Cysteine proteases activity and gene expression studies in soybean nodules during development and drought stress

By

MAGDELEEN DU PLESSIS

Dissertation submitted in partial fulfilment of the requirement for the degree

MAGISTER SCIENTIAE

In the Faculty of Natural and Agricultural Sciences

University of Pretoria

Pretoria

Supervisors: Prof K J Kunert

Dr. B.J. Vorster

March 2013

Declaration

I, Magdeleen du Plessis, the undersigned hereby agree that the thesis submitted herewith, for the degree Magister Scientia, to the University of Pretoria, contains my own independent work.

This work has not been submitted for any other degree at any other University.

Signature:

Date:

TABLE OF CONTENTS

2. MATERIALS AND METHODS

ABSTRACT

Activity and transcription profiles of two classes of cysteine proteases, papain- and legumain-like cysteine proteases, as well as their potential inhibitors, cysteine protease inhibitors (cystatins), were investigated in soybean nodules during nodule development and after drought inducing premature senescence. During nodule development total protease activity increased with major activity bands detected protease zymography in older nodules. Expressed cysteine proteases during nodule development were detected by tagging proteases with the cysteine protease inhibitor DCG-04 with major DCG-04 tagged bands found in both young and old nodules. Increase in protease activity was associated with a significant decrease in nitrogenase activity of nodules measured as acetylene reduction. Semi-quantitative RT-PCR for cysteine protease and cystatin transcription profiling showed a decrease in transcription during development and also after drought treatment of several papain-like cysteine proteases (Glyma04g04400, Glyma17g05670, Glyma10g35100, and Glyma04g03090). In contrast, transcription of three legumain-like cysteine proteases (Glyma17g14680, Glyma05g04230 and Glyma14g10620) increased during nodule development and also after drought treatment. Transcription of two cystatins (Glyma13g27980 and Glyma05g28250) increased during nodule development with Glyma13g27980 strongly up-regulated after drought treatment and Glyma05g28250 constitutively strongly expressed in both well-watered and drought treated nodules. Overall, the study has contributed in establishing an expression profile of cysteine proteases and cystatins in soybean nodules. This knowledge provides a basis which can be used to determine the importance of the individual components of the cysteine protease – cystatin system, during soybean nodule development and during stress-induced premature nodule senescence.

COMPOSITION OF DISSERTATION

Chapter 1 of this dissertation provides an overview about legume nodule biology and the current knowledge of the protease-protease inhibitor system in soybean and its possible biological function in plants. Also outlined are the mechanisms of how plant senescence and stress affects this system. The objectives and aim of the study are provided at the end of this chapter. **Chapter 2** outlines the materials and methods used in this study. This includes growth and drought treatment of soybean plants and the different molecular and biochemical techniques applied to characterize plant tissues and nodules. **Chapter 3** outlines the results obtained from characterizing nodule material of different age and after drought treatment in particular the various expression studies carried out for cysteine proteases and cystatins. **Chapter 4** focuses on the outcomes of the study and new results obtained are discussed in greater detail. The chapter also outlines possible future actions to further advance the function of the cysteine protease-cystatin system in nodule development and under stress. In **Chapter 5** the appendix provides information about all DNA sequences used during this study. Finally, literature cited in this study is listed in **Chapter 6.**

ACKNOWLEDGEMENT

I wish to thank the National Research Foundation for giving me the opportunity to do my MSc by giving me financial support throughout my studies. Sincere gratitude goes out to my supervisors, Prof Karl Kunert who supported, advised and had lots of patience with me throughout the entire course of my study and for giving me great opportunities. A special thanks to Dr Juan Vorster who was always willing to help and advise me with all the possible questions and problems that I had. I also wish to thank Dr Rosita Yocgo and Dr Berhanu Fenta who always had time to help and guide me, without your help I would have been completely lost. Also thank you to all my other Plant Molecular Physiology lab members Stefan van Wyk, Priyen Pillay, Rynhard Smith, Tintswalo Maluleke and Sylvia Phokane. Thank you to Monique Heystek a dear friend, who felt like she was part of our lab, always willing to help and listen. I also wish to thank Prof Gerrit Beemster and my lab members at the University of Antwerp in Belgium for giving me the opportunity to visit your lab, making me feel welcome and the excellent guidance and support. Thank you to my family and friends whose love and encouragement were of great help. A special thanks to Francois Cilliers for his care and love. Lastly but not least, thank you to My Lord Jesus Christ for courage, love, strength and making all this possible.

INDEX OF FIGURES

INDEX OF TABLES

ABBREVIATIONS AND ACRONYMS

3-D	Three Dimensional
Ala	Alanine
Arg	Argenine
ATP	Adenosine triphosphate
bp	Base pair
BSA	Bovine Serum Albumin
cDNA	Complementary DNA
CFU	Colony Forming Unit
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CN	Crown nodules
CP	Cysteine protease
C-Terminus	Carboxyl-terminus
Cys	Cysteine
D	Drought
dH_2O	Sterile distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotide triphosphates

DTT	Dithiothreitol
E-64	Trans-(epoxysuccinyl)-L-leucylamino-4-guanidine butane
EDTA	Ethylenediamine tetra acetic acid
EST	Expressed sequence tag
FU	Fluorescence unit
FW	Fresh mass
Gln	Glutamine
Gly	Glycine
IPTG	Isopropylthio-β-galactoside
kDa	Kilo Dalton
LB	Luria-Bertani
Leu	Leucine
LN	Lateral nodules
MCA	α-amino 4-methylcoumarin
Mr	Relative molecular mass
Nod	Nodulation
N-Terminus	Amino-terminus
OD	Optical density
OL	Old leaves

P	Pods
PAR	Photosynthetically active radiations
PCD	Programmed Cell Death
PCR	Polymerase chain reaction
pH	Log Hydrogen Ion Concentration
PI	Protease inhibitor
Pro	Proline
RNA	Ribonucleic acid
RNAseq	RNA sequencing
rpm	Rotations per minute
RT	Root tips
RT-PCR	Reverse transcriptase PCR
SA	South Africa
SAGs	Senescence associated genes
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SE	Standard error
Semi-qPCR	Semi quantitative PCR

SOC media	Super Optimal broth with Catabolite repression media
TAE	Tris-acetate-EDTA
Taq	*Thermus aquaticus*
TLCK	Tosyllysine Chloromethyl Ketone (hydrochloride)
Trp	Tryptophan
U	Unit
UK	United Kingdom
USA	United States of America
V	Voltage
v/v	Volume per volume
Val	Valine
VPE	Vacuolar Processing Enzymes
w/v	Weight per volume
WW	Well-watered
Xaa	Any amino acid
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
YL	Young leaves

1. Introduction

1.1 Soybean nodules

Soybean (*Glycine max* L.) is a versatile crop, an excellent source of proteins and oil (Keyser and Li, 1992) and is therefore worldwide an important food and feed crop. The symbiotic association between soybean and *Bradyrhizobium japonicum* or *Sinorhizobium* species allows for the fixation of atmospheric nitrogen by the bacteria inside root nodules. This symbiosis will start at the root surface and will result in the formation of nitrogen-fixing nodules (Figure 1.1) (Vincent *et al*., 1980). Nitrogen is an important nutrient and is needed by plants in high amounts as high nitrogen content will have a positive effect on yield as well as the quality of legume plants (Devienne-Barret *et al*., 2000). Root nodules further contribute to the nitrogen content in the soil (Puppo *et al*., 2005).

Figure 1.1: Crown and lateral nodules on soybean roots. Crown nodules develop on the tap roots at an early stage and lateral nodules develop at a later stage on the lateral root hairs. (Right photo taken by Berhanu Fenta, left photo taken from http://www.agnet.org)

1.1.1 Nodule development

Nodules are formed when a mutualistic symbiotic relationship develops between soybean roots and *Rhizobium*. During the symbiosis, the host plant will provide energy in the form of carbon to the *Rhizobium* (Colebatch *et al*., 2004). The bacteria, in return, will bind atmospheric nitrogen to form ammonium which the plant can then transport and use in protein synthesis (Keyser and Li, 1992).

This symbiosis is initiated by the *Rhizobiums's* Nod factors, which releases lipochito-oligosaccharide molecules. Nod-genes will then be expressed to respond to flavonoids (Figure 1.2) and other molecules that are secreted into the soil by the plant (Stougaard, 2000). The soybean plant lowers its endogenous defences which enable the bacteria to invade the organism and colonize the root surface and the root hairs will start to curl. Differentiated cells of the root cortex will form a primordium. The cortical tissue of the root hair will be invaded by an infection thread. This infection thread will then travel into the nodule primordium (Schultze and Kondorosi, 1998). Within the infection thread the bacteria will multiply. The Nod factors will modify the hormones of the plant so that mitosis of the host plant is stimulated and the peribacteroid can be formed (Puppo *et al*., 2005). The bacteria will be released from the infection thread and will be encoiled by a peribacteroid membrane so that the bacteria are separated from the host cell cytoplasm. Other changes that also occur during nodule development are the formation of a vascular network which will transport the photosynthate to the nodules in the form of sucrose. This network will then also in return be used to export the nitrogen containing compounds from the nodules to the plants.

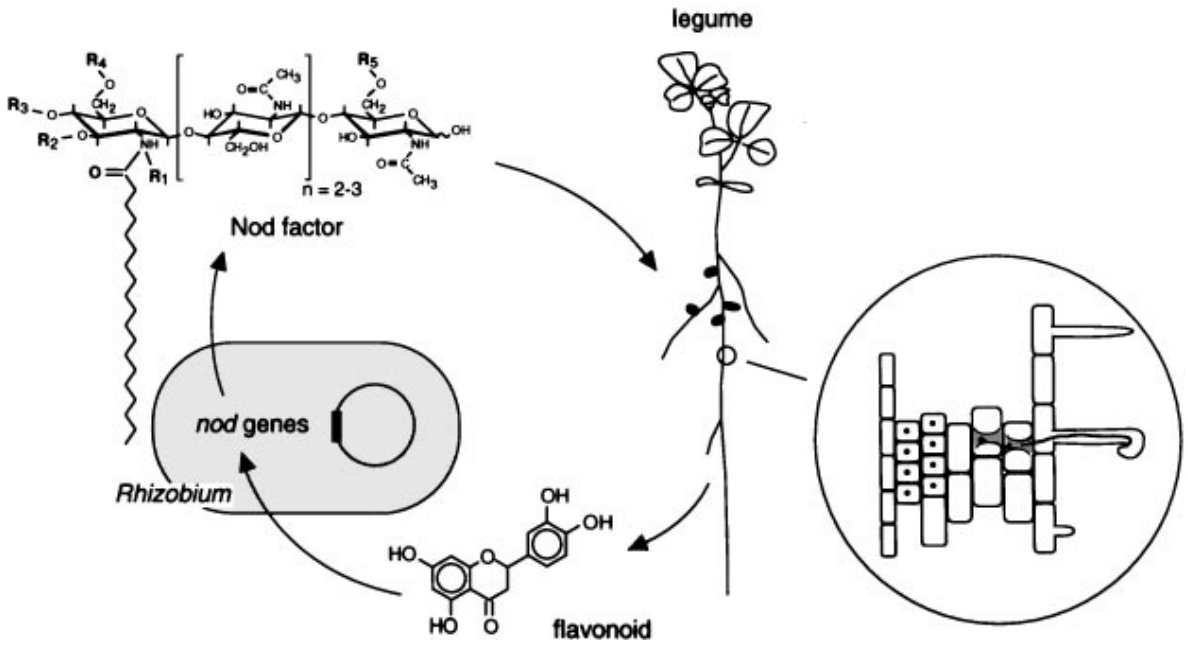

Figure 1.2: Signal exchange in the *Rhizobium*-plant symbiosis. Flavonoids induce the *rhizobia* nod genes. This leads to the production of nodule-inducing (Nod) factors. The insert shows the infection thread that will become a nodule primordium passing the root cortex (Schultze and Kondorosi, 1998).

Two types of nodules can be distinguished based on their development, namely determinate and indeterminate nodules. Determinate nodules (present in soybean) are initiated from the meristem cells in the outer cortex and the division process stops ten days after infection. Determinate nodules (Figure 1.3a) have no active meristem part and have a very different form compared to indeterminate nodules (Puppo *et al*., 2005). The process of senescence starts at the centre of determinate nodules and spreads to the outer edges.

In comparison, indeterminate nodules (Figure 1.3b) have a more cylindrical appearance and can be divided into five distinctive zones (Puppo *et al*., 2005). The first being made up out of small meristematic cells. The second zone consists of the infection zone which leads to the third zone where the bacteria are housed. Zone four will only differentiate when the nodules age and senescence can be observed here first as bacterial cells decline. Zone five has free living bacteria and no symbiosis takes place here (Timmers *et al*., 2000).

