.5). The revelation procedure was based on the method proposed by Blake et al (1984), whereby the phosphatase alkaline when in the presence of it's subtract originates a purple and insoluble pigment. Therefore the membrane was incubated for at 37 ºC with detection buffer (5 mM MgCl2, 0.3 mg/mL NBT (nitro-blue tetrazolium chloride) and 0.15 mg/mL BCIP (5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt)), until the stained polypeptide bands are visible.

Results

Since this was the first attempt to BLAD polypeptide heterologous expression in *E.coli*, we began our work by analyzing the expression protocols of three proteins similar to BLAD, namely: soy bean β-conglycinin, peanut Ara h1 and adzuki bean β-conglycinin. All three protocols used a T7 regulated plasmid associated with a protease deficient *E.coli* strain, BL21DE3 (Maruyama *et al.*, 1998; Fukuda *et al.*, 2008; Cabanos *et al.*, 2010). Considering this, we opted to use the "Champion pET 151 Directional TOPO" kit. This kit allows an easily placement of the target gene into a T 7 regulated pET 151 D TOPO expression plasmid associated to a protease deficient *E.coli* BL 21DE3 strain.

Therefore, we cloned BLAD (173 aa) into the pET151 D-TOPO plasmid obtaining the plasmid pET151BLAD173his and place this new plasmid into the BL21DE3 *E.coli* strain for posterior expression optimization. The protein was expressed in a small volume (5 mL) according with the supplier instructions (Invitrogen) and different samples at time points were taken to determine the optimal time for expression. Bear in mind that BLAD protein is a 20 kDa protein, but when expressed in *E.coli*, due to the different tags associated with the plasmid (his-tag, TEV protease), gains around 4 kDa. As a result the real size of our expressed target protein is in the vicinity of 24 kDa.

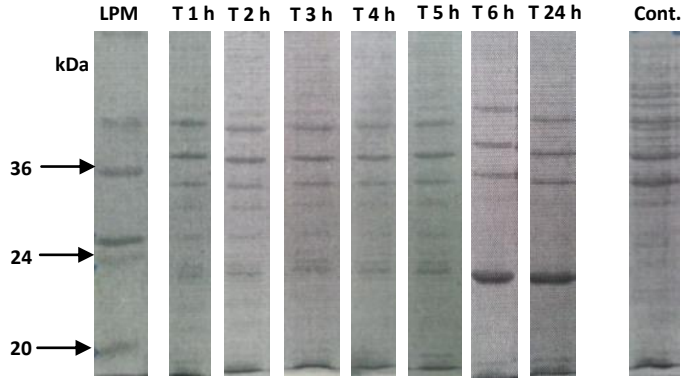


Figure 14: SDS-PAGE (12.5 % of Acrylamide, Coomasie R-250 stained) of the total protein content from the lysis of *E.coli* grown at 37 ºC for 24 hours containing the plasmid pET151BLAD173his. Sample lanes have 10 µL of lysis *E.coli* grown at pellet cells at different time points ("T1h"- 1 hour of growth; "T2h" – 2 hours of growth; "T3h – 3 hours of growth; "T4h" – 4 hours of growth; "T5h" – 5 hours of growth; "T6h" – 6 hours of growth and "T24h" – 24 hours of growth). Control sample, Lane "Cont.", contains a 10 µL sample of *E.coli* without induction of expression. Lane "LPM" refers to the Low protein molecular marker.

Considering this, and analyzing Figure 14, we can conclude that, our target protein was correctly expressed in *E.coli*, and the optimal time point for expression is the six hour mark. In this view, we set out to purify our target protein through metal affinity chromatography, using nickel columns, and determine the extent of its solubility. We extracted the total soluble protein content from a 50 mL, six hour induced cell culture, using the Native Purification procedure (see 2.4.1) and analyzed this through SDS-PAGE (Figure 15, A). The insoluble part, from the Native Purification procedure, had its protein content extracted through the Denaturing purification Procedure and was also purified through metal affinity chromatography (see 2.4.2.), and, as well, analyzed by an SDS-PAGE (Figure 15, C).

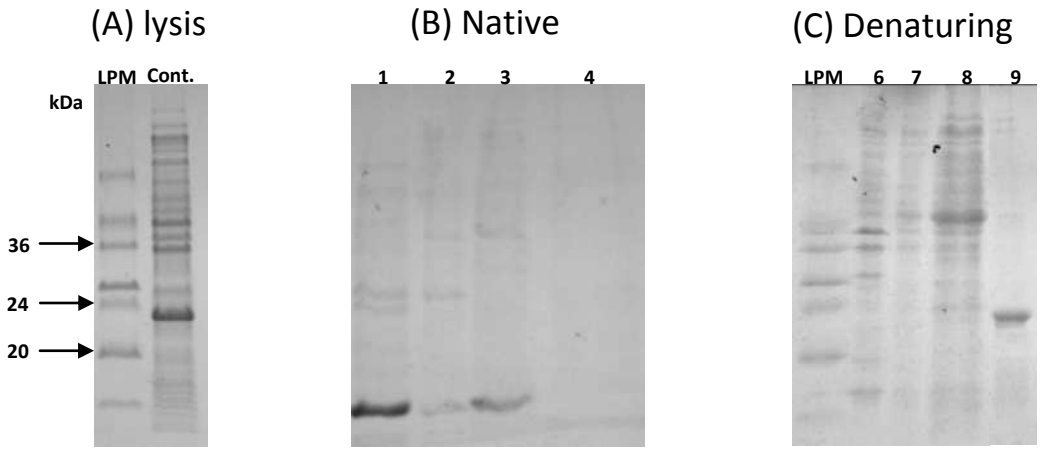


Figure 15: SDS-PAGE (17.5 % of Acrylamide, Coomassie R-250 stained) containing the different fractions from a nickel column used in the Purification procedures: Native Purification procedure (A) and the Denaturing purification procedure (B). Protocol applied to the cells from a 6 hour induced culture of *E.coli* containing the plasmid pET151BLAD173his. Each lane as 30 µg of protein and lanes are in figure 2.A: Cont. – control lane with 100 µL of pellet cells of induced *E.coli* dissolved in the SDS-PAGE sample Buffer; 1- Flow through; 2-wash sample (5 mL); 3- Wash sample (5 mL); 4- Elution. Lanes in figure 2.B are: 6- Flow through; 7- Wash pH 7.8; 8- Wash pH 6; 9- Elution pH 4. Lane "LPM" refers to the Low protein molecular marker.

Examining the SDS-PAGE gels from both these extractions and purification procedures (Figure 15, A, B and C), it is clear that the expressed 24 kDa, BLAD polypeptide is only found in the eluted fraction of the Denaturing Purification protocol, unlike all the three analyzed "BLAD-like" proteins. As a result it can be said that our target protein was insolubly expressed in *E.coli*, possibly as inclusion bodies, and could only be retrieved through the use of high urea concentrations (~6 M) present in the Denaturing purification procedure.

1. Techniques for maximizing the refolding of inclusion bodies

Considering the insoluble expression of the BLAD polypeptide, we set out to remove the high urea concentration, and consequently trying to refold our target protein, using in this manner, this insolubility as a simple purification step (Villaverde and Carrió, 2003).

1.1. One step dialysis, Step-wise dialysis, Dilution/Dialysis and Adjuvant-assisted Dialysis

One documented problem of any refolding procedure, is that usually, the rate of misfolding and aggregation is higher than the rate of refolding, originating a higher concentration of improperly folded proteins (Tsumoto *et al.*, 2003). One way to circumvent this problem is opt to slowly diminish the urea concentration, instead of a rapid removal, thereby allowing for a properly folding process (Singh and Panda, 2005). In view of this, we designed a refolding experiment to test the refolding potential of the insolubly expressed BLAD, when subjected to three separate dialysis and dilution procedures:

1) One step dialysis: the urea was removed through dialysis against a buffer without urea;
2) Step-wise dialysis: the sample was dialyzed against a buffer with ever decreasing urea concentrations (4, 3, 2, 1, 0 M of urea), while maintaining the other constituents of the buffer;
3) Dilution/step-wise dialysis: while maintaining the same buffer the urea was first diluted to a concentration of 1 M and then dialyzed against a buffer containing no urea.

After dialysis the sample was concentrated to the original volume and analyzed. The turbidity (590 nm) and protein concentration was determined for each sample (total protein) prior to a centrifugation step, to remove the insoluble protein fraction. The protein content from the supernatant was quantified, thereby allowing estimating the percentage of soluble protein.

	Turbidity (590 nm)	Protein quantity (µg)		Soluble protein (%) (p/p)
One-step	0,341	Total protein	1200	0,83
		Soluble protein	10	
Step-wise	0,184	Total protein	950	1,02
		Soluble protein	10	
Dilution/Dialysis	0,265	Total protein	850	0,58
		Soluble protein	5	

Table 8: Effect of different refolding procedures of soluble protein obtained (after centrifugation) and on the level of turbidity of the protein sample (as read by absorbance at 590 nm). Protein quantified as a percentage of total protein (before centrifugation). The protein samples are the result of the induced expression of an *E.coli* containing the expression plasmid pET151BLAD173his and the posterior protein extraction and purification using the Denaturing procedure. The percentage of the soluble protein content is the supernatant part obtained through the centrifugation at 5000 *g* for 10 minutes of the total protein content. All protein quantifications were obtained through the Standard Bradford method.

Comparing the three tested refolding methods (Table 8) the Step-wise dialysis produced higher amounts of soluble BLAD (1.02 %) when compared to One-Step and Dilution/dialysis (0.83 % and 0.58 %, respectively). But none of this tested methodologies produced any working level of soluble BLAD. Consider this, to obtain 100 mg of soluble BLAD polypeptide using the Step-wise dialysis alone one needs 500 L of starting culture. Therefore, all these procedures are useless as working refolding procedures.

1.2. Adjuvant-assisted Step-wise dialysis

To improve the refolding percentage using the Step-wise dialysis method, we tried adding small molecules, or adjuvants, to the exchange buffer in order to correctly fold the target protein and decrease aggregation (Singh and Panda, 2005). As mentioned before, there are a great number of refolding adjuvants. To sort out this clutter of data, we used the REFOLD database, responsible for compiling all the submitted refolding protocols. After this database analysis we selected the following adjuvants to test in a Step-Wise dialysis of the BLAD polypeptide: L-Arginine (15.15 % of REFOLD database entries), Glycerol (12.56 % of REFOLD database entries) and Triton 100x (2.75 % REFOLD of database entries) (Chow *et al.*, 2006).

The total protein content from 3 x 50 mL induced *E.coli* culture (containing pET151BLAD173his plasmid) was separately extracted and purified using the Denaturing procedure. These three separate protein samples were Step-wise dialyzed, each in one distinct buffer containing either 2% (v/v) Glycerol, 1% (v/v) Triton or 50 mM of L-Arginine. The turbidity (590 nm) and protein concentration were determined for each sample (total protein) prior to a centrifugation step, to

remove the insoluble protein fraction. The protein content from the supernatant was quantified, thereby allowing an estimate of the percentage of soluble protein.