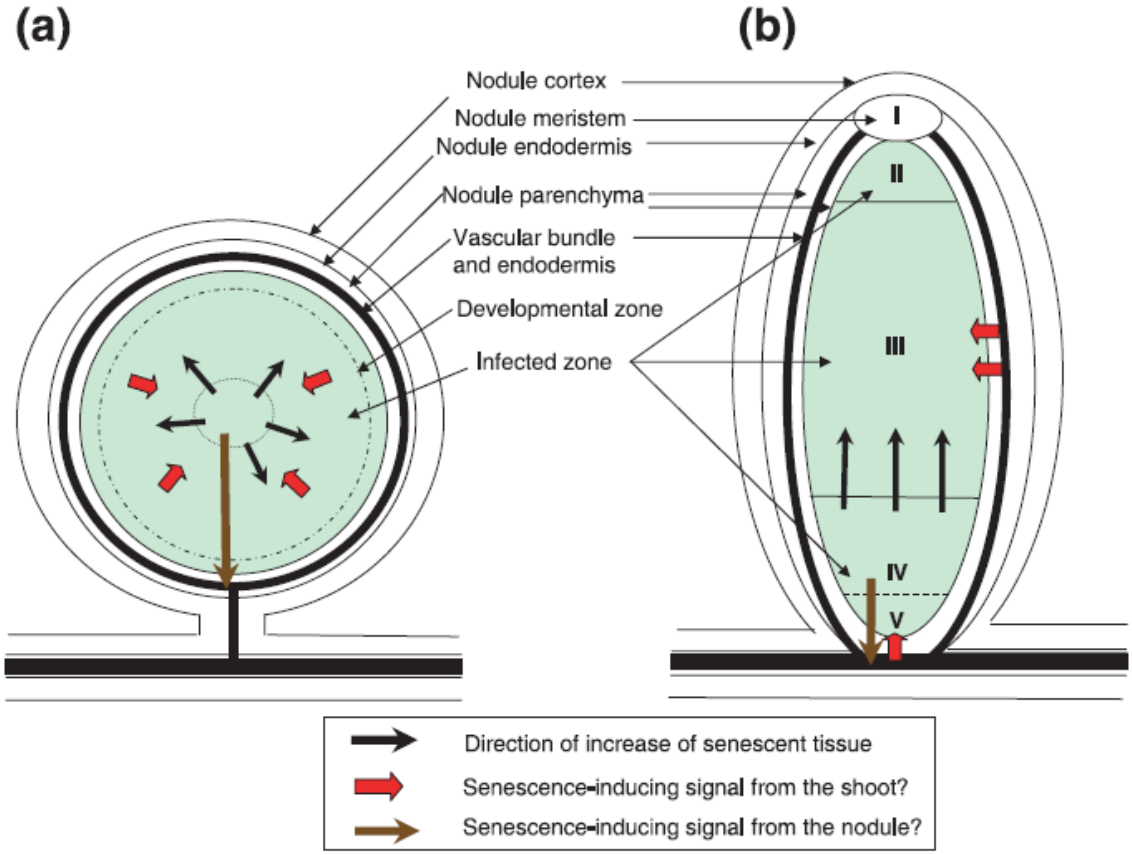

Figure 1.3: The structure of the a) determinate and b) indeterminate nodules. Although the nodules differ, there are regions which are present in both types such as the internal cortex called the nodule parenchyma and the central zone (Puppo *et al*., 2005).

1.1.2 Nodule senescence

Senescence is a programmed degeneration of cells that can lead to the death of a plant or organ. Senescence is an internally programmed process that progresses in an orderly fashion (Noodén *et al*., 1997). It is a genetically controlled process and can either be initiated at a specific time point in the plants life cycle (developmental senescence) or it can be induced by environmental conditions such as drought, limitations of nutrients, change in climate and little light exposure (induced senescence) (Huffaker, 1990). During senescence 2500 genes are activated, 7% of these genes include different kinds of hydrolases and proteases (Martínez *et al*., 2007). An important function of senescence is protein remobilization and salvaging of nitrogen reserves. Proteins in senescent organs will be degraded and remobilized to the reproductive parts of the plants (Huffaker, 1990). The plant may also lose some of its photosynthetic activities, and other macro molecules will also be remobilized (Li *et al*., 2008). Senescence is, however, not an inactive aging process but is controlled by many internal as well as external signals (Huffaker, 1990).

Nodules, like other parts of plants, undergo senescence. The life span of the nodule will depend on the plant species as well as the *Rhizobium* strain and on prevailing environmental conditions (Swaraj and Bishnoi, 1996). Nodules have an average life span of 10-12 weeks after senescence is initiated (Puppo *et al*., 2005). During the senescence process there is a few visible changes that can be observed in the nodule. One of these changes is a colour change from pink to greenish. This is due to the alteration of the functional proteins like leghemoglobin (Swaraj *et al*., 1995). Another change is that the symbiosome will undergo an organ structure change during nodule senescence with cytoplasm in the nodules becoming less dense and the shape and size will then lead to the deterioration of the bacteriod. In

determinate nodules, this decrease in density observed in the cytoplasm does not happen homogenously. Necrosis will start in the centre of the nodule with an increase in hydrogen peroxide and lipid hydroperoxides (Puppo *et al*., 2005). This results in membrane damage early in senescence, especially in the membrane of the symbiosomes that will change in size and shape due to the deteriorating bacterioid. This is also characterised in early senescence by a rapid decline in nitrogen fixation (Alesandrini *et al*., 2003). Other organelle changes such as an increase in peroxisomes and mitochondria forming elongated structures also occur (Puppo *et al*., 2005). In indeterminate nodules, nodules will start to deteriorate due to a mature zone that is established proximal to zone three and which will move to the apical region (Van de Velde *et al*., 2006).

Other factors that also have an effect on nodule senescence are the progress of root and shoot development as well as whether pod-filling have been initiated. Plant nitrogen status is sensed in the shoot and roots and the amount of ammonium assimilate that is produced and exported by the shoot will have a direct result to the nodulation and nodule senescence (Scheible *et al*., 1997). During the pod-filling stage there is an inadequate supply of photosynthetic products to the nodules that result in a decrease in nitrogen fixation.

1.2 Proteases and protease inhibitors

The hydrolysis of proteins is very important for the normal functioning of a cell. Proteolysis of proteins is needed for the normal development, homeostasis and even the death of a plant cell during senescence or programmed cell death (PCD). Proteolytic enzymes are needed to cleave the peptide bonds of proteins so proteins can be degraded. These enzymes are called proteases. Proteases will target mature enzymes by helping with production and assembly of

proteins that gets remobilized (Prins *et al*., 2008). Proteolysis also erases mistakes that occurred during the biosynthesis of proteins. Proteins that will undergo rapid degradation are proteins that are in the wrong cellular compartments or proteins that are not folded correctly (Andersson and Aro, 1997). Most proteins will be degraded at one stage or another due to un-regulated cellular proteolytic enzymes or when the nitrogen and carbon from the amino acids are needed for life in the organism (Estelle, 2001).

Proteases are found in nearly all organisms, from plant, micro-organisms to animals. Plant proteases are responsible for a variety of different processes involved in protein metabolism. There are different types of proteases including endoproteases, aminipeptidases and carboxypeptidases. These proteases are classified according to where they will cleave a protein. Endopeptidases can cleave internal amino acid residues and can usually be classified as having a serine amino acid in its active site (Huffaker, 1990). This group of proteases are usually found in vacuoles according to Boller and Wiemken (1986). Endoproteases are also divided into different classes depending on the specific site where they will cleave the target protein. These include: serine protease, cysteine protease, aspartic protease and metalloprotease. Exopeptidases will work on the termini of the peptides. There are two types of exopeptidases, one that cleaves N-terminal residues, hydrolyzing the bond between the aminoacyl residue and the peptide, and the other cleaves C-terminal residues hydrolyzing the amino acids from peptide chain (Barrett, 1986).

1.2.1 Cysteine proteases

These proteases contain a catalytic sulfhydryl group in its active centre with a cysteine and a histidine residue. Cysteine proteases are responsible for the dismantling of organelles and

different macro molecules so that nutrients can be released and remobilized from senescent organs to actively growing tissues of the plant (Beers *et al*., 2000). Cysteine proteases are involved in a variety of plant processes, such as processing and folding of storage proteins in seed development (Gruis *et al*., 2002), the remobilization of stored proteins during seed germination, fruit ripening (Alonso and Granell, 1995), hormone signalling, embryogenesis and morphogenesis (Salas *et al.*, 2008). A few other examples include the expression of cysteine proteases in the flowers of daylilies, in the senescing leaves of tomato (Drake *et al*., 1996) and maize (Smart *et al*., 1995). Cysteine proteases account for about 90% of the total degradation of storage proteins in cereals like wheat (Bottari *et al*., 1996) and are one of the main proteases responsible for the breakdown of proteins during cell death. Cysteine proteases are also involved in a wide range of processes in response to both biotic and abiotic stress.

Two kinds of cysteine proteases, papain-like and caspase-like, are activated during the process of PCD caused by different abiotic or biotic stress factors (Solomon *et al*., 1999). Abiotic stresses such as cold, heat shock, salt and dehydration can also induce early plant senescence with increased activity of vacuolar cysteine proteases that are involved in PCD (Beyene *et al*., 2006, Martinez *et al*., 2007). Cysteine proteases are eccentric features of the senescence process. The majority of plant cysteine proteases belong to the papain (C1) or legumain (caspases) (C13) families.

1.2.2 Papain- and legumain-like proteases

Papain-like proteases were discovered in the latex and the fruit of the *Carica papaya*. This protease contains 121 amino acids and is used as a model to study cysteine proteases

(Schaller, 2004). They are also the first proteases to have their amino acid sequenced and the three-dimensional (3-D) structure determined (Grudkowska and Zagdanska, 2004). Papain-like proteases can be grouped according to their similarity to mammalian types of cathepsin L-, B-, H- and F-like according to their gene structures and phylogenetic relationship (Martinez and Diaz, 2008). Proteases belonging to the group C1A are synthesised in an inactive form to prevent unnecessary proteolysis. For these proteases to become active they have to undergo self-processing or have to be processed by another enzyme (Wiederanders, 2003). These inactive pro-enzymes are synthesised in the endoplasmic reticulum and will then be transported to the different organelles, such as the vacuole and the lysosomes, by the Golgi-apparatus (Martínez *et al*., 2012). Papain-like proteases have also been isolated in legume nodules and germinating seeds (Lohman *et al*., 1994). Solomon *et al*. (1999) also reported papain-like cysteine proteases being modulators of PCD. To date more than 600 papain-like proteases (C1A) have been included into the MEROPS database (Martínez *et al*., 2012).

Caspases are a very specific type of cysteine proteases that contain an aspartate specific centre that will regulate PCD in animals. Caspase-like proteases have also been found in plants and are involved in PCD. It was long debated if plant caspases exist and it has been established that there are plant caspase-like enzymes that can be the same as animal caspases but it is possible that they might differ structurally. Cysteine proteases known as vacuolar processing enzymes (VPE), also known as the legumain family, show caspase-like activity (Hara-Nishimura *et al*., 2005).

Legumains are a family of aspartic-specific cysteine protease that are located in the vacuoles or cell walls (Müntz and Shutov, 2002). These enzymes are able to cleave any asparagine-

flanked peptide bond, with any of the 20 amino acids in the C-terminal flank of the cleavage region in P1' position (Ishii, 1994). Because of these enzymes location and function, they are called vacuolar processing enzymes (VPE). VPE's sequences are not similar to those of other known cysteine proteases even though the cysteine and histidine residues in the active centre are the same. The legumains are also not as sensitive to the inhibition from E-64 as other papain-like proteases (Müntz and Shutov, 2002). Legumains are only active at acidic pH's that prevail in vacuoles and cell walls (Müntz and Shutov, 2002).

VPE's can be subdivided in two groups: the vegetative type (αVPE and γVPE) and the seed type (βVPE). The βVPE type is involved in the processing of seed storage proteins *in vivo* whereas the vegetative type is up-regulated during PCD, during different stress conditions and will also increase if the hypersensitive response (HR) is initiated, but will decrease before the lesion is formed. VPEs are responsible for the collapse of the vacuole membrane releasing different proteases into the cell which will then be able to complete the PCD or senescence process (Hara-Nishimura *et al*., 2005). It has been proposed that γVPEs are responsible for the maturation of other vacuolar proteins and proteases involved in amino acid recycling during senescence (Roberts *et al*., 2012). This was also found by Sajid *et al*. (2003) when a pre-pro-region of legumain cleaved and activated a cathepsin B from *S. mansoni*.

Vacuolar processing enzymes (VPEs) play an important role in PCD (Fedorova and Brown, 2007) and such VPEs have been isolated in senescing leaves and in root cells of Arabidopsis undergoing PCD. Vegetative VPEs have been found in wounded leaves as a stress response and they also act as processing enzymes in events like senescence and PCD (Müntz and Shutov, 2002).