		Turbidity (590 nm)	Protein quantity (µg)		Soluble protein (%) (p/p)
Step-Wise Dialysis	Glicerol 2%	0,33	Total protein	1600	0,5
			Soluble protein	8	
	Triton 100x 1%	0,26	Total protein	1300	0,76
			Soluble protein	10	
	L-Arginine	0,28	Total protein	1350	0,37
			Soluble protein	5	

Table 9: Effect of different refolding adjuvants on the level of soluble protein obtained (after centrifugation) and on the level of turbidity of the protein sample (as read by absorbance at 590 nm). Protein quantified as a percentage of total protein (before centrifugation). The protein samples are the result of the induced expression of an *E.coli* containing the expression plasmid pET151BLAD173his and the posterior protein extraction and purification using the Denaturing procedure. The soluble protein content is the supernatant part obtained through the centrifugation at 5000 *g* for 10 minutes of the total protein content. All protein quantifications were obtained through the Standard Bradford method.

Comparing these different adjuvants (Table 9) none proved to be an effective refolding adjuvant for the BLAD polypeptide. Oddly, in some cases, the presence of adjuvants (*e.g.* L-Arginine) worsened the level of insoluble protein. Broadly speaking, we were unable to devise a refolding methodology for our target protein. This inability to obtain a refolding protocol for the insolubly expressed BLAD polypeptide led us to try maximizing the soluble expression.

2. Techniques for maximizing soluble protein production

2.1. Low Growth Temperatures

One option to control inclusion bodies production is by decreasing *E.coli* growth temperatures. Low growth temperatures act in two ways on inclusion bodies: i) strongly decrease the level of T 7 promoter expression thereby diminishing the level of expressed protein, which in turn reduces aggregation and inclusion body formation; ii) have an effect on protein stability and correct folding patterns, due to the fact that hydrophobic interactions determining inclusion body formation are temperature dependent (Sahdev *et al.,* 2008).

Therefore, after growing the BL21DE3 *E.coli*, containing the plasmid pET151BLAD173his, it was induced at two temperatures, 25º C and 30 ºC. After six hours the soluble protein was extracted and purified, according to the Native Protocol. Afterword the insoluble pellet from this extraction was further extracted and purified using the Denaturing protocol, to obtain the insoluble protein fraction.

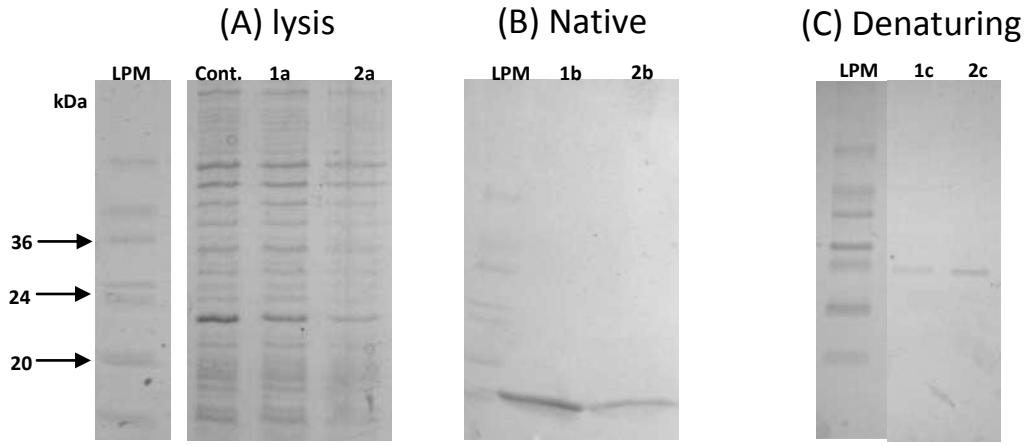


Figure 16: SDS-PAGE gels (17.5 % of Acrylamide, Coomassie R-250 stained) from the polypeptide content of the Lysis (A), and subsequent native purification protocol (B) and Denaturing purification protocol (C) of three *E.coli* containing the plasmid pET151BLAD173his expressed for six hours at three different temperatures: 25ºC, 30ºC and 37ºC (control). Lane 1a – 10 µL cell total cell lysis of *E.coli* grown at 25 ºC; Lane 2a– 10 µL cell total cell lysis of *E.coli* grown at 25 ºC; Lane Cont. – 10 µL total cell lysis of *E.coli* grown at 37 ºC as control; Lane 1b – eluted fraction from the Native extraction protocol from *E.coli* grown at 25 ºC; Lane 2b – eluted fraction from the Native extraction protocol from *E.coli* grown at 30 ºC; Lane 1c – eluted fraction from the Denaturing extraction protocol from *E.coli* grown at 25 ºC; Lane 2c – eluted fraction from the Denaturing extraction protocol from *E.coli* grown at 30 ºC; Lane "LPM" refers to the Low protein molecular marker.

Analyzing the temperature effect (Figure 16, A, B and C), we see that by lowering the induction temperature, there is a sharp decrease in the level of protein expression. The *E.coli* grown at 37 ºC (Lane "Cont.") had a higher level of protein expression than the others grown at 25 ºC (Lane 1a) and 30 ºC (Lane 2a). However, lower expression level had no effect on the level of soluble protein (Figure 16, C). As in control temperature (37 ºC), we were only able to isolate the BLAD polypeptide on the insoluble fraction of the Native procedure and using high urea concentrations (Denaturing protocol) (Figure 16 C).

Considering this, we attempted to lower the growth temperature even further, using the specific pCold I plasmid. The pCold I vectors or Cold-shock expression vectors, are specifically designed to perform efficient protein expression at very low temperatures, 15 ºC, using a promoter derived from cspA gene, one of the cold-shock genes. By lowering the induction temperature to 15 ºC, the promoter becomes highly stable resulting in improved translation efficiency, allowing for efficient expression at low temperatures (15 ºC) (Mitta *et al.*, 1997; Xia *et al.,* 2001).

Therefore we induced an *E.coli* containing the expression plasmid pColdIBLAD173his and allowed it to grow at 15 ºC. After six hours we extracted and purified the soluble protein, according with the Native Protocol. The insoluble pellet from this extraction was extracted and purified using the Denaturing protocol, to obtain the insoluble protein fraction. Again, similar to the pET151 plasmid due to the different tags associated with the plasmid (his-tag, TEV protease) the BLAD polypeptide gains around 4 kDa, therefore the real size of our expressed target protein is in the vicinity of 24 kDa.

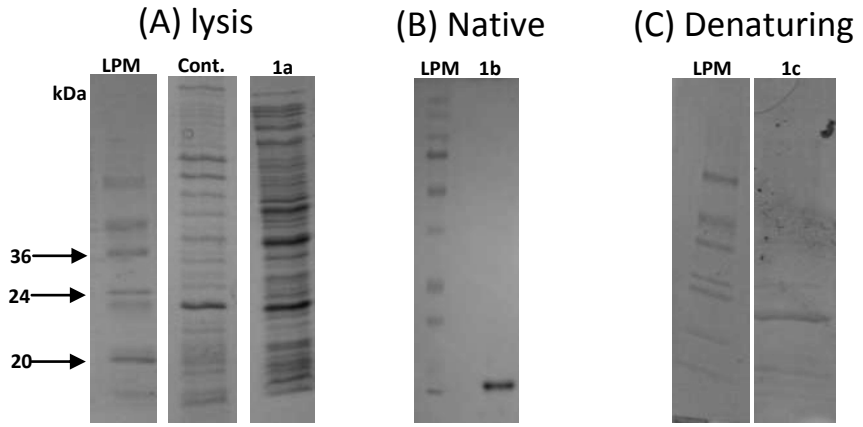


Figure 17: SDS-PAGE gels (17,5 % of Acrylamide, Coomassie R-250 stained) from the polypeptide content of the Lysis (A), and subsequent native purification protocol (B) and Denaturing purification protocol (C) of an *E.coli* containing the plasmid pColdIBLAD173his expressed for six hours at two different temperatures: 37 ºC and 15 ºC (control). Lane 1 – 10 µL cell total cell lysis of *E.coli* grown at 15 ºC; Lane Cont. – 10 µL total cell lysis of *E.coli* grown at 37 ºC as control; Lane 1a – eluted fraction from the Native extraction protocol from *E.coli* grown at 15 ºC; Lane 1b – eluted fraction from the Native extraction protocol from *E.coli* grown at 15 ºC; Lane 1c – eluted fraction from the Denaturing extraction protocol from *E.coli* grown at 15 ºC; Lane "LPM" refers to the Low protein molecular marker.

Similar to previous results, lowering the culture growth temperature to 15 ºC had no effect on the expression of soluble BLAD, (Figure 17) we were only able to extract the BLAD polypeptide by using the Denaturing protocol (Figure 17, C). Therefore, the soluble expression of the BLAD polypeptide is independent of temperatures, at least the range that we used in this study. Furthermore, up to this point the BLAD polypeptide proved impossible to be either soluble expressed in *E.coli*, or correctly refolded upon urea removal. Taking into consideration all these unsuccessful results we began to wonder if this inability to obtain soluble BLAD might be the result of this polypeptide specific nature, and if that nature could be altered to improve solubility.

3. Protein engineering to improve protein solubility

Several proteins, relatively similar to BLAD, have been successfully over-expressed in *E.coli* (e.g. β-conglycinin, Ara h1, adzuki bean β-conglycinin) (Maruyama *et al.*, 1998; Fukuda *et al.*, 2008; Cabanos *et al.*, 2010). So, what are the distinct characteristics of the BLAD polypeptide that make it insoluble?

BLAD is a 20 kDa polypeptide, a subunit of a larger heteroligomer, and is therefore inexistent in nature as an isolated entity (Monteiro *et al.*, 2010). Can this polypeptide, by itself, independent of *E.coli*, due to hydrophobic character of the primary structure and/or self-aggregation became insoluble upon expression?

A cursory look at soluble and insoluble proteins, reveals that its primary sequence is the most important determinant factor of protein solubility (Idicula-Thomas *et al.*, 2006), specifically, the presence of hydrophobic residues their number and their location in the proteins structure. (Luan *et al.,* 2004). As a rule, a protein tends to "hide" its hydrophobic residues, by folding, to minimize its interaction with water. Obviously, the higher the number of these residues the harder it is to conceal them and the more hydrophobic the protein (Agostini *et al.*, 2012). One way to analyze the hydrophobic character of a given protein is to have a plot depicting the distribution of polar and apolar residues along a protein given sequence. A thoroughly insoluble protein will have large areas of its structure categorized as hydrophobic, the opposite for soluble proteins (Idicula-Thomas and Balaji, 2005). The most commonly used software for this type of analysis is the Kylie-Doolittle plot in which each amino acid is given a hydrophobicity score between 4.6 and -4.6. A window size is the number of amino acids whose hydrophobicity scores will be averaged and assigned to the first amino acid in the window. The y axis represents the hydrophobicity scores and the x axis represents the window number, regions with values above 0 are hydrophobic in character (Kyte and Doolittle, 1982). Using the Kylie-Doolittle software we compared our structure hydrophobicity with two more similar proteins that had already been over-expressed in a soluble form in *E.coli*: β-conglycinin - BLAD similarity 56% (accession number: t1 UIK_A) and Ara h1 protein - BLAD similarity 55% (accession number: 3SMH_A).

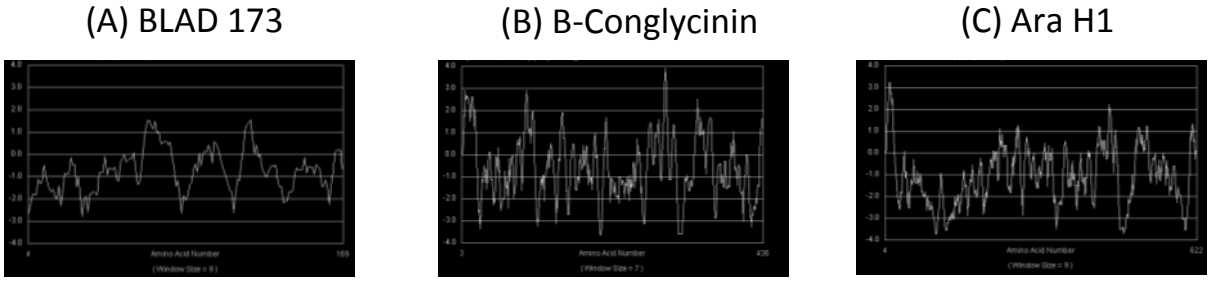


Figure 198: Kyte-Doolittle plots depicting the hydrophobic character of the BLAD polypeptide (A) and two proteins similar to BLAD but already soluble over-expressed in *E.coli*: B-conglycinin (B) and Ara H1 (C). Each amino acid in the proteins sequence was given a hydrophobicity score between 4.6 and -4.6. A window size is the number of amino acids whose hydrophobicity scores will be averaged and assigned to the first amino acid in the window. The y axis represents the hydrophobicity scores and the x axis represents the window number. Regions with values above 0 are hydrophobic in character.

Comparing these three proteins hydrophobicity plots (Figure 18) we see that BLAD, although highly insoluble upon over-expression, presents a highly soluble chart profile, possessing only two distinct hydrophobic regions in the middle of its primary sequence. On the contrary the soluble expressed β-Conglycinin and Ara H 1 have highly irregular plots with several hydrophobic patches, conveying the idea of possessing a more hydrophobic primary structure. Examining these plots we can say that BLAD's primary structure is not the key reason for its high level of insolubility upon over-expression.

Our second hypothesis for BLAD's insolubility upon expression in *E.coli*, is its innate aggregation. Globulins are known to self-aggregate, forming very large molecular mass

aggregates, probably to facilitate packaging in storage vacuoles (Freitas *et al*, 2000). The polypeptide BLAD being the subunit of a seed storage globulin (Monteiro *et al.*, 2010), might be structurally pre-disposed, when over-expressed, to self-aggregate, and in consequence become insoluble. To study this innate characteristic, we began by computing BLAD polypeptide predictive tri-dimensional structure using the PHYRE 2 software (Kelley and Sternberg, 2009), and attempted to determine specific sequence or folds in the structure that might induce aggregation. The PHYRE 2 software uses a library of known protein structures taken from the Structural Classification of Proteins and the Protein Data Bank (PDB). A user-submitted primary sequence, or 'query', is scanned against these databases, and a predictive tri-dimensional structure is obtained (Kelley and Sternberg, 2009).

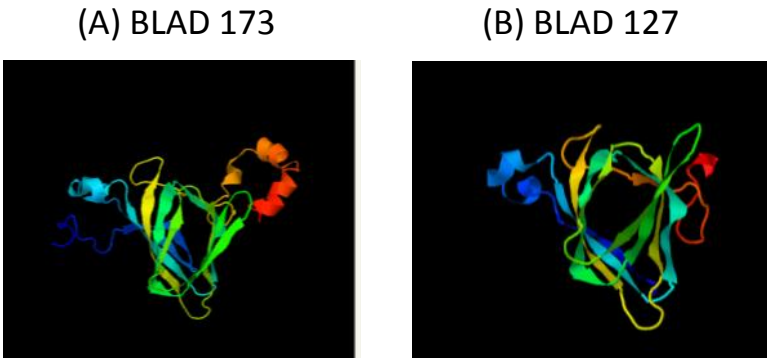


Figure 19: Image of the predicted tri-dimensional structures, according with the PHYRE 2 software, of the original BLAD 173 aminoacids (A); and a BLAD designed for higher solubility lacking part of its native N and C terminal sequence (B). Structure is oriented from the N terminal (Blue) to the C terminal (Red). Prediction performed by the PHYRE 2 software with confidence 100 % and 56 % Identity and based Chain A crystal structure of soybean beta-conglycinin homotrimer (NCBI PDB ID: 1IPJ).

Looking at BLAD 173 predicted tri-dimensional structure (Figure 19, A), and after conversations with Dr. Isabel Bento (Structural Genomics Macromolecular Crystallography Unit" located at ITQB), we hypothesized if the C-terminal region with its two alpha helix could be able to connect with the N-Terminal region from another BLAD polypeptide, generating an endless chain of interlinked BLAD monomers and therefore being this the cause for aggregation, regardless of inclusion bodies formation. Previous works with β-conglycinin from soya, specifically by Maruyiama and colleagues, showed that different variants of the β-conglycinin polypeptide, with shortened N and C extremities, exhibited highly different solubility patterns when compared with its native counterpart, proving that these regions might be relevant for solubility (Maruyama *et al.*, 1999; Prak *et al.,* 2007; Tandang-Silvas *et al.*, 2011). Consequently, we designed a new BLAD gene with shortened N and C extremities, BLAD 127 (Annex III.1, for gene sequence), and placed it into the pET151plasmid, generating the plasmid pET151BLAD127his. This *E.coli* was grown at 37 ºC and after six hours post-induction we extracted and purified the soluble protein, according to the Native Protocol. The insoluble pellet from this extraction was extracted and purified using the Denaturing protocol, to obtain the insoluble protein fraction. Bear in mind that BLAD 127 when expressed in *E.coli*, due to the different tags associated with the plasmid

(his-tag, TEV protease), gains around 4 kDa. Therefore the real size of our expressed target protein is in the vicinity of 19 kDa.

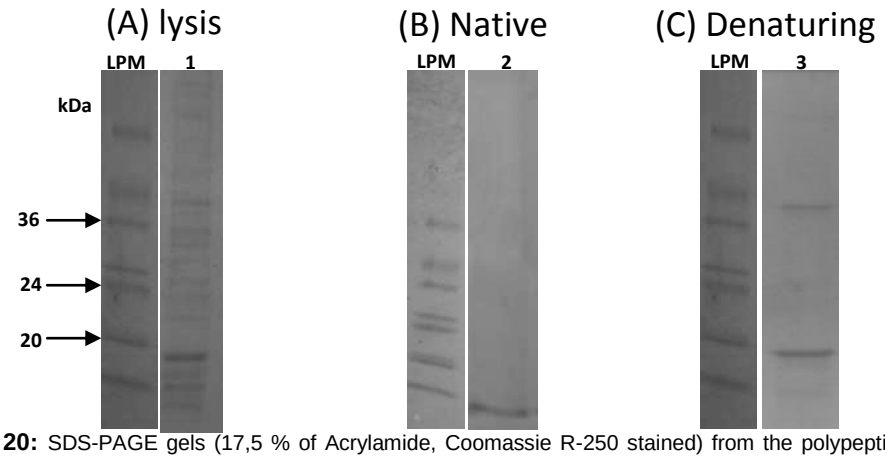


Figure 20: SDS-PAGE gels (17,5 % of Acrylamide, Coomassie R-250 stained) from the polypeptide content of the Lysis (A), and subsequent native purification protocol (B) and Denaturing purification protocol (C) of an *E.coli* containing the plasmid pET151BLAD127his expressed for six hours at 37ºC. Lane 1 – 10 µL cell total cell lysis of *E.coli* grown at 37 ºC; Lane 2 – eluted fraction from the Native extraction protocol from *E.coli* grown at 37 ºC; Lane 3 – eluted fraction from the Denaturing extraction protocol from *E.coli* grown at 37 ºC. Lane "LPM" refers to the Low protein molecular marker.

Examining Figure 20, above, we see that resembling the behavior of BLAD 173 polypeptide, BLAD127 is only successfully extracted from *E.coli* using the Denaturing extraction and purification protocol. Evidently BLAD's aggregation plays no role in this polypeptides soluble expression in *E.coli*. Therefore, neither BLAD's primary structure, nor its possible aggregation, plays a role on the soluble expression in *E.coli*. Clearly BLAD polypeptide, independently of the size, is insolubly expressed in *E.coli* in the form of inclusion bodies, and further expression optimizations must follow. But can the shorter version of BLAD (BLAD127), by being less aggregation prone, improve our refolding of BLAD's inclusion bodies?

3.1. Protein engineering and refolding of inclusion bodies

To understand if BLAD 127 might improve the protein refolding rate, we carried out an experiment, and compared the refolding percentage of two different BLAD constructs, 127 and 173, using some of the already tested refolding protocols, and an alternate version of the Dilution/ Step-Wise dialysis procedure. This alternate version was developed with the idea that a lower protein concentration might result in lower aggregation, thus higher levels of soluble protein. After the two distinct BLAD polypeptides, BLAD 127 and BLAD 173 (full gene), were over-expressed and extracted/purified from *E.coli* using the Denaturing protocol, in the case of the alternate Dilution/ Step-Wise dialysis procedure each sample was diluted tenfold, using the standard denaturing elution buffer with 6 M urea, and immediately after, Step-wise dialyzed against a buffer with ever decreasing urea concentrations (4, 3, 2, 1, 0 M of urea), while maintaining the other constituents of the buffer. For the remaining methods the procedure followed was the same. After dialysis the samples from the dilution/step wise protocol were concentrated (10x) to the original volume. The turbidity (590 nm) and protein concentration were determined for each sample (total protein) prior to a centrifugation step, to remove the insoluble

protein fraction. The protein content from the supernatant was quantified, thereby allowing to estimate the percentage of soluble protein.

		Turbidity (590 nm)	Protein quantity (µg)		Soluble protein (%) (p/p)
One-Step	BLAD 127	0,034	Total protein	580	14 %
			Soluble protein	82,5	
	BLAD 173	0,145	Total protein	955	1 %
			Soluble protein	10	
Dilution/Step-Wise Dialysis	BLAD 127	0,024	Total protein	495	25 %
			Soluble protein	122	
	BLAD 173	0,157	Total protein	1195	0,76 %
			Soluble protein	9	
Alternate Dilution/Step-Wise Dialysis	BLAD 127	0,026	Total protein	1160	38 %
			Soluble protein	445	
	BLAD 173	0,074	Total protein	1550	9 %
			Soluble protein	135	

Table 10: Effect of different refolding procedures on the amount of soluble protein obtained (after centrifugation) and on the level of turbidity of the protein sample (as read by absorbance at 590 nm). Protein was quantified as a percentage of total protein (before centrifugation). The protein samples are the result of the induced expression of an *E.coli* containing the expression plasmid pET151BLAD127his and the pET151BLAD173. Posterior protein extraction and purification was using the Denaturing procedure. The soluble protein content is the supernatant part obtained through the centrifugation at 5000 *g* for 10 minutes of the total protein content. All protein quantifications were obtained through the Standard Bradford method.