1.2.3 Nodule cysteine proteases

Proteolytic enzyme activities have been detected in developing and senescent nodules (Asp *et al*., 2004). Proteases identified in the cytoplasm of nodules have increased expression during nodule senescence. These proteases in the nodules will target the cytosolic proteins that will affect the bacteriods. Early senescence in alfalfa showed increased activity of acidic proteases. These acidic proteases affect the peribacteriod membrane stability and will lead to the rupture of the peribacteriod membrane and the elimination of the microbial partner (Pladys and Vance, 1993). Proteolytic enzymes are usually limited to the infected cells of the nodules and will thus limit the bacterial symbiosis. Proteases that digest the cell wall of the peptidoglycan of the bacteriod were identified in French beans. These proteases were only expressed during senescence (Pladys and Vance, 1993). The increased activity of proteases in the nodule during senescence leads to reduced nitrogen fixation and ammonia production (Pfeiffer *et al*., 1983).

Host cells are able to adapt to stress and Manen *et al*., (1991) reported that cysteine proteases could be involved in this adaptation process. Cysteine proteases are abundant during senescence in late nodules (Lee *et al*., 2004). According to Li *et al*. (2008), nodule cysteine proteases can be categorized into two classes namely nodule specific cysteine proteases and nodule enhanced proteases. A nodule specific protease, *As*NODF32, isolated from *A. sinicus* (a winter growing Chinese milk vetch) is also activated during nodule senescence and belongs to the papain protease family. These proteases were detected in 14 day old nodules and it was found that their expression increased up to five-fold at the age of 30 days. *In situ* hybridization indicates that these findings are age-dependent (Naito *et al*., 2000). Hybridization showed that *As*NODF32 is expressed in the senescent and inter-zone implying

a link to nodule senescence (Li *et al*., 2008). Strong hybridization signals of proteases were observed in the senescence zone of the nodule. Hybridization showed that cysteine protease activity increased in senescent indeterminate nodules. Lievens *et al*. (2004) found that the nodule lifespan is prolonged through the silencing of *As*NODF32 leading to increased nitrogen fixation. However, it is crucial to regulate the expression of certain proteases that could be damaging to plant cells and speed up the process of nodule senescence.

A previous study also showed that anti-sense inhibition of CYP15A, a cysteine protease originally found in wilted pea shoots, delayed the senescence process in *Medicago trunctula* (Sheokand *et al*., 2005). Pladys and Vance (1993) reported that cysteine protease activity increased clearly at the onset of senescence *(*Lee *et al*., 2004). A microarray analysis of cDNA libraries that were prepared from soybean nodules at different stages of nodulation (2, 5, and 10 weeks), showed that the transcript level increased in mature nodules and senescent nodules. After cDNA markers were differentially expressed by dot-blot analysis (Alesandrini *et al*., 2003), the cysteine protease clone DD15 encoding a protein involved in remobilisation occurring during natural or stress induced senescence, showed high expression in senescent nodules.

Root nodule proteases are involved in the adaptation to physiological stresses and the control of root senescence (Kardailsky and Brewin, 1996), but few genes have been isolated so far from senescent nodules. Cysteine proteases do get expressed in younger nodules but they have very low proteolytic activity and are strictly compartmentalized to stop proteolysis in fixing nodules (Pladys and Vance, 1993). Hydrolytic enzymes will be activated in the nodule as they become older. Whole ranges of proteases are involved in the breakdown of nodule

proteins with thiol-type proteases to serve a very important function in senescing soybean (Malik *et al*., 1981).

1.2.4 Protease inhibitors

Protease inhibitors are small proteins that are found in high concentrations that are able to regulate the activity of proteases. These proteins have been found in storage tissues but also in aerial parts of the plant (Lea and Miflin, 1980). All plant protease inhibitors can form reversible complexes with the proteases so that the proteases are left inactive**.**

Protease inhibitors are involved in a wide range of defence mechanisms (Lievens *et al*., 2004). Protease inhibitors protect the plant against herbivorous predators as well as fungi and bacteria that uses digestive and hydrolytic enzymes to gain entry into the plant (Johnson *et al*., 1989). These inhibitors are resistant to extreme heat and pH, which gives them the function as storage proteins so that they are immune to digestion until germination occurs (Richardson, 1977).

Plant protease inhibitors are active against all the different classes of protease, cysteine, serine, metallo and aspartic proteases (Valueva and Mosolov, 1999). Some families contain inhibitors of more than one type of protease, for example the Kunitz family, and are able to inhibit serine proteases (family S1) and cysteine proteases (C1 family) (Lea and Miflin, 1980).

Cystatins are inhibitors of cysteine proteases. They play very important roles in physiological and cellular processes and also the inhibition of exogenous cysteine proteases secreted by

pathogens, herbivores and other insects (Benchabane *et al*., 2010) Plant cystatins, also called phytocystatins, are inhibitors of the papain-like type cysteine proteases that contain the Gln–Xaa–Val–Xaa–Gly motif in the centre of the polypeptide chain (where Xaa is any amino acid), a Pro–Trp (or Leu–Trp) dipeptide motif in the C-terminal region, and a conserved Gly residue in the N-terminal (Benchabane *et al*., 2010). The cystatins will attach to the active site of cysteine proteases (Figure 1.4) and will leave the enzyme inactive so that it can't access the protein substrates by forming an irreversible complex with the cysteine proteases. A surface hairpin loop with the motif Gln–Xaa–Val–Xaa–Gly will interact with the target site of the enzyme. Another hairpin loop from the C-terminal with the motif Pro-(Leu)–Trp will also enter the active site. The conserved Gly residue at the N-terminal of the cystatin will help with the binding process to the enzyme as well as the specificity toward cysteine proteases (Benchabane *et al*., 2010).

Cystatins can be grouped into four groups namely the stefin, cystatin, kininogens and phytocystatin group (Turk and Bode, 1991). Most plant cystatins fall into the phytocystatin group. Inhibitors are classified into these groups according to their molecular masses, disulphide bonds and sequence similarities. Inhibitors from the stefin family are single-chained, they lack a disulphide bond and have a molecular mass (Mr) of approximately 11000, whereas the cystatin group has a Mr of 13000 and two disulphide bonds (Turk and Bode, 1991). Plant cysteine proteases inhibitors can then be further grouped regarding to their domains. The first group is consisting of single-domain phytocystatins, and the second group are multi-cystatins found in potato tubers (Walsh and Strickland, 1993).

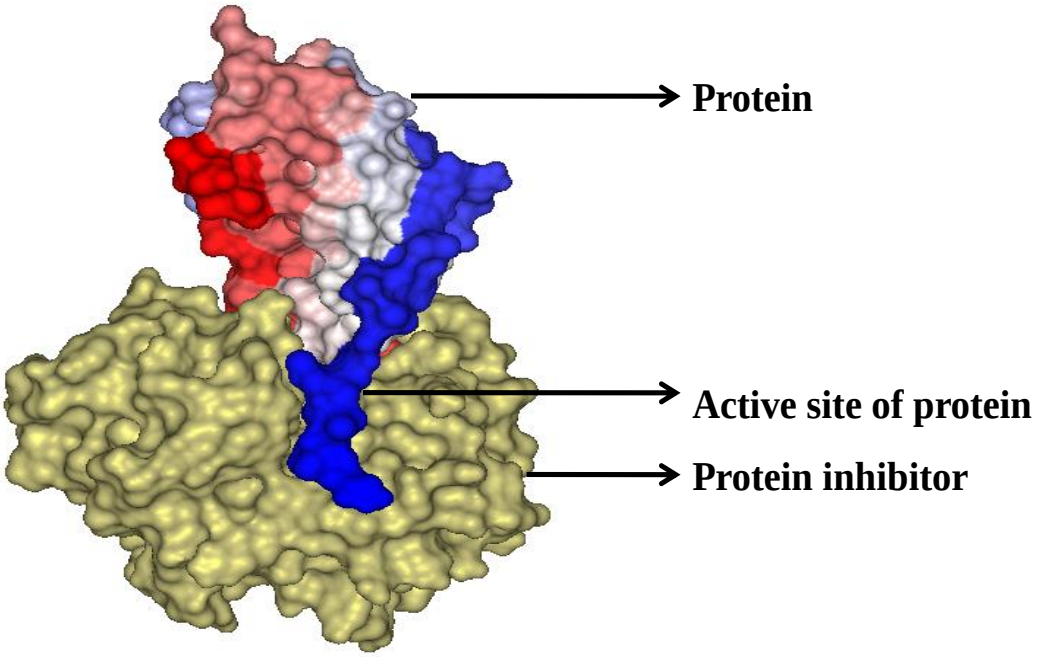

Figure 1.4: Cysteine protease inhibitor binding to a cysteine protease (Image provided by Dr Juan Vorster, UP).

Oryzacystatin from rice was the first plant cystatin to be characterized. Oryzacystatin (Figure 1.5a, b) has a five-stranded anti-parallel ß-sheet wrapped around a central a-helix (Nagata *et al*., 2000). Since cystatins inhibit cysteine proteases, they play a vital role in the plant protease regulatory system. Studies have shown that the physiological processes involving cysteine proteases inhibitors include deposition and mobilization of protein pools in storage and senescent organs and the down-regulation of cysteine protease activities in storage organs (Turk and Bode, 1991). Cystatins will be synthesised in young vegetative storage organs and will, at a later stage, be out-numbered by the cysteine proteases in theses organs and thus promoting the deposition of proteins (Benchabane *et al*., 2010).

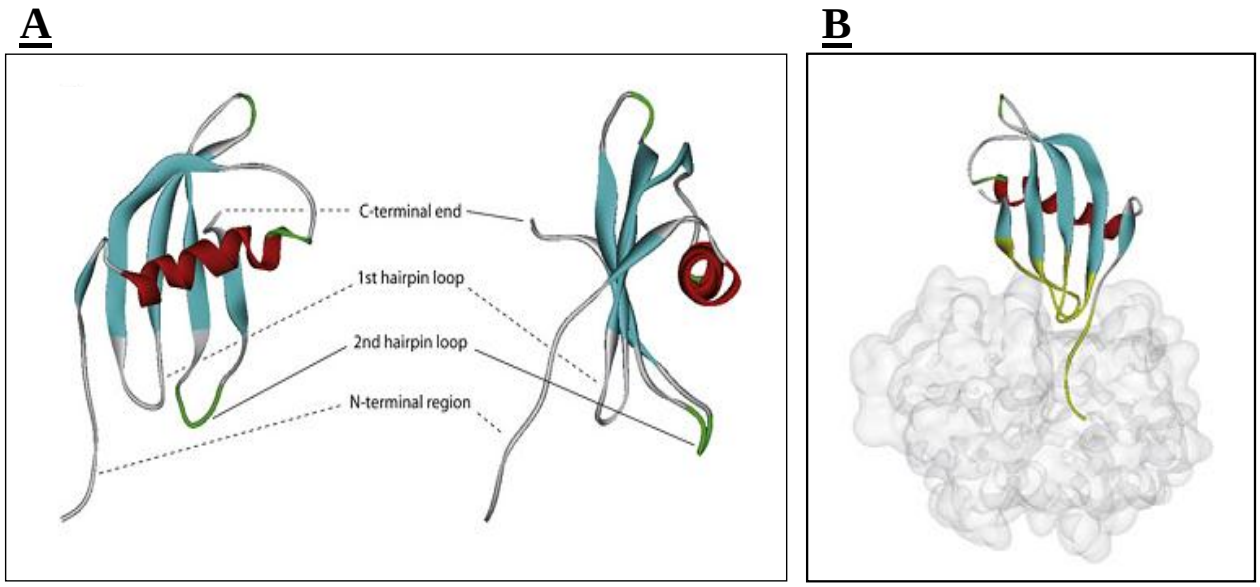

Figure 1.5: 3-D Representation of Oryzacystatin a) has a five-stranded anti parallel ß-sheet wrapped around a central a-helix and b) Oryzacystatin binds to a cysteine protease (Benchabane *et al*., 2010)

Hong *et al*. (2007) suggested that there are also protease inhibitors that are involved in nodulation. Lievens *et al*. (2004) showed that a serine inhibitor (SrPI1), belonging to the Kunitz-type family, is expressed in the nodule primordium in the early stages of nodulation in the plant *Sesbania rostrata* (Lievens *et al*., 2004). Transcripts of this inhibitor were also found in the central nodule tissue. Lievens *et al*. (2004) suggested that the inhibitory role of this inhibitor is involved in the control of proteases, such as cathepsins B, H and L, as the expression decreases in mature nodules. Maximal expression of the inhibitor was two days after nodulation but expression greatly decreased 12 days after nodulation. The inhibitor was detected in the younger part of the fixation zone according to an *in situ* localization study.

The transcriptional regulation of cystatin biosynthesis can play an important role in the process of protein recycling during the senescence process and under abiotic stress conditions (Benchabane *et al*., 2010). Studies have shown that cystatins and serine-protease inhibitors are down-regulated and cysteine proteases are up-regulated in senescent organs (Sugawara *et al*., 2002). However, other studies have shown that cystatin mRNA transcripts increase in leaves experiencing abiotic stress such as drought, salinity or low temperature. Coupe *et al*. (2003) also suggested that inhibitors can be modulators of PCD.