Comparing the two BLAD constructs, it is clear that the less aggregation-prone BLAD 127 is more easily refolded, with percentages of refolded protein in-between 14 and 38 % (Table 10). Evidently, the presence of the N and C- Terminal regions, affects its ability to refold back to its native form, and is critical in obtaining workable amounts of soluble BLAD. Focusing on the refolding methodologies alone, the alternate protocol of dilution/ step-wise dialysis is the better refolding methodology producing the higher amounts of soluble BLAD. For example, for the highly insoluble BLAD 173 construct, we obtained a value of 9 % soluble protein, an improvement, when compared with the < 1 % values obtained with the previous methods. Therefore, protein concentration, aggregation propensity and a slow removal of urea are determinant factors in a successful refolding procedure for the BLAD polypeptide.

In all successful refolding methodologies, there is the possibility that the target proteins, although soluble, might be inactive, due to incorrect folding patterns (Sahdev *et al.*, 2008). Consequently, to understand whether our new refolding procedure was adequate to refold the BLAD polypeptide back to its native form, we tested one of its biological activities. The BLAD polypeptide is a stable breakdown product of *Lupinus* β-conglutin with a proven lectin-like activity. In the work conducted by Ramos *et al.*, (1997), it was shown that the BLAD 20 kDa subunit strongly recognizes non-specific anti-Ubiquitin antibodies, therefore exhibiting a lectin-like activity. To verify the BLAD polypeptide correct refolding, in this alternate Dilution/Dialysis protocol, we subjected the BLAD 127, with the higher amount of soluble protein, to a

immunoblot. The polypeptide was immobilized in a PVDF membrane and probed alternately with, anti-BLAD antibodies (specific for the BLAD polypeptide) and with a non-specific rabbit anti-ubiquitin antibodies.

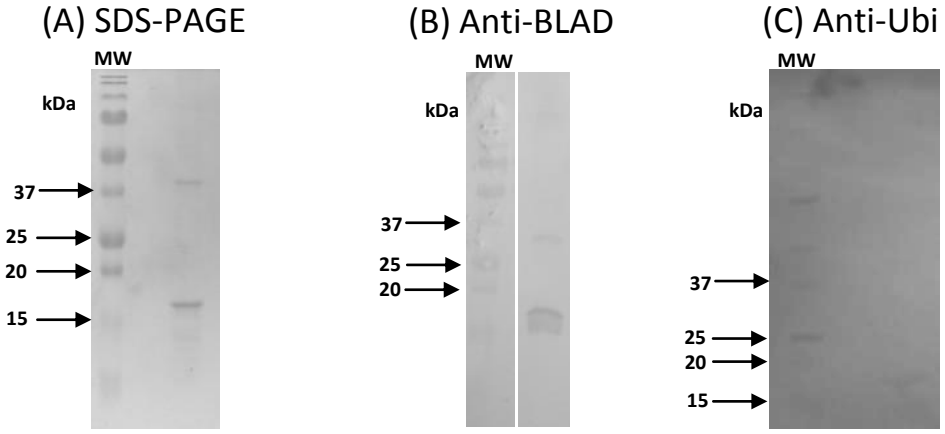


Figure 21: SDS-PAGE gel (17,5 % of Acrylamide, Coomassie R-250 stained) from 100 µl the polypeptide content of the Denaturing extraction and purification of BLAD 127 (A), and subsequent immunoblot with specific BLAD antibodies (anti-BLAD) (B) and immunoblot with specific ubiquitin antibodies (anti-Ubi) (C). The *E.coli* contained the plasmid pET151BLAD127his expressed for six hours at 37ºC. Lane 1 – 10 µL cell total cell lysis of *E.coli* grown at 25 ºC; Lane 2 – 10 µL cell total cell lysis of *E.coli* grown at 25 ºC; Lane Cont. – 10 µL total cell lysis of *E.coli* grown at 37 ºC as control; Lane 3 – eluted fraction from the Native extraction protocol from *E.coli* grown at 25 ºC; Lane 4 – eluted fraction from the Native extraction protocol from *E.coli* grown at 30 ºC; Lane 5 – eluted fraction from the Denaturing extraction protocol from *E.coli* grown at 25 ºC; Lane 6 – eluted fraction from the Denaturing extraction protocol from *E.coli* grown at 30 ºC; Lane "MW" refers to the protein molecular marker (Biorad).

As expected BLAD 127 polypeptide is strongly detected in the immunoblot procedure when probed with specific BLAD antibodies (Figure 21, B). When probed with the non-specific antibodies, Anti-Ubi (Figure 21, C), the refolded BLAD 127 displayed a distinct lectin-like behavior, interacting with the carbohydrate compounds present in the glycosyilated antibodies, and therefore generating in the immunoblot a visible polypeptide band under the 20 kDa marker. This specific activity resembles the one found in the native form of the BLAD oligomer (Ramos *et al,* 1997), although less intense. Thus this refolding methodology (alternate dilution/Step-dialysis) does not affect BLAD's 127 lectin-like activity, and therefore constitutes a valid method for the refolding of our target protein in its active form.

Conclusions

From the beginning we had one clear defined objective to develop a working protocol for the heterologous expression of the BLAD polypeptide in *E.coli*. Although apparently a simple task, BLAD, unlike similar proteins, proved to be insolubly expressed in *E.coli* in the form of inclusion bodies. We approached this problem by using two different methodologies, namely: i) using the inclusion bodies as a rather effective purification step and refold them back to its native form or ii) by maximizing the soluble protein expression, eliminating inclusion body formation. First we opted to refold BLAD's inclusion bodies, and tested several procedures, one step dialysis, Step-Wise dialysis, Dilution/stepwise and adjuvant assisted Step-Wise dialysis, but none was able to refold BLAD into its native and soluble form. Considering this, we changed our approach and attempted to increase the level of soluble protein expression in *E.coli* by lowering growth temperatures. We tested 30º C, 25 ºC and 15 ºC, the last one in a special low temperature expression plasmid, pCold I. However, even at temperature as low as 15 ºC, there was no effect on the level of soluble protein expression. Given these different, yet negative results, we focused on BLAD´s intrinsic characteristics that could be responsible for its insolubility, namely its hydrophobicity and possible propensity for self-aggregation. In terms of hydrophobicity BLAD has a highly hydrophilic primary structure, even more, than the closely related proteins (β-conglycinin, Ara h1) already expressed in *E.coli*. Consequently BLAD's hydrophobicity is unrelated with its insolubility. Regarding BLAD aggregation we postulated that BLAD N and C terminal extensions played a role in BLAD aggregation, to test this a shorter version of the BLAD polypeptide was constructed, BLAD 127. This construct continued to be insolubly expressed in *E.coli*, proving that BLAD's aggregation is irrelevant for the soluble expression in *E.coli*. In another perspective, we tested whether this BLAD 127 was more easily refolded with some of the already tested refolding procedures. At the same time we used an alternate version of the Dilution/Step-wise dialysis, which focuses on protein dilution, thereby diminishing protein interactions and aggregation. BLAD 127 proved to be easier to refold in any of the methodologies, achieving refolding percentages as high as 38 % (w/w), while maintaining BLAD lectin-like activity, therefore probably possessing a correct folding pattern. This improvement allowed for the production of higher amounts of soluble BLAD, probably with the correct fold, however it is still, very low amounts, only satisfactory for a laboratory scale and useless for any future industrial application. Considering that there is no reason why BLAD cannot be soluble expressed in *E.coli*, our present efforts are focused in maximizing BLAD's soluble expression. To achieve this we have set up collaboration with the company "Hitag", which specializes in protein expression in *E.coli*. With the help of Hitag we intend to test a combination of 120 variables for maximizing the level of soluble BLAD expression in *E.coli*, thereby bypassing inclusion body and refolding problems. Variables will include up to three fusion proteins, the co-expression of foldases and chaperones, four growth temperatures and the optimization of native extraction buffers.

Bibliography

Agostini, F., Vendruscolo, M., Tartaglia, G.G. (2012). Sequence-based prediction of protein solubility. *Journal of Molecular Biology.* **421(2-3):** 237-41.

Arene, L., Pellegrino, C., Gudin, S. (1993). A comparison of the somaclonal variation level of *Rosa hybrid* L. cv. *Meirutal* plants regenerated from callus or direct induction from different vegetative and embryonic tissues. *Euphytica.* **71:** 83–90.

Balls, A.K., Hale, W.S., Harris, T.H. (1942). A crystalline protein obtained from a lipoprotein of wheat flour. *Cereal Chemistry Journal*.**19:** 279–288.

Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*.Current *Opinion in Biotechnology.* **10:**410-421.

Baneyx, F., Palumbo, J. L. (2003). Improving Heterologous Protein Folding via Molecular Chaperone and Foldase Co-Expression. *Methods in Molecular Biology.* **205:** 171-197.

Bao, Y., Liu, G., Shi, X., Xing, W., Ning, G., Liu, J., Bao, M. (2012). Primary and repetitive secondary somatic embryogenesis in *Rosa hybrida* 'Samantha'. *Plant Cell, Tissue and Organ Culture*. **109(3):** 411-418

Bensadoun, A., Weinstein, D. (1976). Assay of proteins in the presence of interfering materials. *Analytical Biochemistry*. **70:** 241-250.

Bischof**,** J.C**.**, He**,** X. (2005). Thermal stability of proteins. *Annals of the New York Academy of Sciences*. **1066:** 12-33.

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

Broekaert, W., Cammue, B., De Bolle, M., Thevissen, K., De Samblanx, G., Osborn, R. (1997). Antimicrobial peptides from plants. *Critical Reviews in Plant Sciences*.**16(3):** 297-323.

Brown, D. C. W., Thorpe, T. A. (1995). Crop improvement through tissue culture. *World Journal of Microbiology and Biotechnology.* **11(4):** 409-415.

Brown, G. G. (2011). Molecular Cloning - Selected Applications in Medicine and Biology. *Rijeka, Croatia: InTech*.

Cabanos, C., Urabe, H., Masuda, T., Tandang-Silvas, M.R., Utsumi, S., Mikami, B., Maruyama, N. (2010). Crystallization and preliminary X-ray analysis of the major peanut allergen Ara h 1 core region. *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications.* **66(Pt 9):** 1071-1073.

Cabanos, C., Urabe, H., Tandang-Silvas, M.R., Utsumi, S., Mikami, B., Maruyama, N. (2011). Crystal structure of the major peanut allergen Ara h 1. *Molecular Immunology*. **49(1-2):** 115-23.

Cabrera, E., García-Granados, A. (1981). Fitoquimica de Thyme-leaces (III): componentes cumarinicos y flavonicos en hojas de Daphne gnidium L. *Anales de Quimica*. **77:** 31–34.

Cândido, E.S., Porto, W.F., Amaro, D.S., Viana, J.C., Dias, S.C., Franco, O.L. (2011). Structural and functional insights into plant bactericidal peptides. In: Méndez-Vilas A, editor. Science against microbial pathogens: communicating current research and technological advances. Formatex, 951–960.

Cantón, E., Espinel-Ingroff, A., Pemán. J. (2009). Trends in antifungal susceptibility testing using CLSI reference and commercial methods. *Expert Review of Anti-Infective Therapy*. **7(1):** 107-119.

Ceasar, S. A., Ignacimuthu, S. (2012). Genetic engineering of crop plants for fungal resistance: role of antifungal genes. *Biotechnology Letters*. **34(6):** 995-1002.

Chandrashekhara, Niranjan-Raj, S., Deepak S., Manjunath, G., Shekar Shetty, H. (2010). Thionins (PR protein-13) mediate pearl millet downy mildew disease resistance. *Archives of Phytopathology and Plant Protection*. **43(14):** 1356-1366.

Chevallet, M., Luche, S., Rabilloud, T. (2006). Silver staining of proteins in polyacrylamide gels. *Nature protocols*. **1(4):** 1852-1858.

Chow, M. K., Amin, A. A., Fulton, K. F., Fernando, T., Kamau, L., Batty, C., Louca, M., Ho, S., Whisstock, J. C., Bottomley, S. P., Buckle, A. M. (2006). The REFOLD database: a tool for the optimization of protein expression and refolding. *Nucleic Acids Research*. 34 (Database issue): D207-12.

CLSI. (2002). Performance Standards for Antimicrobial Disk and Dilution Susceptibility Tests for Bacteria Isolated from Animals; Clinical and Laboratory Standards Institute proved Standard—Second Edition.CLSI document M31-A2.

CLSI. (2008a). Reference Method for Broth Dilution Antifungal Susceptibility Testing of Yeasts. Clinical and Laboratory Standards Institute aproved Standard—Third Edition. CLSI document M27-A3.

CLSI. (2008b). Reference Method for Broth Dilution Antifungal Susceptibility of Filamentous Fungi. Clinical and Laboratory Standards Institute aproved Standard—Second Edition.CLSI document M38-A2.

Cutler, S., Cutler, H. (1999). Biologically Active Natural Products: Pharmaceuticals. England: CRC Press Edition.

Condliffe, P. C., Davey, M. R., Power, J. B., Koehorst-van Putten, H., Visser, P. B. (2003). An Optimised Protocol for Rose Transformation Applicable to Different Cultivars. *ISHS Acta Horticulturae.* **612:** 115-120

Colgrave, M. L., Craik, D. J. (2004). Thermal, chemical, and enzymatic stability of the cyclotide kalata B1: the importance of the cyclic cystine knot. *Biochemistry*. **43(20):** 5965-75.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research.* **16(22):** 10881-10890.

Costa, F., Carvalho, I. F., Montelaro, R. C., Gomes, P., Martins, M. C. (2011). Covalent immobilization of antimicrobial peptides (AMPs) onto biomaterial surfaces. *Acta Biomaterialia*. **7(4):**1431-40.

Cottiglia, F., Loy, G., Garau, D., Floris, C., Casu, M., Pompei, R., Bonsignore, L. (2001). Antimicrobial evaluation of coumarins and flavonoids from the stems of Daphne gnidium L. *Phytomedicine*. **8(4):**302-5.

Cottiglia, F., Bonsignore, L., Loy, G., Garau, D., Floris, C., Casu, M. (2002). Structure elucidation and antibacterial activity of a new coumarinolignoid from *Daphne gnidium* L. *Magnetic Resonance in Chemistry***. 40(8):** 551–553.

Cottiglia, F., Lucca, A. J., De, A., Cleveland, T. E., Wedge, D. E. (2005). Plant-derived antifungal proteins and peptides, *Canadian Journal of Microbiology*. **1014(2):** 1001–1014.

Cotton, C. M. (1996). Ethnobotany: principles and applications. *Chichester*, UK: John Wiley.

Cowan, M. M. (1999). Plant products as antimicrobial agents. *Clinical Microbiology Reviews*. **12(4):** 564.

Curtis, M., Grossniklaus, U. (2003). A Gateway™ cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiology*. **133(2):** 462–469.

Deiana, M., Rosa, A., Casu, V., Cottiglia, F., Bonsignore, L., Dessì, M. A. (2003). Chemical Composition and Antioxidant Activity of Extracts from *Daphne gnidium* L. *Journal of the American Oil Chemists' Society.* **80 (1):** 65-70.

De La Riva, G. A., González-Cabrera, J., Vázquez-Padrón, R., Ayra-Pardo, C. (1998). *Agrobacterium tumefaciens*: a natural tool for plant transformation. *Electronic Journal of Biotechnology*. **1(3):** 24-25.

De Lucca, A. J., Cleveland, T. E., Wedge, D. E. (2005). Plant-derived antifungal proteins and peptides. *Canadian Journal of Microbiology*. **51(12):** 1001-1014.

De Wit, J. C., Esendam, H. F., Honkanen, J. J., Tuominen, U. (1990). Somatic embryogenesis and regeneration of flowering plants in rose. *Plant Cell Reports*. **9:** 456-458.

Dohm, A., Ludwig, C., Schilling, D., Debener, T. (2001). Transformation of Roses with Genes for Antifungal Proteins. Rose research and cultivation International symposium; 3rd, Rose research and cultivation. *Acta Horticulturae*. **547:** 27-34.

Doon, S. (2002). Peptide and Proteins. *Royal Society of Chemistry*. **15**: 20-107.

Ramos, P. C., Ferreira, R. M., Franco, E., Teixeira, A. R. (1997). Accumulation of a lectin-like breakdown product of beta-conglutin catabolism in cotyledons of germinating *Lupinus albus L.* seeds. *Planta*. **203(1):** 26-34.

Duranti, M., Restani, P., Poniatowska, M., Cerletti, P. (1981). The seed globulins of *Lupinus albus*. *Phytochemistry.* **20(9):** 2071–2075.

Eijsink, V. G., Bjørk, A., Gåseidnes, S., Sirevåg, R., Synstad, B., van den Burg, B., Vriend, G. (2004). Rational engineering of enzyme stability. *Journal of Biotechnology*. **113(1-3):** 105-20.

Epand, R. M., Vogel, H. J. (1999). Diversity of antimicrobial peptides and their mechanisms of action. *Biochimica et Biophysica Acta*.**1462(1-2):** 11-28.

Fabricant D. S., Farnsworth N. R. (2001). The Value of Plants Used in Traditional Medicine for Drug Discovery. *Environmental Health Perspectives*. **109 (1):** 69–75.

Ferreira P. (2007). Avaliação da bioactividade de extractos de plantas das famílias Thymelaeaceae, Euphorbiaceae e Zingiberaceae em modelos microbianos e em culturas *in vitro* de células vegetais. *Graduation thesis*. Universidade do Minho.

Firoozabady E., Moy Y., Courtney-Gutterson N., Robinson C. (1994). Regeneration of transgenic rose (*Rosa hybrida*) plants from embryogenic tissue. *BioTechnology*.**12:** 609-613.

Foley, R. C.**,** Gao, L. L.**,** Spriggs, A.**,** Soo, L. Y. C.**,** Goggin, D. E.**,** Smith, P. M. C.**,** Atkins, C. A., Singh, K. B. (2011). Identification and characterisation of seed storage protein transcripts from *Lupinus Angustifolius*. *BMC Plant Biology*. **11:** 59.

Freitas, R. L., Ferreira, R. B., Teixeira, A. R. (2000). Use of a single method in the extraction of the seed storage globulins from several legume species. Application to analyze structural comparisons with the major classes of globulins. *International Journal of Food Science and Nutrition*. **51:** 342-352.

Fukuda, T., Maruyama, N., Salleh, M. R., Mikami, B., Utsumi, S. (2008). Characterization and crystallography of recombinant 7S globulins of Adzuki bean and structure-function relationships with 7S globulins of various crops. **56(11):** 4145-4153.

Gao, G. H., Liu, W., Dai, J. X., Wang, J. F., Hu, Z., Zhang, Y., Wang, D.C. (2001). Solution structure of PAFP-S: a new knottin-type antifungal peptide from the seeds of *Phytolacca americana*. *Biochemistry*. **40(37):** 10973-10978.

García-Olmedo, F., Molina, A., Alamillo, J. M., Rodríguez-Palenzuéla, P. (1998). Plant defense peptides. *Biopolymers*. **47(6):** 479-91.

Geer, L. Y., Marchler-Bauer, A., Geer, R. C., Han, L., He, J., He, S., Liu, C., Shi, W., Bryant, S.H. (2010). The NCBI BioSystems database. *Nucleic Acids Research*. **38 (D492-6).**

Georgiou, G., Valax, P. (1996). Expression of correctly folded proteins in *Escherichia coli*. *Current Opinion in Biotechnology*. **7(2):** 190-197.

Gileadi, O., Burgess-Brown, N. A., Colebrook, S. M., Berridge, G., Savitsky, P., Smee, C. E., Loppnau, P., Johansson, C., Salah, E., Pantic, N.H. (2008). High throughput production of recombinant human proteins for crystallography. *Methods in Molecular Biology*. **426:** 221-246.

Haider, S. R., Sharp, B. L., Reid, H. J. (2011). A comparison of Tris-glycine and Tris-tricine buffers for the electrophoretic separation of major serum proteins. *Journal of Separation Science*. **34(18):**2463-2467.

Hall, J. (2007). Applying Fusion Protein Technology to *E. coli*. *BioPharm International*. http://www.biopharminternational.com/biopharm/article/articleDetail.jsp?id=423193

Hall, G. S. (2012). Interactions of Yeasts, Moulds, and Antifungal Agents: How to Detect Resistance. *Humana Press*.

Hames, B. D. (1998) Gel Electrophoresis of Proteins: A Practical Approach. 3rd edition. *Oxford University Press, Springer*.

Hammerschlag, F. A. (2004). In Vitro Inhibitory Activity of Antimicrobial Peptides Cecropin, α-Thionin DB4, and γ-Thionin RsAFP1 Against Several Pathogens of Strawberry and Highbush Blueberry. *Horticultural Science.* **39 (5):** 1053-1055.

Hannig, G., Makrides, S.C. (1998). Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends in Biotechnology*.**16(2):** 54-60.

Harshberger, J. W. (1896). The Purposes of Ethno-Botany. *Botanical Gazette*. **21:** 146-158.

Hellens, R., Mullineaux, P., Klee, H. (2000) Technical Focus: a guide to *Agrobacterium* binary Ti vectors. *Trends in Plant Science*. **5(10):** 446-51.