1.3 Cysteine protease and cystatin expression under drought

Drought is a serious environmental abiotic stress severely limiting plant growth (Simova-Stoilova *et al*., 2010). Plant growth rates are directly proportional to the amount of water that is available to the plant from the soil (Kameli and Lösel, 1995). A plant will suffer drought stress if the rate of transpiration exceeds the water uptake (Lawlor and Cornic, 2002). Bray (1997) suggested that plants response to drought stress will depend on factors like the

genotype of the species, the length and severity of the drought stress and even age and development. Responses to drought stress can be on a metabolic or structural level (Bohnert and Sheveleva, 1998). Changes can be seen in the root and shoot ratio, accumulation of reserves in petioles and even the anatomical changes in the leaf area (Pinheiro *et al*., 2001). Other changes in the plant also include a change in the transcriptional activation of different genes. One of these responses to water stress is to synthesise protective proteins and alter the senescence process (Simova-Stoilova *et al*., 2010).

Soybean is greatly affected by drought (Clement *et al*., 2008). Drought leads to a decrease in nitrogen fixation (Serraj *et al*., 1999) decreasing seed yield. In particular the symbiotic relationship between the plant and the bacteria is very susceptible to environmental stresses (Zahran, 1999). It is already known for 40 years that drought stress negatively affects nitrogen fixation as well as nitrogenase activity. Pimratch *et al*. (2008) reported that biological nitrogen fixation and plant biomass declines under water stress. A correlation was found between biomass of the plant and nitrogen fixation suggesting that if nitrogen fixation is maintained during drought a higher yield in the plant can be expected.

Drought has an effect on the regulation of cysteine proteases expression (Groten *et al*., 2006). Sprent (1972) reported that drought affects azo-caseinolytic activity in the winter wheat *Triticum aestivum* L. If the plant recovers from drought stress, the cysteine protease activity is reversed. Cysteine proteases that are induced by drought can function as an adaptation protein to the stress condition or can have other functions like activating specific proteins by proteolysis or protein denaturisation (Jones and Mullet, 1995). Cysteine protease activity generally increases during drought, but exceptions exist. Simova-Stoilova *et al*. (2010) observed that protein loss in wheat is irreversible after drought treatment by withholding

irrigation for seven days from eight-days old seedlings planted in the same pot. When investigating the protease WCP2 using a RT-PCR approach, the protease was suppressed during drought in drought-tolerant cultivars. In less tolerant cultivars, WCP2's expression remained constant (Simova-Stoilova *et al*., 2010). Khanna-Chopra *et al*. (1999) showed that in cowpea, not watered for 7 days, an increase in papain-like polypeptides were detected by Western blotting using polyclonal antiserum raised against papain. Three of these detected papain-like proteases appeared only after drought treatment. Further, two cysteine proteases investigated in our group in tobacco had different sensitivity to drought treatment. The cysteine protease *Nt*CP1 expressed in senescent leaves was not induced by drought whereas the cysteine protease *Nt*CP2, expressed in mature leaves, was expressed after withholding water for 10 consecutive days (Beyene *et al*., 2006)

A number of cysteine proteases are induced by abiotic stress associated with induction of nodule senescence. One such protease is *Pisum sativum* drought-induced *Ps*Cys15a which is a cysteine protease up-regulated during the symbiosis with *Rhizobium* in pea nodules. The first target of this protease in the root nodules are cytosolic proteins, such as leghemoglobin, and degradation of the bacteriod. The expression of this protease was also found in senescing alfalfa nodules (Vincent *et al*., 1980).

Cystatins are also affected by drought in different ways. A multi-cystatin, VuC1, identified in cowpea, had higher expression after drought when analysed by Northern blotting (Diop *et al*., 2004). A cystatin from *Brassica napus* (*Bn*D22) was also induced by drought treatment (Downing *et al*., 1992). Massonneau *et al*. (2005) found five cystatins, namely CII, CC3, CC4, CC5 and CC9, to be repressed by drought in maize.

1.4 Problem statement and research objectives

Although components of cysteine protease-cysteine protease inhibitor system have been previously studied in legume nodules, there is still a lack of detailed knowledge about the function of this system particularly in crown nodules during nodule development and in stressed nodules. For ultimate characterization of functionality of this nodule system, this study was carried out to determine expression of various components of the system during nodule development as well as to test if expression of these components is changing in nodules after drought treatment which causes premature senescence. The particular focus of this study was on the analysis of cysteine protease and cystatin expression in crown nodules with the specific objectives to first relate crown nodule performance during development with a total protease activity profile and secondly to detect cysteine proteases expressed during crown nodule development which can be tagged with cysteine protease inhibitor. A third specific objective was to establish a transcription profile of cysteine proteases (papain- and legumain-like) and cystatins during crown nodule development and after drought stress to identify any uniquely expressed proteases or cystatins. A fourth and final objective was to identify any changes in cysteine protease activity in crown nodules due to drought treatment and relate activity changes to changes in cysteine protease transcription after drought treatment.

2. Materials and Methods

2.1 Plant material, growth and treatment

2.1.1 Plant material

Commercial soybean seed (*Glycine max* L. Merr.; cultivar Prima 2000) were obtained from Pannar Seed (Greytown, South Africa). Seeds were inoculated with *Bradyrhizobium japonicum* (strain WB 74-1) powder at 109 CFU g^{-1} (Soygro bio-fertilizer Limited, South Africa). Seedlings were grown in pots [17.5 cm x 20 cm diameter (top) and 13.1cm] in fine-grade vermiculite (Mandoval PC, South Africa). Plants were grown under natural light conditions which were extended with artificial lights up to a 13 hr photoperiod at 600 mM $m^{-2}s^{-1}$ photosynthetically active radiation (PAR) and 60% relative humidity in an environmentally controlled green-house at 25°C / 16°C day / night temperature. Plants were watered twice a week with de-ionised water and three-times a week with a nitrogen-free Hoagland solution to obtain nodule formation. Pots were re-arranged periodically to allow uniform exposure to light in the growth chamber.

2.1.2 Plant material collection

For sample collection, eight plants (for each time point) were harvested after 4, 6, 8, 10, 12, 14, 16 and 18 weeks. Crown nodules were collected over this period and immediately stored after each harvest in separate falcon tubes in a -80°C freezer until the samples were used for further analysis.

2.1.3 Nodule activity, fresh mass and nodule number determination

Shoots, roots and nodules from three soybean plants were collected at 4, 6, 8, 10, 12, 14, 16 weeks and weighed on an AB104-5 Mettler-Toledo balance to determine the fresh mass (FW). For determination of nodule number, both crown and lateral nodules were removed from the roots by hand and counted. Crown nodules were dissected in to halves with a blade. A pink color in the nodule indicated an active nodule whereas as greenish or brown color indicated a senescent crown nodule.

2.1.4 Exposure of plants to water deficit

To evaluate the effect of drought, 30 plants were grown for two months until the third trifoliate leaf vegetative growth stage. Half (15) of all grown plants were then exposed to drought by completely withholding water for 10 days. Drought for the rest of this dissertation refers to water deficit. Water deficit experienced in a pot differs greatly from that of a field of soybeans. This can been seen as a catastrophic water stress due to the fact that the plant will be able to extract water with ease from the vermiculite to were no water is available unlike as in the field. Control plants were further watered every second day using a nitrogen-free Hoagland nutrient solution. After 10 days, plant material (plants were nine weeks of age) from both well-watered and drought-treated plants was harvested and samples were kept at -80°C for further analysis.

2.2 Protein studies

2.2.1 Protein extraction from nodules

Collected crown nodules of four plants were separately mixed with liquid N_2 and ground into a fine powder using a mortar and pestle. The powder was dissolved in 1 ml extraction buffer (50 mM Tris-HCl, pH 8.0). The suspension was centrifuged at 13 000 rpm for 10 min at 4°C using an Eppendorf centrifuge. The resulting protein-containing supernatant was stored for further analysis at -80°C.

2.2.2 Protein determination

The protein content of nodule extracts was determined using a commercial determination kit (Bio-Rad, UK) based on the method described by Bradford (1976). Bovine serum album (BSA; Sigma, South Africa) were used to set up a standard curve. A blank sample was prepared by using a 50 mM Tris-HCl, pH 8.0 buffer instead of a protein solution. Reactions were carried out on a micro-titre plate incubated for 30 min at room temperature. All reactions were performed in triplicate for both biological and technical replicates. A BMG FluoStar plate reader (Germany) was used to determine the colour reaction at 595 nm and three technical replicates were carried out for each sample.

2.2.3 Enzymatic assays

2.2.3.1 *Nitrogenase activity*

For determination of symbiotic N_2 fixation activity, plants were removed from pots and all crown and lateral nodules were carefully harvested. Nodules were placed small air-tight Erlenmeyer flask with a 43 ml capacity. An estimate of 4 ml acetylene was injected into flasks and after a 10 min incubation period ethylene production was measured by extracting 1 ml of gas from the headspace of each flask in a gas chromatograph (Varian 3900; Varian Inc., USA) according to the method described by Turner and Gibson (1980). A flame ionization detector was used with a oven temperature of 80°C (50s) → 160°C (20°C/min), 1177:180 ^{0}C and a gas flow of : air (300), H_2 (30), N_2 carrier gas (25) and running time was 4.8 minute**.** For calibration a standard curve were made by injecting ethylene from 100 to 2000 ml concentration in to the gas chromatography and use for estimation of acetylene reduction.

2.2.3.2 *Protease zymography*

Activity gel electrophoresis for protease activity measurement in nodules was based on a method described by Michaud and Asselin (2005). A 15% (w/v) sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE) according to the method of Laemmli, (1970) containing 1% (w/v) of gelatin. An amount of 50 µg of protein was loaded onto the gel after mixing with a 5X sample buffer (31.25 ml Tris-HCl, pH 6.8, 1% (w/v) SDS powder, 25 ml 50% (v/v) glycerol, 75 µl of 1 mM bromo-phenol blue dissolved in 2% (v/v) ethanol and 5 µl of 2-mecraptoethanol). A Bio-Rad (UK) electrophoresis system was used with a running buffer

consisting of 0.038 M glycine, 0.1 M of Tris-HCl, pH 8.0, and 1% (w/v) of SDS. The gel was run at 100V on ice for 2 hrs.

After the run, the gel was incubated with 2.5% (v/v) Triton X-100 for 30 min at room temperature for protease re-naturing. The gel was then rinsed three-times for 10 min each in distilled water to remove the excess Triton X-100. Protease activity was allowed to occur by incubating the gel overnight in a proteolytic buffer (0.1 M citrate phosphate buffer, pH 6.0, containing 10 mM L-cysteine) at 37°C. The gel was then transferred to a staining solution (25% isopropanol, 10% acetic acid, 0.5% coomassie blue) overnight and then de-stained for 2 hrs with a de-staining solution (10% acetic acid, 40% dH_2O and 50% methanol). Clear (proteolytic) bands indicating protease activity was captured using a digital camera.

2.2.3.3 *Assay of protease activity*

The activity of different types of proteases was measured using their specific substrates. For papain-like cysteine proteases, the cathepsin L substrate Z-Phe-Arg-MCA was applied (Sigma-Aldrich, Germany). For detection, 2 µg of protein was added to 95 µl reaction buffer (50 mM sodium phosphate buffer, pH 6.0, containing 5mM L-cysteine). After mixing, 1 µl of Z-Phe-Arg-MCA substrate solution was added. Trypsin-like proteases were measured using 0.1 M Tris-HCl (pH 8.0) and substrate Z-Arg-MCA, subtisilin-like proteases in 0.1 M Tris-HCl buffer (pH 7.5) and substrate Z-Ala-Ala-Phe-MCA, and legumain-like proteases in a buffer containing 39.5 mM citric acid, 121 mM Na_2HPO_4 (pH 7.0), 1 mM dithiothreitol (DTT), 1 mM EDTA, and 0.01% CHAPS (3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate) with substrate Z-Ala–Ala-Asn–MCA (Bachem, Switzerland). All substrates were dissolved in DMSO at a concentration of 10 mM.

Hydrolysis of the fluorogenic substrate was monitored by measuring the released α-amino 4-methylcoumarin (MCA) that is able to fluoresce, using a hydrolysis progress curve as described by Salvesen and Nagase (1989). Hydrolysis was monitored at 25^{o}C with a spectro-fluorometer (BMG FluoStar Galaxy) with excitation and emission at 340-10 nm and 450-10 nm, respectively. The reaction was measured over a period of 20 min. Reaction rates represented by the slope of the curve were measured as Fluorescence Units (FU/min). All reactions were performed triplicate for both biological and technical replicates. The cysteine protease inhibitor E-64 (10 µM) was added to the reaction mixture (10 µl) to inhibit cysteine protease activity. For drought treated plants, the percentage inhibition (percentage decrease) was determined by adding 100 uM E-64 and 10mM of TLCK into the buffer used to measure papain-like and legumain like proteases. Slope values were determined before addition and after addition of the inhibitor so that the percentage inhibition could be calculated.

2.2.4 Detection of expressed cysteine proteases

For labeling and detection of DCG-04 bound cysteine proteases with a streptavidin-based detection system, the method described by Martinez *et al*. (2007) was applied. Five plant samples were pooled. Total nodule protein was extracted by homogenizing crown nodules in distilled water on ice followed by centrifugation of extracts at 13 000 rpm for 15 min at 4^{o}C in an Eppendorf centrifuge. One volume of the protein supernatant was then mixed with three volumes of DCG-04 (5 µM) dissolved in 25 mM sodium-acetate buffer, pH 5.0, containing 10 mM cysteine, pH 5.0. The mixture was agitated at room temperature for 5 hrs. To control the specificity of the DCG-04 labeling, a competition experiment with E-64 was carried out on aliquots of the protein samples. The mixture was incubated at room temperature under gentle shaking on a shaker for 5 hrs. Chilled (-20^{o}C) acetone was then added to the mixture to

precipitate all proteins which were re-suspended in Laemmli's buffer (25 mM Tris-HCl pH 6.8, 2 % w/v SDS, 10 % v/v glycerine) and boiled for 1 min before loading onto a 15% SDS–PAGE gel. After protein separation, proteins were electro-transferred onto a nitrocellulose membrane for immuno-detection after blocking with 10% (w/v) nonfat milk, dissolved in PBS-T (8.0 mM potassium phosphate buffer, pH 7.4, 150 mM NaCl, and 0.02% v/v Tween-20). DCG-04-labelled proteins were detected with a streptavidin–alkaline phosphatase conjugate (1:2000 v/v, Ultrasensitive Streptavidin–Alkaline Phosphatase conjugate, Sigma Chemical Corp., St Louis, MO, USA) and an alkaline phosphatase detection kit.

2.3 DNA and RNA based studies

2.3.1 RNA extraction

RNA was extracted from all plant materials collected using the Qiagen RNeasy® plant mini kit (Qiagen, UK) according to the manufacturer's protocol. Tissues of four plants were pooled for RNA extractions. Extracted RNA was dissolved in 30 µl of diethylpyrocarbonate-(DEPC)-treated water. DNA contamination present in the RNA was eliminated by adding 3 µl of DNase 1, 4 µl of 10X reaction buffer (Fermentas, Canada) and 2 µl of RNAsin® Ribonuclease inhibitor (Promega, USA). This reaction was incubated at 37°C for 30 min after which the reaction was stopped by adding 1 µl of EDTA to each of the tubes followed by incubation at 65°C for 10 min. All samples were placed onto ice for cooling down.

2.3.2 Determination of RNA quality and quantity

RNA was visualized on a 1% agarose gel in 1X Tris-acetate-EDTA (TAE) buffer. RNA samples (2 µl), for gel electrophoresis were prepared in a 10 µl reaction volume containing 6X DNA loading dye (Fermentas, Canada) and gelred (Biotium, USA) for easy visualization of samples during the electrophoresis process. Samples were individually loaded into the wells of the agarose gel and the electrophoresis was carried out at 100 V for 30 min. Presence and quality of RNA was assessed under ultra-violet light. The amount and quality of RNA in each sample was also determined using the Nanodrop® Spectrophotometer (Fermentas, Canada) with 1 µl of RNA following the instructions provided by the manufacturer.

2.3.3 cDNA synthesis

Complimentary deoxyribonucleic acid (cDNA) was synthesized in a 38 µl reaction mixture using 2 µg RNA sample and a first-strand cDNA synthesis kit following the manufacturer's instructions (Promega, USA). The quality of cDNA was determined via a polymerase chain reaction (PCR) using a 40S ribosomal protein S8 gene from *Medicago sativa* (tentative consensus EST cluster TC100533 of MTGI from TGIR) (Van de Velde *et al*., 2006). Primers for this gene were named GM40S. For the reaction, 1 µl of the synthesized cDNA was used in a PCR reaction containing 2 µl of 10X PCR buffer, 1.6 µl of 25 mM $MgCl_2$, 2 µl of 2 mM deoxyribonucleotide triphosphates (dNTPs), 0.4 µl of primers (10 µM) and 0.2 µl of 5 units *Taq* polymerase. The final volume of the reaction was adjusted to 20 µl with sterile distilled water (sdH_2O). Reactions were placed into a thermal cycler (Biorad, USA) set to denature DNA at 94°C for 5 min followed by 39 cycles consisting of 94°C for 30 sec, 55°C for 30 sec for DNA annealing and 72°C for 1 min for DNA extension. Final elongation was then carried out at 72°C for 10 min. PCR samples were run on a 2% (w/v) agarose gel and DNA was stained with GelRed™ (Biotum, USA) added to 6X Loading Dye (Fermentas, Canada). Following verification of cDNA quality, cDNA concentrations were determined using the Nanodrop Spectrophotometer and adjusted to 200 ng/µl.

2.3.4 DNA primer design

Primers used to amplify cysteine proteases and cystatins were designed using the primer3 software (http://frodo.wi.mit.edu/primer3/). Selected primers were set to anneal between 50°C and 60°C with a primer length of between 18 and 22 nucleotides. Amplified products had a size below 500 bp. Selected primers are shown in Table 2.1.

2.3.5 Semi-quantitative RT-PCR

Semi-quantitative real-time polymerase chain reactions (semi-qRTPCR) were performed using between 200 to 400 ng/µl of cDNA and 25 to 46 cycles of amplification (Table 2.2). Reactions were set up in 20 µl reaction mixture. After amplification, the entire 20 µl reaction mixture was run on a 2% (w/v) agarose gel to visualize the amplification product with GelRed™ (Biotum, USA) added to 6X Loading Dye (Fermentas, Canada).

Table 2.1: Sequences for the forward and reverse primers used for different cysteine proteases and cysteine protease inhibitors. All Accession numbers are in accordance with the Phytozome database (http:www.phytozome.net/search.php?show=b;ast&method=org-Gmax).

Primer name	Accession number	Forward Primer	Reverse Primer	Amplicon (bp)
GmVPE1	Glyma17g14680	CTACGGAAACTACAGGCATC	GTTCTCCGTCGTCACATTAT	217
GmVPE2	Glyma05g04230	GGTCGTGGATGTTGCTGAGG	ATCTGCTTGATGCCTGTAGTTTCC	191
GmVPE3	Glyma14g10620	CACCATCCCTTGTAAATTGT	GGGGTTTCAGTGCATAATAA	247
GmCP1	Glyma04g04400	GATCTTTAATGGCCACGATCCTCAT	CAGCACCTTGAAAGGGGTAATCCT	678
GmCP2	Glyma17g05670	GCTTGTCACTGCTCATTTTCGC	TTTTCCGGTGTAGGGATATGC	687
GmCP3	Glyma10g35100	GAGGCCATGCCCTCATGT	TCACCTCTCTCCCCAGTGTAGG	716
GmCP4	Glyma14g40670	ATATGGAGCGTGTGACTCGG	GTAATATCCATTCTCTCCCCAGCTC	430
GmCP5	Glyma04g03090	AAGCTGTGGTGCATGTTGGG	AGTGGCGCTTGTCTTTGCAG	791
GmCC1	Glyma15g36180	TGCTGGAATTTGTGAAAG	GCCAATCAGTTTAAATTCC	125
GmCC2	Glyma14g04250	CACCGAAAGAGGATTAACAG	GGAGTTTGTGAGGGTGATTA	180
GmCC3	Glyma13g27980	GAAAAGAGATAGACCCGAAGGAT	TACAATCGCCTGATAATTGTTGCT	240
GmCC4	Glyma05g28250	GGATTAAAGCATACTAAACCA	GAATATTCGAATCCGTTGT	161
*Gm*40S	TC100533 (MGI)	GCCAGCCTGCTAACACTAAG	AAGAGTCTGAGTACGCACAAG	250

Table 2.2: Annealing temperature and number of cycles used for different primer pairs GM40S primers were used as a control for all primers at those primers annealing temperature and amount of cycles. All Accession numbers are in accordance with the Phytozome database (http:www.phytozome.net/search.php?show=b;ast&method=org-Gmax).

Primer	Accession number	Annealing temperature (°C)	Cycles	DNA (ng)
GmVPE1	Glyma17g14680	55	25	200
GmVPE2	Glyma05g04230	50	46	500
GmVPE3	Glyma14g10620	55	25	200
GmCP1	Glyma04g04400	63	27	400
GmCP2	Glyma17g05670	61	39	400
GmCP3	Glyma10g35100	64.5	39	400
GmCP4	Glyma14g40670	63	26	200
GmCP5	Glyma04g03090	62.4	46	400
GmCC1	Glyma15g36180	55	39	400
GmCC2	Glyma14g04250	50	30	400
GmCC3	Glyma13g27980	61	46	200
GmCC4	Glyma05g28250	55	30	400

2.3.6 Cloning of PCR products

For amplification product purification, bands were excised from the 2% agarose gel using the Qiagen gel extraction kit following the steps recommended by the manufacturer (Qiagen, UK). The extracted DNA was eluted with 20 µl of distilled water. A ligation reaction was set up using the pGEM®-T Easy vector (Figure 2.1) cloning kit (Promega, USA) following the steps recommended by the manufacturer with some modifications. A standard reaction was set up using 3 µl of PCR product and 1 µl of pGEM®-T easy vector which was added to a mixture of 5 µl of 2X Rapid Ligation Buffer containing 60mM Tris-HCl (pH 7.8), 20mM $MgCl_2$, 20mM DTT, 2mM ATP and 10% polyethylene glycol (Promega, USA) and T4 DNA ligase (Promega, USA). The mixture was incubated for 1 hr at room temperature. Transformation of bacterial cells was done using 5 µl of the ligation reaction with 50 µl of JM109 competent cells and incubation on ice for 20 min. The cells were then heat-shocked for 45 sec at 42°C and placed onto ice for 2 min. SOC medium (300 µl containing 20% tryptone, 0.5% yeast extract, 10mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$, 20 mM glucose, 10 mM $MgCl_2$) was added to cells and the mixture was incubated for 1.5 hrs at 37°C in an incubator under shaking at 200 rpm. Heat-shocked cells (100 µl) were then grown on LB plates (10 g tryptone, 5 g yeast, 10 g NaCl and 12 g agar, 1 ml of Isopropyl β-D-1-thiogalactopyranoside (IPTG), 1 ml of X-Gal solution (Fermentas, Canada), and 1 ml of ampicillin (100 µg/ml)). Plates were incubated over night at 37°C.

Putative transformed cells, which appeared as white colonies, were selected and grown by agitating overnight in 10 ml liquid broth with 100 µl of ampicillin at 37°C under shaking at 200 rpm. Plasmid isolation was done using the GenElute™ Plasmid Miniprep (Sigma-Aldrich, Germany) extraction kit following the manufacturer's instructions.