Hernandez-Lucas, C., Carbonero, P., Garcia-Olmedo, F. (1978). Identification and purification of a purothionin homolog from rye (*Secale cereale* L.). *Journal of Agriculture and Food Chemistry.* **26(4):**794–796.

Horstmann, V., Huether, C., Jost, W., Reski, R. and Decker, E. L. (2004) Quantitative promoter analysis in *Physcomitrella patens*: a set of plant vectors activating gene expression within three orders of magnitude. *BMC Biotechnology*. **4:** 13-26.

Hsia, C., Korban, S. S. (1996). Organogenesis and somatic embryogenesis in callus cultures of *Rosa hybrida* and *Rosa chinensis minima. Plant Cell, Tissue and Organ Culture*. **44:** 1–6.

Iauk, L., Aleo, G., Caccamo, F., Rapisarda, A., Ragusa, S., Speciale, A. M. (1996). Antibacterial and antimycotic activities of *Daphne gnidium* L. extracts. *Phytotherapy Research*. **10:** 166–168.

Iauk, L., Aleo, G., Caccamo, F., Rapisarda, A., Ragusa, S., Speciale, A. M. (1997). Comparative Evaluation of Antibacterial and Antimycotic Activities of *Daphne gnidium* Leaf and Bark Extracts. *Farmaci & Terapia.***14:** 37–43.

Idicula-Thomas, S., Balaji, P. V. (2005). Understanding the relationship between the primary structure of proteins and its propensity to be soluble on overexpression in *Escherichia coli*. *Protein Science*. **14(3):** 582-92.

Idicula-Thomas, S., Kulkarni, A. J., Kulkarni, B. D., Jayaraman, V. K., Balaji, P. V. (2006). A support vector machine-based method for predicting the propensity of a protein to be soluble or to form inclusion body on overexpression in *Escherichia coli. Bioinformatics*. **22(3):** 278-84.

Kanokwiroon, K., Teanpaisan, R., Wititsuwannakul, D., Hooper, A. B., Wititsuwannakul, R. (2008). Antimicrobial activity of a protein purified from the latex of *Hevea brasiliensis* on oral microorganisms. *Mycoses*. **51(4):** 301-307.

Keil, B. (1971). *The Enzymes*. 3rd edition. Vol. III. *Academic Press*. USA.

Kelley,L. A., Sternberg, M. J. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols*. **4(3):** 363-371.

Kim, T. D., Ryu, H. J., Cho, H. I., Yang, C. H., Kim, J. (2000). Thermal behavior of proteins: heat-resistant proteins and their heat-induced secondary structural changes. *Biochemistry.* **39(48):** 14839-14846.

Kim, C. K., Chung, J. D., Park, S. H., Burrell, A. M., Kamo, K. K., Byrne, D. H. (2004). *Agrobacterium tumefaciens*-mediated transformation of *Rosa hybrida* using green fluorescent protein (GFP) gene. *Plant Cell, Tissue and Organ Culture*. **78:** 107–111.

Kintzios, S., Manos, C., Makri, O. (1999). Somatic embryogenesis from mature leaves of rose (*Rosa sp*.). *Plant Cell Reports.* **18(6):**,467-472.

Kolmar, H. (2008). Alternative binding proteins: biological activity and therapeutic potential of cystine-knot miniproteins. *FEBS Journal*. **275(11):** 2684-90.

Koo, J. C., Chun, H. J., Park, H. C., Kim, M. C., Koo, Y. D., Koo, S. C., Ok, H. M., Park, S. J., Lee, S. H., Yun, D. J., Lim, C. O., Bahk, J. D., Lee, S. Y., Cho, M. J. (2002). Over-expression of a seed specific hevein-like antimicrobial peptide from *Pharbitis nil* enhances resistance to a fungal pathogen in transgenic tobacco plants. *Plant Molecular Biology*. **50(3):** 441-52.

Korban, S. S. (2006). Somatic Embryogenesis in Rose: Gene Expression and Genetic Transformation. *Somatic Embryogenesis Plant Cell Monographs.* **2:** 247-257.

Kouhei, T., Daisuke, E., Izumi, K., Tsutomu. (2003). Practical considerations in refolding from inclusion bodies. *Protein Expression and Purification*. **28(1):**1-8.

Kragh, K. M., Nielsen, J. E., Nielsen, K. K., Dreboldt, S., Mikkelsen, J. D. (1995). Characterization and localization of new antifungal cysteine-rich proteins s from *Beta vulgaris* L. *Molecular Plant Microbe Interactions*. **8:** 424-434.

Kunitake, H., Imamizo, H., Mii, M. (1993) Somatic embryogenesis and plant regeneration from immature seed-derived calli of rugosa rose (*Rosa rugosa* Thunb). *Plant Science*. **90:** 187–194.

Kyte, J., Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*. **157(1):** 105-132.

Li, X., Krasnyanski, S. F., Korban, S. S. (2002). Somatic embryogenesis, secondary somatic embryogenesis, and shoot organogenesis in Rosa. *Journal of Plant Physiology*. **159(3):** 313–319.

Li, X., Gasic, K., Cammue, B., Broekaert, W., Korban, S. (2003). Transgenic rose lines harboring an antimicrobial protein gene, *Ace-AMP1*, demonstrate enhanced resistance to powdery mildew (*Sphaerotheca pannosa*). *Planta*. **218(2):** 226-232

Loeza-Ángeles, H., Sagrero-Cisneros, E., Lara-Zárate, L., Villagómez-Gómez, E., López-Meza, J.E., Ochoa-Zarzosa, A. (2008). Thionin Thi2.1 from *Arabidopsis thaliana* expressed in endothelial cells shows antibacterial, antifungal and cytotoxic activity. *Biotechnology Letters*. **30(10):** 1713-1719.

Loeza-Ángeles, H.; López-Meza, J. E., Ochoa-Zarzosa, A. (2011). Antimicrobial effects of plant defense peptides expressed by bovine endothelial cells on intracellular pathogens. *Electronic Journal of Biotechnology*. **14(5)** (http://dx.doi.org/10.2225/vol14-issue5-fulltext-1).

Luan, C. H., Qiu, S., Finley, J. B., Carson, M., Gray, R. J., Huang, W., Johnson, D., Tsao, J., Reboul, J., Vaglio, P., Hill, D. E., Vidal, M., DeLucas, L. J., Luo, M. (2004). High-Throughput Expression of *C. elegans* Proteins. *Genome Research*. **14(10b):** 2102-2110.

Maistrello, L., López, Mª. A., Soria, F. J., Ocete, R. (2005). Growth inhibitory activity of *Daphne gnidium* L. (Thymelaeaceae) extracts on the elm leaf beetle (Col., Chrysomelidae). *Journal of Applied Entomology*. **129(8):** 418–424.

Makrides, S. C. (1996). Strategies for achieving high-level expression of genes in *Escherichia coli. Microbiology and Molecular Biology Reviews.* **60(3):** 512-538.

Manojlović, N. T., Mašković, P. Z., Vasiljević, P. J., Jelić, R. M., Jusković, M. Ž., Sovrlić, M., Mandić, L., Radojković, M. (2012). HPLC analysis, antimicrobial and antioxidant activities of *Daphne cneorum* L.. *Journal Hemijska Industrija*. **66 (5):** 709-716.

Marchant, R., Davey, M. R., Lucas, J. A. and Power, J. B. (1996). Somatic embryogenesis and plant regeneration in Floribunda rose (*Rosa hybrida L*.) cvs. Trumpeter and Glad Tidings. *Plant Science*. **120(1):** 95–105.

Marchant, R., Power, J. B., Lucas, J. A., Davey, M. R. (1998). Biolistic transformation of rose (Rosa hybrida L.). *Annals of Botany*. **81:** 109–114.

Marcos, J. F., Muñoz, A., Pérez-Payá, E., Misra, S., López-García, B. (2008). Identification and rational design of novel antimicrobial peptides for plant protection. *Annual Review of Phytopathology*. **46:**273-301.

Marcos, J. F., Gandía, M. (2009). Antimicrobial peptides: to membranes and beyond. *Expert Opinion on Drug Discovery*. **4(6):** 659-671.

Marques, R., Sousa, M. M., Oliveira, M. C., Melo, M. J. (2009). Characterization of weld (*Reseda luteola* L.) and spurge flax (*Daphne gnidium* L.) by high-performance liquid chromatography–diode array detection–mass spectrometry in Arraiolos historical textiles. *Journal of chromatography A.* **1216 (9):**1395-402.

Marr, A. K., Gooderham, W. J., Hancock, R. E. (2006). Antibacterial peptides for therapeutic use: obstacles and realistic outlook. *Current Opinion in Pharmacology*. **6(5):**468-72.

Maruyama, N., Katsube, T., Wada, Y., Oh, M. H., Barba De La Rosa, A. P., Okuda, E., Nakagawa, S., Utsumi, S. (1998). The roles of the N-linked glycans and extension regions of soybean beta-conglycinin in folding, assembly and structural features. *European Journal of Biochemistry.* **258(2):** 854-62.

Maruyama, N., Sato, R., Wada, Y., Matsumura, Y., Goto, H., Okuda, E., Nakagawa, S., Utsumi, S. (1999). Structure-physicochemical function relationships of soybean beta-conglycinin constituent subunits. *Journal of Agriculture Food Chemistry.* **47(12):** 5278-84.

Melo, T. S., Ferreira, R. B., Teixeira, A. N. (1994).The seed storage proteins from *Lupinus albus. Phytochemistry* **37**: 641–648.

Mitta, M., Fang, L, Inouye, M. (1997). Deletion analysis of cspA of *Escherichia coli*: requirement of the AT-rich UP element for cspA transcription and the downstream box in the coding region for its cold shock induction. *Molecular Microbiology*. **26(2):** 321-335.

Monteiro, S., Freitas, R., Rajasekhar, B. T., Teixeira, A. R., Ferreira, R. B. (2010). The unique biosynthetic route from lupinus beta-conglutin gene to blad. *PLoS One*. **5(1):** 8526-8542.

Montesinos, E., Bardají, E. (2008). Synthetic Antimicrobial Peptides as Agricultural Pesticides for Plant-Disease Control. *Chemistry & Biodiversity*. **5(7):** 1225–1237.

Noriega, C., and Söndahl, M. R. (1991). Somatic Embryogenesis in Hybrid Tea Roses. *Nature Biotechnology*. **9:** 991-993.

Office of the Gene and Technology Regulator. (2009). The Biology and Ecology of Rosa x hybrida (Rose). *Australian Governement, Department of Health and Ageing*. **2:** 6-20**.**

Oliveira, E., Amara, I., Bellido, D., Odena, M. A., Domínguez, E., Pagès, M., Goday, A. (2007). LC-MSMS identification of Arabidopsis thaliana heat-stable seed proteins: enriching for LEA-type proteins by acid treatment. *Journal of Mass Spectrometry*. **42(11):**1485-1495.

Palomares, L., Kuri-Brena, F., Ramirez, O. (2005). Industrial Recombinant Protein Production. Biotecnhology. *Encyclopedia of Life Support Systems*. Volume V.

Parijs, J. V., Broekaert, W. F., Goldstein, I. J., and Peumans, W. J. (1991). Hevein: an antifungal protein from rubber-tree (*Hevea brasiliensis*) latex. *Planta*. **183:**258-264.

Pastor, M. T., Esteras-Chopo, A., López de la Paz, M. (2005). Design of model systems for amyloid formation: lessons for prediction and inhibition. *Current Opinion in Structural Biology*. **15(1):** 57-63.

Peleg, Y., Unger, T. (2012). Resolving bottlenecks for recombinant protein expression in *E. coli*. *Methods in Molecular Biology.* **800:**173-86.

Pelegrini, B., Del Sarto, P., Perseghini, S., Franco, O., Luiz, O. (2011). Antibacterial peptides from plants: what they are how they probably work. *Biochemistry Research International.* **2011:** 250-349.

Pereira J. (2008). Avaliação da bioactividade de extractos vegetais de *Daphne gnidium, Euphorbia hyberna, Hedychium gardnerianum* e de extractos de propolis em modelos microbianos e em culturas *in vitro* de linho. Graduation thesis Universidade do Minho.

Peternel, S., Grdadolnik, J., Gaberc-Porekar, V., Komel, R. (2008). Engineering inclusion bodies for non denaturing extraction functional proteins. *Microbial Cell Factories*, **7:**34.

Portieles, R., Ayra, C., Borras, O. (2006). Basic insight on plant defensins. *Biotecnologia Aplicada*. **23:**75–78.

Powers, J. P., Hancock, R. E. (2003)**.** The relationship between peptide structure and antibacterial activity. *Peptides.* **24(11):**1681-91.

Prak, K., Nakatani K., Maruyama, N., Utsumi, S. (2007). C-terminus engineering of soybean proglycinin: improvement of emulsifying properties. *Protein Engineering, Design & Selection*. **20(9):** 433-442.

Pratap, K., Siba, R., Madhu, S., Sood, A., Paramvir, A. (2006). In vitro propagation of rose: a review. *Biotechnology Advances.* **24(1):** 94-114.

Rahnamaeian, M. (2011). Antimicrobial peptides: modes of mechanism, modulation of defense responses. *Plant Signaling & Behaviour*. **6(9):**1325-32.

Ramagli, L. S., Rodriguez, L. V. (1985). Quantitation of microgram amounts of protein in two-dimensional polyacrylamide gel electrophoresis sample buffer. *Electrophoresis*. **6(11):** 559–563.

Rivier, L., Bruhn, J. (1979). Editorial. *Journal of Ethnopharmacology***. 1:** 1

Roccotiello, E., Casazza, G., Galli, L., Cornara, L., Moncalvo, A., Minuto, L. (2009). The flower biology of Daphne gnidium L. (Thymelaeaceae). *Plant Systematics and Evolution*. **279 (1-4):** 41-49.

Ramos P., Ferreira R., Franco E., Teixeira A. (1997). Accumulation of a lectin-like breakdown product of β-conglutin catabolism in cotyledons of germinating *Lupinus albus* L. seeds. *Planta.* **203:**26-34.

Rogozhin, E. A., Ryazantsev, D. Y., Grishin, E. V., Egorov, T. A., Zavriev, S. K. (2012). Defense peptides from barnyard grass (*Echinochloa crusgalli* L.) seeds. *Peptides*. **38(1):**33-40.

Roodveldt, C., Aharoni, A., Tawfik, D. S. (2005). Directed evolution of proteins for heterologous expression and stability. *Current Opinion in Structural Biology.* **15(1):** 50-6.

Rout, G. R., Debata, B. K., Das, P. (1991). Somatic embryogenesis in callus culture of *Rosa hybrida* L. cv. Landora. *Plant Cell, Tissue and Organ Culture*. **27:** 65–69.

Sagaram, U. S., Pandurangi, R., Kaur, J., Smith, T. J., Shah, D. M. (2011). Structure Activity Determinants in Antifungal Plant Defensins MsDef1 and MtDef4 with Different Modes of Action against *Fusarium graminearum. Plos One*. **6(4):** 18446-18550.

Sahdev, S., Khattar, S. K., Saini, K. S. (2008). Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and Cellular Biochemistry*. **307(1-2):** 249-64.

Sarasan, V., Roberts, A. V., Rout G. R. (2001). Methyl laurate and 6-benzyladenine promote the germination of somatic embryos of a hybridrose. *Plant Cell Reports*. **20:** 183–186.

Schägger, H. (2006). Tricine-SDS-PAGE. *Nature protocols*.**1(1):**16-22.

Schmidt, F.R. (2004). Recombinant expression systems in the pharmaceutical industry. *Applied Microbiology and Biotechnology.* **65(4):** 363-72.

Selitrennikoff, C. P. (2001). Antifungal Proteins. *Applied and Environmental Microbiology*. **67(7):** 2883.

Shao, F., Hu, Z., Xiong, Y. M., Huang, Q. Z., Wang, C. G., Zhu, R.H., Wang, D. C. (1999). A new antifungal peptide from the seeds of *Phytolacca americana*: characterization, amino acid sequence and cDNA cloning. *Biochimica et Biophysica* Acta. **1430(2):** 262-8.

R Shukurov, R., D Voblikova, V., Nikonorova, A. K., Komakhin, R. A., V Komakhina, V., A Egorov, T., V Grishin, E., V Babakov, A. (2012). Transformation of tobacco and Arabidopsis plants with Stellaria media genes encoding novel hevein-like peptides increases their resistance to fungal pathogens. *Transgenic Research*. **21(2):**313-25.

Singh, S. M., Panda, A.K. (2005). Solubilization and refolding of bacterial inclusion body proteins. *Journal of Bioscience and Bioengineering*. **99(4):** 303-310.

Stec, B. (2006). Plant thionins – the structural perspective. *Cellular and Molecular Life Sciences*. **63(12):**1370-1385.

Takara. (2008). Cold Shock Expression System, user manual. Cat. 3360-3364.

Tam, J. P., Lu, Y. A., Yang, J. L., Chiu, K. W. (1999). An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. *Biochemistry*. **96(16):** 8913–8918.

Tanaka, Y., Katsumoto, Y., Brugliera, F. and Mason, J. (2005). Genetic engineering in floriculture. *Plant Cell, Tissue and Organ Culture*. **80(1):** 1-24.

Tandang-Silvas, M. R., Tecson-Mendoza, E. M., Mikami, B., Utsumi, S., Maruyama, N. (2011). Molecular design of seed storage proteins for enhanced food physicochemical properties. *Annual Review of Food Science and Technology*. **2:** 59-73.

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology.* **72(2):** 211-22.

Thevissen, K**.**, Ghazi, A**.**, De Samblanx, G.W**.**, Brownlee, C**.**, Osborn, R. W**.**, Broekaert, W. F. (1996). Fungal membrane responses induced by plant defensins and thionins. *Journal of Biological* Chemistry. **271(25):** 15018-25.

Thomma, B. P., Cammue, B. P., Thevissen, K. (2002). Plant defensins. *Planta*. **216:** 193–202.

Tivy, J. (1993). Biogeography: a study of plants in the ecosphere**.** 3rd ed. London: Longman Scientific and Technical. *Progress in Physical Geography.* **19:** 157-158.

Tokatlidis, K., Salamitou, S., Bbguin, P., Dhurjati, P., Aubert, P. (1991). Interaction of the duplicated segment carried by *Clostridium thermocellum* cellulases with cellulosome components. *FEBS Letters*. **291:** 185-188.

Towbin, H., Staehelin, T., Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences of the* United States *of America.* **76(9):** 4350-4354.

Trop, M., Birk, Y. (1970). The specificity of proteinases from *Streptomyces griseus* (Pronase). *Biochemistry Journal*. **116(1):** 19-25.

Tsumoto, K., Ejima, D., Kumagai, I., Arakawa, T. (2003). Practical considerations in refolding proteins from inclusion bodies. *Protein Expression and Purification*. **28(1):** 1-8.

USDA, NRCS. (2013). The PLANTS Database (http://plants.usda.gov). *National Plant Data Team*, Greensboro, USA.

van Loon, L. C., van Strien, E. A. (1999). The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiological and Molecular Plant Pathology*. **55:** 85-97.

Van der Salm T. P. M, Hänisch C., Dons J. J. M. (1996). Prospects for applications of rol genes for crop improvement. *Plant Molecular Biology Reporter.* **14:** 207–28.

Vergne, P., Maene, M., Gabant, G., Chauvet, A. (2010). Somatic embryogenesis and transformation of the diploid *Rosa chinensis* cv Old Blush. *Plant Cell, Tissue and Organ Culture.* **100(1) :** 73-81.

Villaverde, A., Carrió, M. M. (2003). Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnology Letters.* **25(17):** 1385-1395.

Visessuwan R, Kawai T, Mii, M. (1997). Plant regeneration systems from leaf segment culture through embryogenic callus formation of *Rosa hybrida* and *R. canina*. *Breeding Science*. **47:** 217–222.

Vlasak,J., Ionescu, R. (2011). Fragmentation of monoclonal antibodies. *MAbs*. **3(3):** 253-63.

Wang, L., Maji, S. K., Sawaya, M. R., Eisenberg, D., Riek, R. (2008). Bacterial inclusion bodies contain amyloid-like structure. *PLoS Biology*. **6(8):** 174-195.

Weigel, D., Glazebrook, J. (2006a). Vectors and *Agrobacterium* hosts for *Arabidopsis* transformation. *Cold Spring Harbour Protocols.* **2006(7):** 145-176.

Weigel, D., Glazebrook, J. (2006b). Transformation of *Agrobacterium* using the freeze-thaw method*.* **2006(7):** 167-189.

Xia, B., Etchegaray, J.P., Inouye, M. (2001). Nonsense mutations in cspA cause ribosome trapping leading to complete growth inhibition and cell death at low temperature in *Escherichia coli* . *Journal of Biological Chemistry*. **276:** 35581-35588.

Yasukawa, T., Kanei-Ishii, C., Maekawa, T., Fujimoto, J., Yamamoto, T., Ishii, S. (2005). Increase of solubility of foreign proteins in *Escherichia coli* by coproduction of the bacterial thioredoxin. *The Journal of Biological Chemistry.* **270(43):** 25328-25331.

Yawson, R., Yawson, I. (2008). Policy options of agricultural biotechnology R&D in Sub-Saharan Africa: key issues and aspects*. University of Minnesota*. **34880**.

Zimmerman, J. (1993). Somatic embryogenesis: a model for early development in higher plants. *The Plant Cell*. **5:** 1411–1423.

Zuker, A., Tzfira, T., Vainstein, A. (1998). Genetic engineering for cut-flower improvement. *Biotechnology Advances*. **16(1):** 33–79.

Annex I

1. Different microbiology media used in this work

PCA (Plate Count Agar), approximate formula per Litre:
- Pancreatic Digest of Casein- 5.0 g;
- Yeast Extract- 2.5 g;
- Dextrose- 1.0 g;
- Agar- 15.0 g.