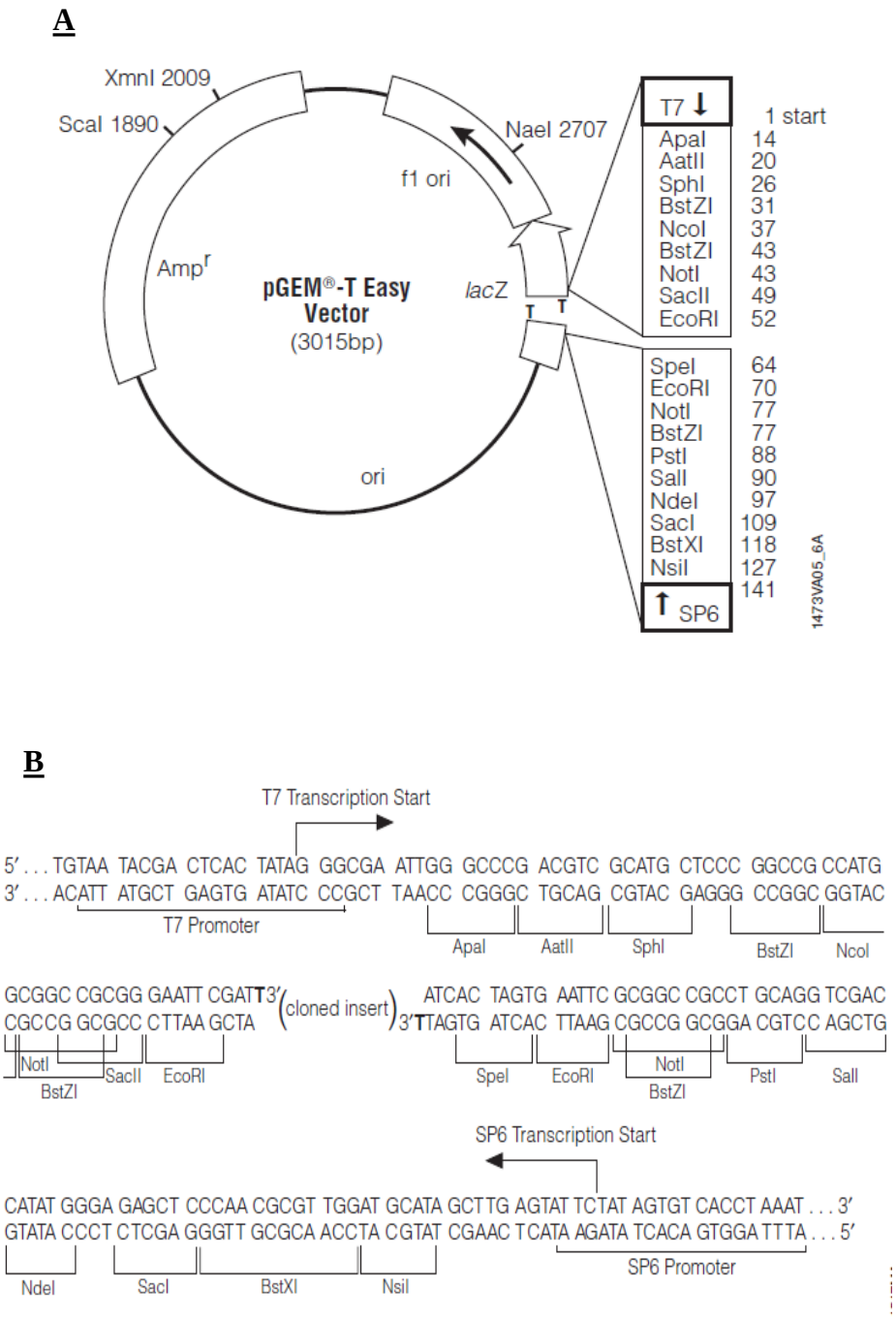

Figure 2.1: pGEM®-T Easy cloning system. a) The vector map of the pGEM®-T Easy cloning vector. b) The multiple cloning site and promoter region of the vector (Promega, USA)

2.3.7 Screening for positive transformants

The presence of the insert in cells was verified from isolated plasmid DNA using the FastDigest® EcoRI system (Fermentas, Canada). The digestion was carried out at 37°C for 30 min with 50-100 ng of DNA (plasmid extraction) incubated with 5 µl of FastDigest® EcoRI using a buffer recommended by the manufacturer (Fermentas, Canada). The presence of the insert was determined on a 1.5% agrose gel and the amplification product was visualized with GelRed™ (Biotum, USA) added to 6X Loading Dye (Fermentas, Canada).

2.3.8 Sequencing analysis

Sequencing of inserts was done at Inqaba Biotec™ (Pretoria, South Africa). The program, ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) was used to align sequenced products with gene templates for which primers were designed to confirm identity. All amplified sequences identity was verified through database comparisons using Blastn. (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAMS=blastn&BLAST_PROGRAMS=megaBlast&Page_type=BlastSearch&SHOW_DEFLAUTS=on&LINK_LOC=blastchrome). Sequences used are shown in Appendix 1.

2.4 Statistical methods

The statistical significance of the variations in mean was determined by using the GraphPad Prism5 software and a Student's t-test with a stringency level of $P \leq 0.05$.

3. Results

3.1 Soybean plant growth

Since the focus of the dissertation was on studying the cysteine protease – cystatin system in soybean crown nodules during nodule development, soybean seeds (cultivar Prima 2000) were inoculated with *Bradyrhizobium japonicum.* After germination seedlings were grown in vermiculite over a period of 16 weeks at 25^{o}C (Figure 3.1) in a temperature-controlled green-house under natural light with an extended light period (2 hrs. artificial light in the evening) to produce determinate crown nodules. Plants were watered twice a week with de-ionised water and three-times a week with a nitrogen-free Hoagland solution to obtain nodule formation.

Plants of different ages were first characterized regarding their productivity and nodule formation. Fresh shoot biomass increased over 16 weeks from 5.38 ± 0.95 g to 128.35 ± 15.92 g and fresh root biomass from 2.89 ± 0.75 g to 69.0 ± 21.78 g (Figure 3.2). Although fresh crown nodule biomass and crown nodule number steadily increased from 4 to 14 weeks with 0.34 ± 0.08 g to 1.14 ± 0.18 g and 22 ± 3 to 33 ± 5 nodules, respectively, both fresh nodule biomass and nodule number decreased after 14 weeks during pod-filling and flowering of plants to 0.93 ± 0.14 g and 29 ± 4 nodules, respectively, at 16 weeks (Figure 3.2).

Figure 3.1: Growth of soybean plants from 4 to 14 weeks after germination under green-house conditions (photos taken by A Mashamba, University of Pretoria).

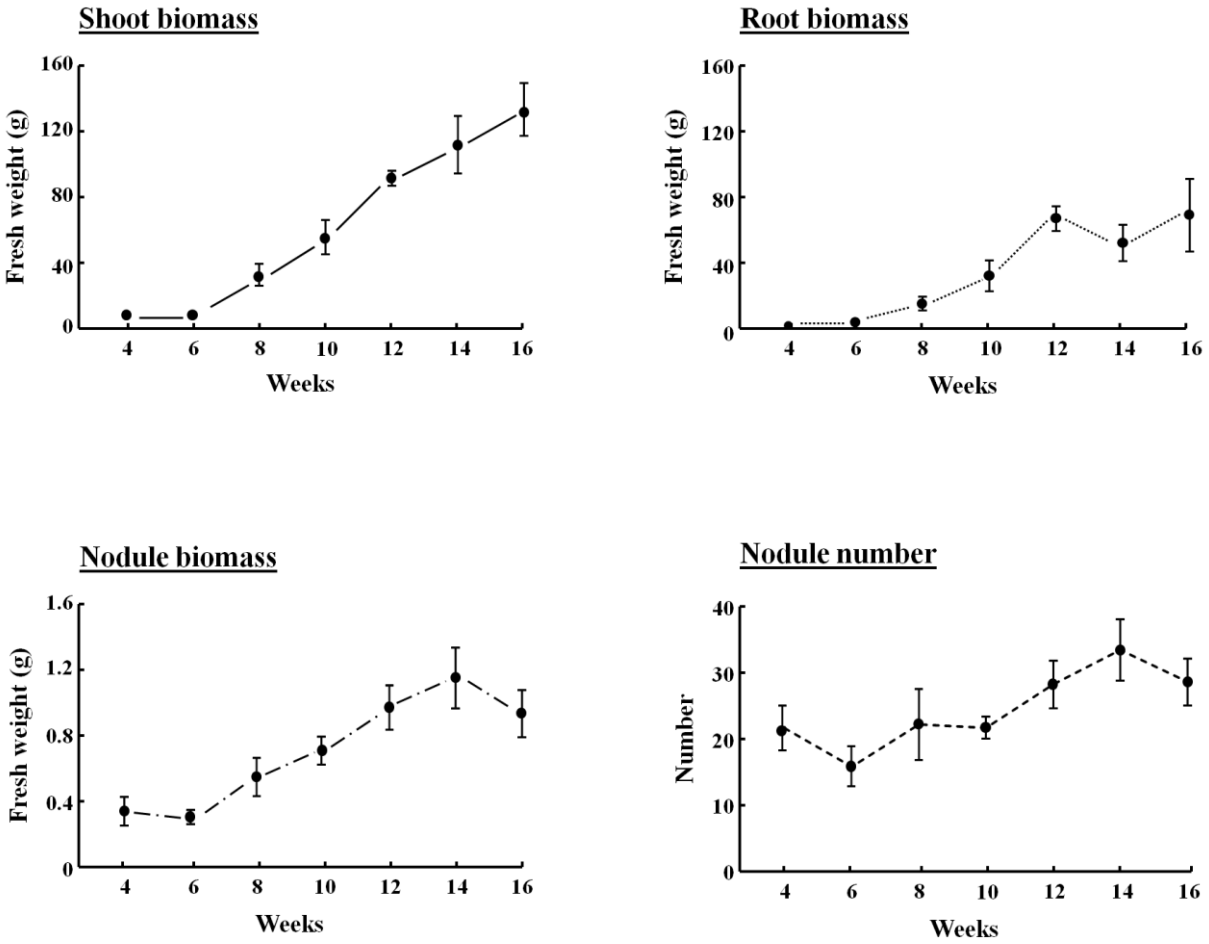

Figure 3.2: Production of fresh shoot biomass, root biomass, nodule biomass number of crown nodules over a period of 16 weeks. Data represent the mean ± SD of fresh mass and nodule number derived from three individual plants.

3.2 Analysis of soybean nodule enzymes

3.2.1 Nitrogenase activity during nodule development

Nitrogenase activity of different age crown nodules was determined using the acetylene reduction assay. Nitrogen fixation decreased in older senescent crown nodules. In crown nodules, nitrogenase activity was maximal at 6 weeks and then declined over time with 16 weeks old crown nodules having almost no measurable nitrogenise activity (Figure 3.3).

Nodules were dissected into two halves so that the colour change due to degrading leghemoglobin could be investigated. Leghemoglobin is responsible to deliver oxygen to the *Rhizobium* bacteria and has a pink colour if leghemoglobin is active, resulting in active nitrogen fixation as seen in 4 week old nodules (Figure 3.4). If a green/brownish colour is seen it indicates that the nodules have started to senesce (14wks and 18wks) and that nitrogen fixation is inactive due to degraded leghemoglobin.

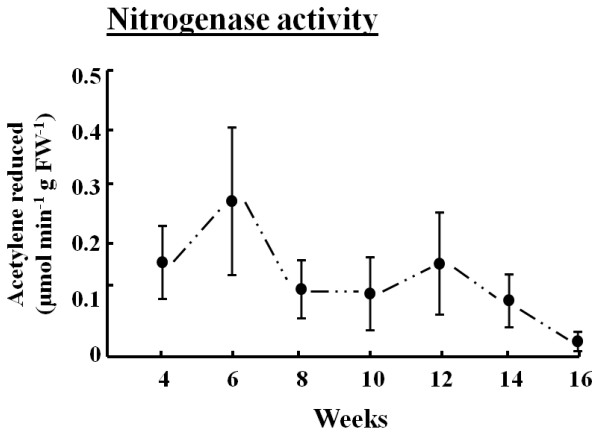

Figure 3.3: Nitrogenase activity of crown nodules measured as µmol acetylene reduced min^{-1} g FW^{-1} over a period of 16 weeks. Data represent the mean ± SD of activity derived from nodules of three individual plants.

Figure 3.4: Active nitrogen fixating crown nodules has a pink colour (4 wks) due to the heme cofactor of leghemoglobin and brown/greenish crown nodules (14 wks and 18 wks) are inactive nitrogen fixating senescent nodules. (Photo taken from Vorster *et al.*, 2013 - accepted)

3.2.2 Protease activity during nodule development

Since investigation of the expression of cysteine proteases and their inhibitors during crown nodule development was the major research interest, the overall protease activity profile of crown nodules was investigated. Different techniques were used to measure protease activity including a protease activity measurement on a mildly denaturing gelatine SDS-PAGE gel and measuring protease activity using various fluorogenic protease substrates in a fluorometric assay.

Total proteases activity was measured on a mildly denaturing gelatine SDS-PAGE gel where lighter, unstained bands, due to gelatine degradation, indicated the presence of protease activity. When soluble proteins were extracted from soybean crown nodules of different ages and total protease activity was determined, total protease activity greatly increased in 16 weeks old nodules (Figure 3.5). This increase was in contrast to younger, 4 to 12 week old nodules with rather low protease activity indicated by a faint activity band.

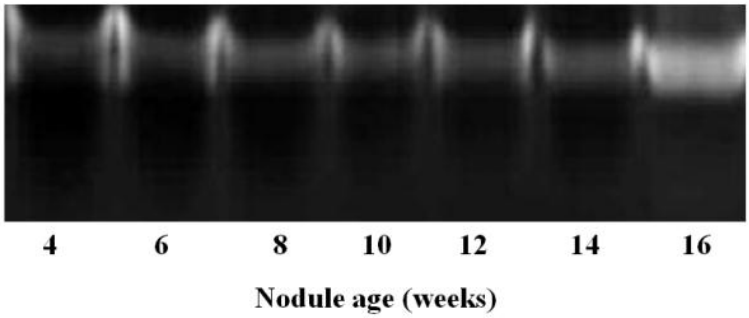

Figure 3.5: Total protease activity in crown nodules visualised on a mildly denaturing gelatine-containing SDS-polyacrylamide gel over a period of 16 weeks. Lighter, unstained bands indicate protease activity.

3.2.2.1 *Different classes of proteases activity during nodule development*

For determination of activity of papain-like cysteine protease during crown nodule development, cathepsin L-like cysteine protease activity was measured using the fluorogenic substrate Z-Phe-Arg-MCA. For comparison, activities of other types of proteases were also measured establishing a partial nodule protease profile during development (Figure 3.6). Fluorogenic substrates used were, Z-Ala–Ala-Asn–MCA for legumain-like cysteine proteases measured as well as Z-Ala-Ala-Phe-MCA for subtilisin-like and Z-Arg-MCA for trypsin-like serine protease activity. Since greatly increased subtilisin activity was only found in 18 weeks old crown nodules, activities of all proteases were measured for direct comparison in 4 to 18 weeks old crown nodules. Except for subtilisin activity, significantly higher ($P \leq 0.05$) activity of all tested proteases was found in both senescent, 16 and 18 weeks old nodules when compared to the activities measured in younger nodules (Figure 3.5). However, higher cathepsin L-like protease activity was observed in younger, (4 weeks old) nodules when compared to 6 weeks old nodules even though activity in 4 week old nodules were still less than in 16 and 18 week old nodules.

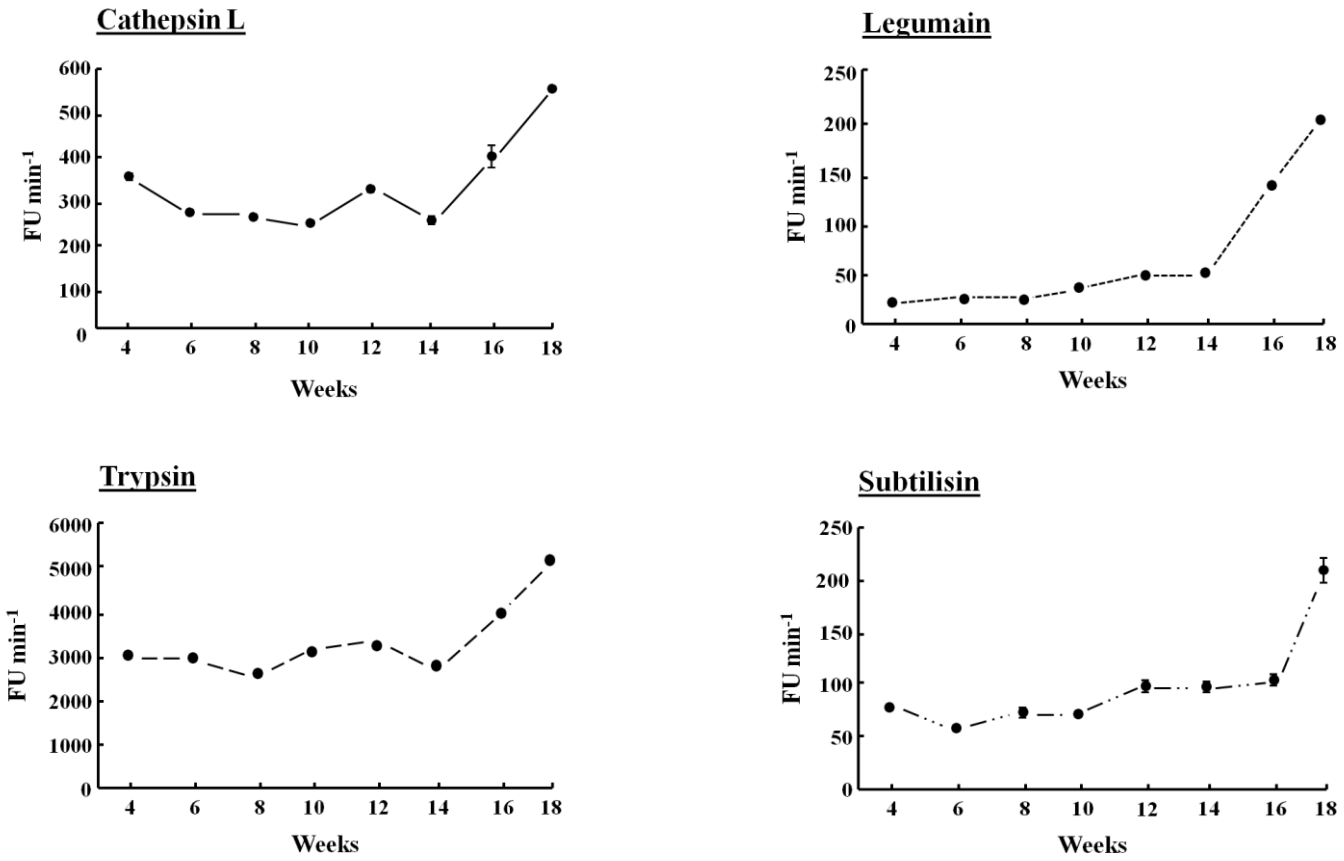

Figure 3.6: Determination of protease activity in crown nodules over a period of 18 weeks. Data represent the mean ± SD of three individual activity measurements with 2 µg of nodule protein extract added to each assay.

3.2.2.2 *Detection of expressed cysteine proteases*

To specifically detect the presence of expressed cysteine protease in extracts of crown nodules during nodule development, expressed cysteine proteases were also enriched by tagging them with the inhibitor DCG-04. Tagged proteases were separated on a SDS-polyacrylamide gel and finally blotting tagged proteins onto a membrane for detection of cysteine proteases with a streptavidin-based detection system (Figure 3.8; Martinez *et al.*, 2007).

DCG-04 is an analogue of the irreversible inhibitor of cysteine protease E-64, preventing cysteine protease activity. DCG-04 carries a biotin residue allowing tagging cysteine proteases and detection of biotin carrying tagged proteins with peroxidase-labelled streptavidin (Figure 3.7). Several DCG-04 tagged protein bands with various sizes, very likely representing cysteine proteases, were detected in crown nodules during development. Most prominently was detection of a strong tagged band in 4 week old nodules with a molecular mass of about 25 kDa (Figure 3.8). Tagged proteins with a weaker detection signal were further found in 6 to 14 week old crown nodules with a general increase in intensity in older, senescent nodules (16 and 18 weeks old) with two prominent bands at about 25 and 30 kDa (Figure 3.7).

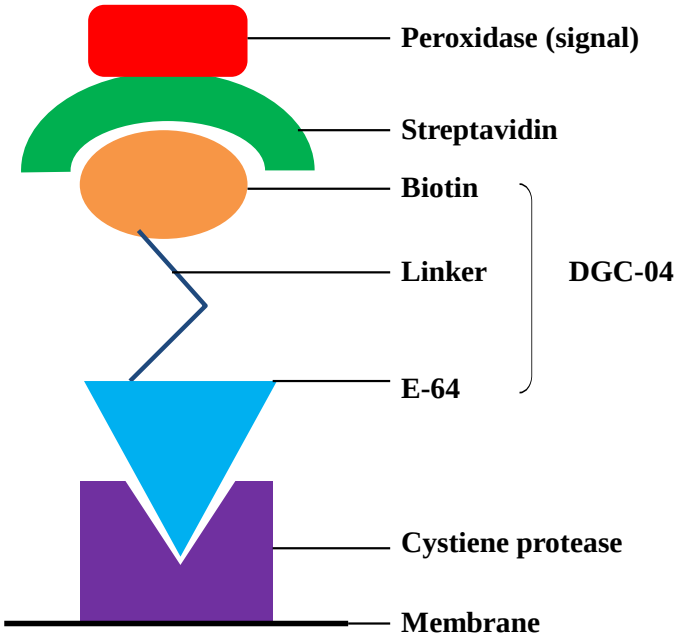

Figure 3.7: Streptavidin-based detection system for cysteine proteases where peroxidase-labeled streptavidin binds to biotin linked with DGC-04 which is an E-64 analogue. DGC-04 binds to the active site of the cysteine proteases allowing detection of DCG-04-bound cysteine proteases via peroxidase activity measurement.

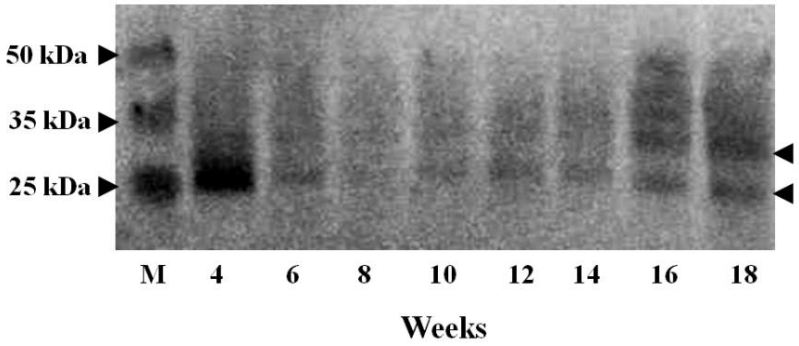

Figure 3.8: Detection of DCG-04-labelled cysteine proteases produced in 4 to 18 weeks old crown nodules. Arrows on the right indicate position of detected cysteine protease bands with major intensity in senescent nodules. M represents different size protein markers.

3.3 Transcription of cysteine proteases and cystatins

The semi-quantitative real-time PCR (RT-PCR) method was applied to determine transcription of various cysteine proteases in crown nodules of different ages. Sequences of papain-like and legumain-like proteases as well as cystatins were identified from RNA sequencing of 4, 8 and 14 week old crown nodules and found to be expressed in the crown nodules after transcriptome data analysis. Gene specific primers were designed for semi-quantitative RT-PCR analysis so as to establish the relative transcript abundance of different proteases and cystatins.

3.3.1 RNA isolation

Gel electrophoresis revealed the presence of suitable total RNA from various soybean tissues for cDNA production with no DNA contamination suitable for semi-quantitative RT-PCR technique (Figure 3.9). Table 3.1 outlines the concentration of total RNA isolated and cDNA synthesized from different tissue samples of plants including crown nodules harvested at either 4, 8 or 14 weeks (Table 3.1).

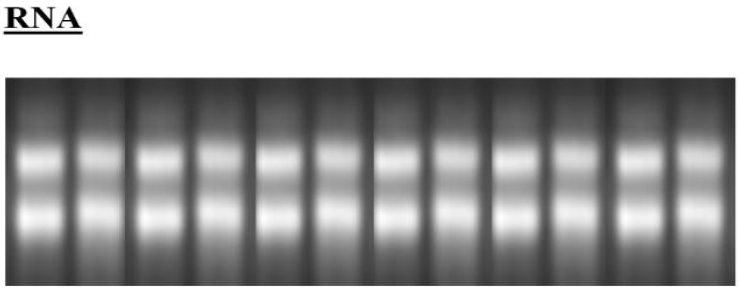

Figure 3.9: Assessment of purity of total isolated RNA used for cDNA synthesis on a 1% agrose gel run at 100V for 30 min.

Table 3.1: Concentration of total RNA isolated and cDNA synthesized from different tissue samples of soybean plants harvested at 4, 8 or 14 weeks and RNA and cDNA quantified by the Nanodrop® Spectrophotometer.

Organ	Weeks	RNA (ng/µl)	cDNA (ng/µl)
Crown nodules	4 weeks	910.1	1021.8
	8 weeks	1152.8	1008.5
	14 weeks	972.2	1009.5
Lateral nodules	8 weeks	921.3	1009.5
	14 weeks	621.0	1083.4
Root tips	4 weeks	514.1	1057.4
	8 weeks	623.6	1046.5
	14 weeks	388.2	1048.5
Young leaves	4 weeks	369.1	1078.5
	8 weeks	908.1	1053.0
	14 weeks	621.2	1099.0
Old leaves	4 weeks	2045.9	998.6
	8 weeks	199.6	756.4
	14 weeks	1018.2	1079.9
Pods	14 weeks	884.4	1081.9

3.3.2 Cysteine proteases and cystatin transcription profiles

Relative transcript abundance of cysteine proteases and cystatins in crown nodules from plants of different ages (4, 8 and 14 weeks old) was determined using semi-quantitative RT-PCR. To identify any gene sequence uniquely expressed in crown nodules, the presence of transcripts of cysteine proteases and cystatins were also determined for comparison in other tissues (Figure 3.10). Accession numbers provided in figures and tables are in accordance with the Phytozome (http:www.phytozome.net/search.php?show=b;ast&method=org-Gmax). All amplified sequences with selected primer pairs were sequenced, and sequence identity was verified through alignment and database comparisons using ClustalW and Blastn respectively.(http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAMS=blastn&BLAST_PROGRAMS=megaBlast&Page_type=BlastSearch&SHOW_DEFLAUTS=on&LINK_LOC=blastchrome). All the amplified products had a 100% homology to accessions used as templates (see appendix for sequences of cysteine proteases and cystatins).

None of the papain-like or legumain-like cysteine proteases investigated was uniquely transcribed in crown nodules. Except for Glyma10g35100 in older leaves, which showed no expression in 8 weeks, all papain-like cysteine proteases evaluated had either a higher transcript abundance in 4 and 8 weeks old tissues compared to 14 weeks old tissues, or had equal transcript amounts at all three time points (4, 8, 14 weeks) with no visible increase in transcript amounts in 14 weeks old tissues (Figure 3.11, Table 3.2). In pods, only papain-like cysteine proteases Glyma10g35100 and Glyma04g03090 were transcribed (Figure 3.10; Table 3.2).