MH (Mueller Hinton), approximate formula per Litre:
- Beef infusion solids- 2.0 g;
- Starch- 1.5 g;
- Casein hydrolysate- 17.5 g.

PDA (Potato Dextrose Agar), approximate formula per Liter:
- Potato extract- 4.0 g;
- Dextrose- 20.0 g;
- Agar- 15.0 g.

PDB (Potato Dextrose Broth), approximate formula per Liter:
- Potato extract- 4.0 g;
- Dextrose- 20.0 g.

GYP (Glucose, Yeast, Peptone), approximate formula per Liter:
- Yeast extract- 5 g;
- Peptone- 5 g;
- Glucose- 20 g;
- Agar- 15 g.

2.Elution volumes on the Superose 12 HR 10/30 column of different proteins:

Ferritin (440 kDa)- Elution volume, 9.5 mL
Catalase (282 kDa)- Elution volume, 10.9 mL
Aldolase (158 kDa)- Elution volume, 11.4 mL
Alcohol dehidrogenase (150 kDa)- Elution volume, 11.7 mL
Trypsin (23.8 kDa)- Elution volume, 15.3 mL
Citochorme C (12.4 kDa)- Elution volume, 15.7 mL.

Annex II

1. Microbiology media used in this work

LBA (Luria Bertoni Agar), approximate formula per Litre:

Bacto tryptone- 10.0 g;
Yeast Extract- 5 g;
NaCl- 10.0 g;
Agar- 15.0 g.

LB (Luria Bertoni Broth), approximate formula per Litre:

Bacto tryptone- 10.0 g;
Yeast Extract- 5 g;
NaCl- 10.0 g;

2. Primer pair used for BLAD expression

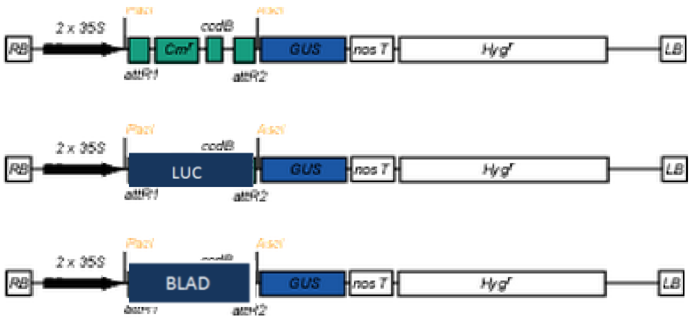
attBLAD primer FW: ACAAGTTTGTACAAAAAAGCAGGCTCGTAGACAAAGGAACCCTTA
attBLAD primer RV: ACCACTTTGTACAAGAAAGCTGGGTATCCTCATTCCCTAAAATAATCCTTTGTA

3. Plasmid vector cassette maps

pMDC107

pEXPLUC

pEXPBLAD



4. Multialign alignment between the BLAD gene published sequence and the pEXPBLAD173 sequencing result

```
           1        10        20        30        40        50        60        70        80        90        100       110       120       130
           |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
     Blad      CGTAGACAAAGGAACCCTTATCACTTCAGCTCTCAAAGATTCCAAACTCTTTACAAAAATAGGAATGGCAAAATC-CGTGTGCTCGAGAGGTTTGACCAAAGAACCAATAGACTTGAGAATCTCCA
    Blad3  TATGTGTAGACAAAGGAACCCTTATCACTTCAGCTCTCAAAGATTCCAAACTCTTTACAAAAATAGGAATGGCAAAATC-CGTGTGCTCGAGAGGTTTGACCAAAGAACCAATAGACTTGAGAATCTCCA
    Blad5                                                  ACTCTTTACAAAAATAGGAATGGCAAAATCNCGTGTGCTCGAGAGGTTTGACCAAAGAACCAATAGACTTGAGAATCTCCA
Consensus  .....gtagacaaaggaacccttatcacttcagctctcaaagattccaaACTCTTTACAAAAATAGGAATGGCAAAATC.CGTGTGCTCGAGAGGTTTGACCAAAGAACCAATAGACTTGAGAATCTCCA

           131      140       150       160       170       180       190       200       210       220       230       240       250       260
           |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
     Blad  AAACTACCGCATTGTTGAGTTCCAATCAAAACCTAACACTCTCATTCTCCCTAAACACTCTGATGCTGACTACGTCCTCGTTGTACTCAATGGTAGAGCCACAATCACGATAGTAAACCCTGATAGAAGA
    Blad3  AAACTACCGCATTGTTGAGTTCCAATCAAAACCTAACACTCTCATTCTCCCTAAACACTCTGATGCTGACTACGTCCTCGTTGTACTCAATGGTAGAGCCACAATCACGATAGTAAACCCTGATAGAAGA
    Blad5  AAACTACCGCATTGTTGAGTTCCAATCAAAACCTAACACTCTCATTCTCCCTAAACACTCTGATGCTGACTACGTCCTCGTTGTACTCAATGGTAGAGCCACAATCACGATAGTAAACCCTGATAGAAGA
Consensus  AAACTACCGCATTGTTGAGTTCCAATCAAAACCTAACACTCTCATTCTCCCTAAACACTCTGATGCTGACTACGTCCTCGTTGTACTCAATGGTAGAGCCACAATCACGATAGTAAACCCTGATAGAAGA

           261      270       280       290       300       310       320       330       340       350       360       370       380       390
           |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
     Blad  CAAGCATATAACCTTGAGTATGGCGATGCTCTCAGAATCCCAGCTGGCTCAACTTCATATATCCTTAACCCGGATGACAACCAGAAGCTTAGAGTAGTCAAGCTCGCAATACCCATCAACAATCCTGGCT
    Blad3  CAAGCATATAACCTTGAGTATGGCGATGCTCTCAGAATCCCAGCTGGCTCAACTTCATATATCCTTAACCCGGATGACAACCAGAAGCTTAGAGTAGTCAAGCTCGCAATACCCATCAACAATCCTGGCT
    Blad5  CAAGCATATAACCTTGAGTATGGCGATGCTCTCAGAATCCCAGCTGGCTCAACTTCATATATCCTTAACCCGGATGACAACCAGAAGCTTAGAGTAGTCAAGCTCGCAATACCCATCAACAATCCTGGCT
Consensus  CAAGCATATAACCTTGAGTATGGCGATGCTCTCAGAATCCCAGCTGGCTCAACTTCATATATCCTTAACCCGGATGACAACCAGAAGCTTAGAGTAGTCAAGCTCGCAATACCCATCAACAATCCTGGCT

           391      400       410       420       430       440       450       460       470       480       490       500       510       520
           |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
     Blad  ACTTTTATGATTTCTATCCATCGAGTACTAAAGACCAACAATCCTACTTCAGTGGCTTCAGCAGGAACACTTTAGAGGCCACCTTCAATACTCGTTATGAAGAGATACAAAGGATTATTTTAGGGAATGA
    Blad3  ACTTTTATGATTTCTATCCATCGAGTACTAAAGACCAACAATCCTACTTCAGTGGCTTCAGCAGGAACACTTTAGAGGCCACCTTCAATACTCGTT
    Blad5  ACTTTTATGATTTCTATCCATCGAGTACTAAAGACCAACAATCCTACTTCAGTGGCTTCAGCAGGAACACTTTAGAGGCCACCTTCAATACTCGTTATGAAGAGATACAAAGGATTATTTTAGGGAATGA
Consensus  ACTTTTATGATTTCTATCCATCGAGTACTAAAGACCAACAATCCTACTTCAGTGGCTTCAGCAGGAACACTTTAGAGGCCACCTTCAATACTCGTTatgaagagatacaaaggattattttagggaatga

           521      530       54042
           |---------+---------+-|
     Blad  GGAT
    Blad3
    Blad5  GGATTAGCTAGGAAATCTAGCA
Consensus  ggat
```

Annex III

1. Nucleotide sequences of the different BLAD genes used

BLAD 173 (GenBank: DQ142920.1):

1 cgtagacaaa ggaaccctta tcacttcagc tctcaaagat tccaaactct ttacaaaaat

61 aggaatggca aaatccgtgt gctcgagagg tttgaccaaa gaaccaatag acttgagaat

121 ctccaaaact accgcattgt tgagttccaa tcaaaaccta acactctcat tctccctaaa

181 cactctgatg ctgactacgt cctcgttgta ctcaatggta gagccacaat cacgatagta

241 aaccctgata gaagacaagc atataacctt gagtatggcg atgctctcag aatcccagct

301 ggctcaactt catatatcct taacccggat gacaaccaga agcttagagt agtcaagctc

361 gcaataccca tcaacaatcc tggctacttt tatgatttct atccatcgag tactaaagac

421 caacaatcct acttcagtgg cttcagcagg aacactttag aggccacctt caatactcgt

481 tatgaagaga tacaaaggat tattttaggg aatgaggat

BLAD 127:

1 atg ag gaatggca aaatccgtgt gctcgagagg tttgaccaaa gaaccaatag acttgagaat

121 ctccaaaact accgcattgt tgagttccaa tcaaaaccta acactctcat tctccctaaa

181 cactctgatg ctgactacgt cctcgttgta ctcaatggta gagccacaat cacgatagta

241 aaccctgata gaagacaagc atataacctt gagtatggcg atgctctcag aatcccagct

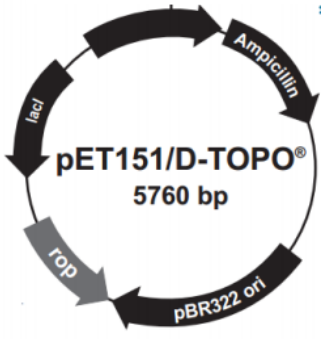
301 ggctcaactt catatatcct taacccggat gacaaccaga agcttagagt agtcaagctc

361 gcaataccca tcaacaatcc tggctacttt tatgatttct atccatcgag tactaaagac

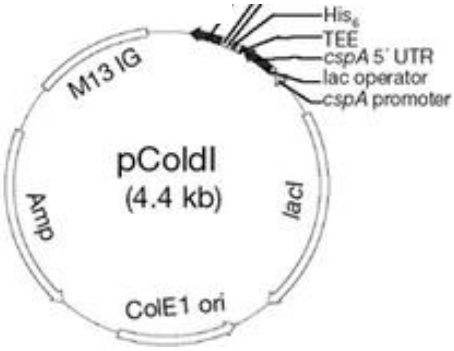
421 caacaatcct acttcagtgg ttag

2. Expression plasmid vector maps

pET151 D-TOPO plasmid



pColdI plasmid



3. BLAD 173 primers

Primer Fw: CACCCGTAGACAAAGGAACCCTTA

Rev compl.: ATCCTCATTCCCTAAAATAATCCTTTGTA

4. BLAD 127 primers

Primer Fw: CACCATGAGGAATGGCAAAATCC

Rev compl.: CTAACCACTGAAGTAGGATTGTTG