Transcript amounts of legumain-like cysteine proteases (Glyma17g14680, Glyma05g04230 and Glyma14g10620) increased during crown nodule development. A similar increase was also found in root tips (Glyma17g14680). Glyma17g14680, Glyma05g04230 showed a constant expression in lateral nodules and root tips which is in contrast to the sharp increase in expression in crown nodules. Transcription of legumain-like cysteine proteases Glyma17g14680 and Glyma14g10620 was further detected in pods of 14 weeks old plants (Figure 3.12; Table 3.3).

None of the cystatins tested was specifically transcribed in crown nodules. A higher amount of cystatin transcripts (Glyma14g04250 and Glyma05g28250) was detected in 14 weeks old crown nodules when compared to the two other cystatin transcripts tested (Glyma15g36180 and Glyma13g27980). Further, transcription of cystatin Glyma14g04250 increased during root tip development and showed progressively increased expression in leaves matured from 4-14 weeks (Figure 3.13 and Table 3.2). Transcription of cystatin Glyma05g28250 was also found in root tips of 14 weeks old plants as well as in old leaves. Table 3.2 provides a general overview about the level of transcription in different tissues at specific time points.

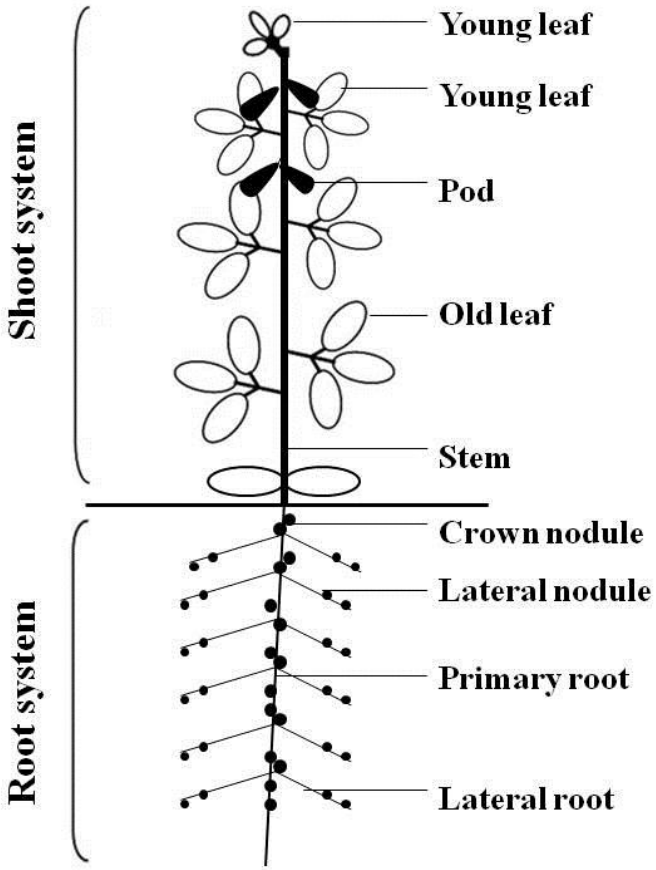

Figure 3.10: Morphology of a soybean plant showing the above ground shoots system and the below ground root system with different types of nodules which have been used for semi-quantitative RT-PCR.

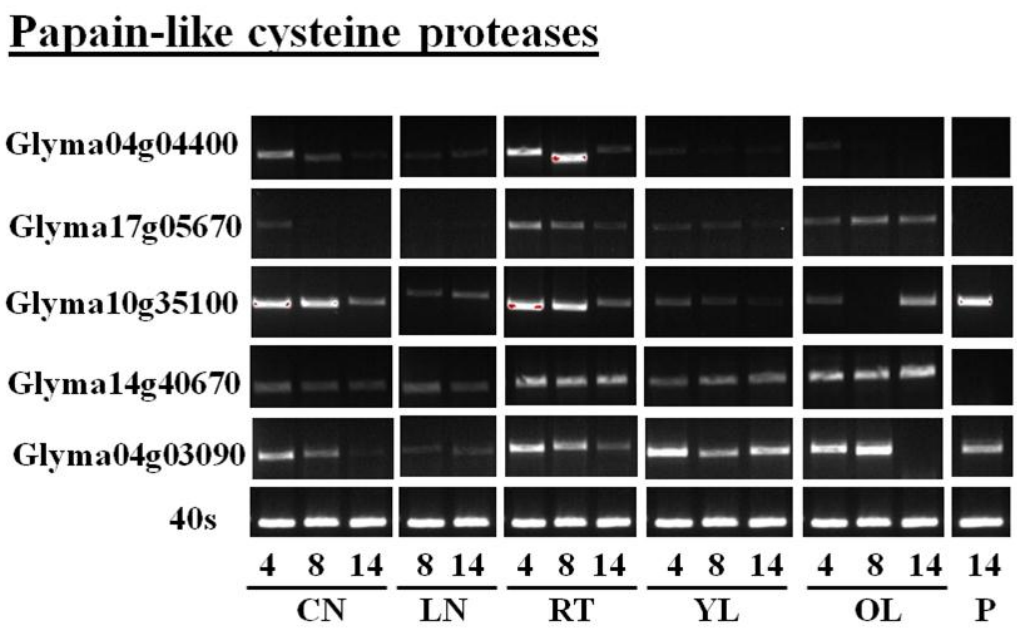

Figure 3.11: Transcription profiles of five papain-like cysteine proteases in soybean tissues of plants of different age (4, 8 and 14 weeks) determined by semi-quantitative RT-PCR. Amplification of the 40S gene sequence was used as a control. CN represents crown nodules, LN lateral nodules, RT root tips, YL young leaves, OL old leaves, and P pods.

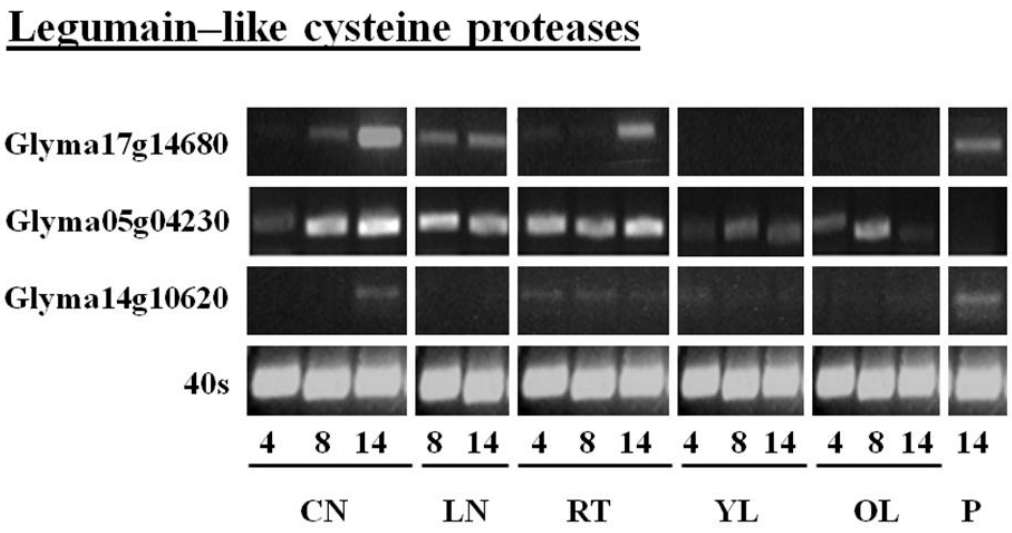

Figure 3.12: Transcription profiles of three legumain-like cysteine proteases in soybean tissues of plants of different age (4, 8 and 14 weeks) determined by semi-quantitative RT-PCR. Amplification of the 40S gene sequence was used as a control. CN represents crown nodules, LN lateral nodules, RT root tips, YL young leaves, OL old leaves, and P pods.

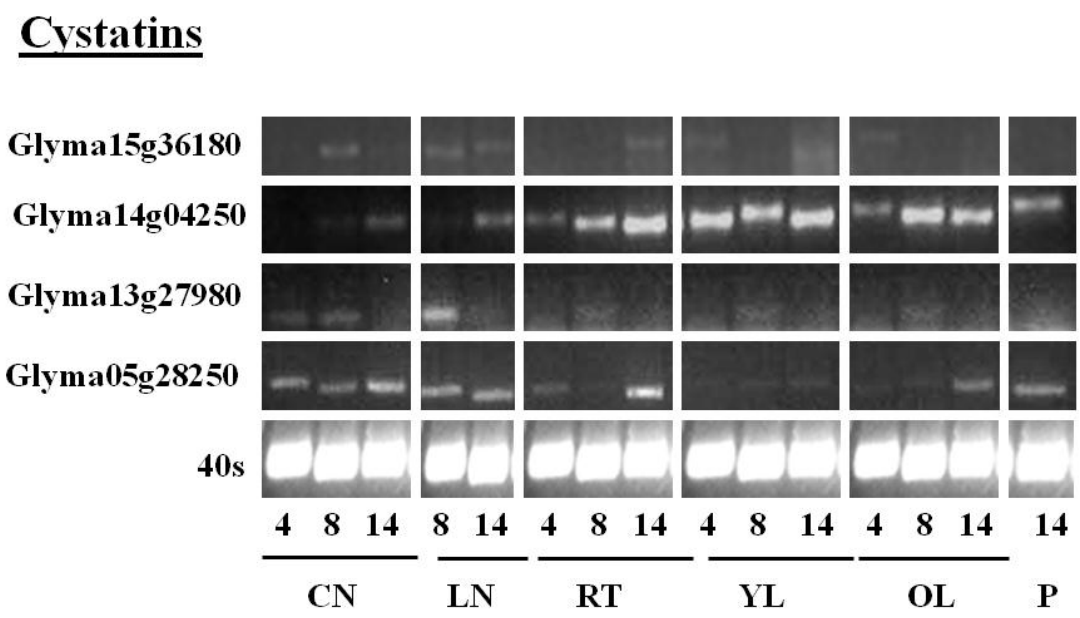

Figure 3.13: Transcription profiles of cystatins in soybean tissues of plants of different age (4, 8 and 14 weeks) determined by semi-quantitative RT-PCR. Amplification of the 40S gene sequence was used as a control. CN represents crown nodules, LN lateral nodules, RT root tips, YL young leaves, OL old leaves, and P pods.

Table 3.2: Transcription of cysteine proteases (CPs) and cystatins in tissues of plants 4, 8 and 14 weeks old. CN represents crown nodules, LN lateral nodules, RT root tips, YL young leaves, OL old leaves, and P pods. +++ represents strongest transcription and (+) weakest but still detectable transcription.

Accession Number	CN			LN		RT			YL			OL			P
	4	8	14	8	14	4	8	14	4	8	14	4	8	14	14
Papain-like CPs															
Glyma04g04400	++	+		(+)	(+)	+++	+++	(+)							
Glyma17g05670	(+)					++	+	(+)	(+)			+	+	+	
Glyma10g35100	+++	+++	++	(+)	+	+++	+++	+	+	(+)		(+)		++	+++
Glyma14g40670	+	+	+	+	+	++	++	++	+	+	+	++	++	++	
Glyma04g03090	++	+	(+)	+	+	+++	++	+	+++	++	++	+++	+++		++
Legumain-like CPs															
Glyma17g14680		+	+++	+	+			++							++
Glyma05g04230	+	++	+++	+++	+++	+++	++	+++	(+)	+	+	+	++		
Glyma14g10620			+			(+)	(+)	(+)							+
Cystatins															
Glyma15g36180			+	(+)	(+)			(+)	+++	+++	+++				
Glyma14g04250			+	(+)	+	+	++	+++				+	+++	+++	++
Glyma13g27980	(+)		(+)	++											
Glyma05g28250	+	+	++	++	++	(+)		++		(+)	(+)		(+)	+	+

3.4 Drought treatment of soybean plants

Both cysteine protease activity and transcription as well as cystatin transcription were affected by drought which causes premature senescence. Experiments were carried out with 8 weeks old nodulated soybean plants exposed to drought conditions by withholding watering for 10 days and plants used for analysis are shown in Figure 3.14. Leaves, shoot and root fresh biomass of drought treated plants was significantly lower ($P \leq 0.05$) than in well-watered plants (Table 3.3). There was no significant difference ($P \leq 0.05$) between crown nodule fresh biomass of well-watered and drought treated nodules (Table 3.3).

Nodules were dissected into two halves. As explained previously a pink colour (Figure 3.15) indicates active leghemoglobin and active nitrogen fixation. It was clearly seen that well-watered nodules (Figure 3.15a) still had active leghemoglobin whereas drought treated nodules had a green/brownish colour (Figure 3.15b) which indicates denatured leghemoglobin in senescing nodules.

Well-watered Drought-treated

Figure 3.14: Soybean plants (8 weeks old) were either well-watered plants or exposed to drought by withholding water for 10 days.

A B

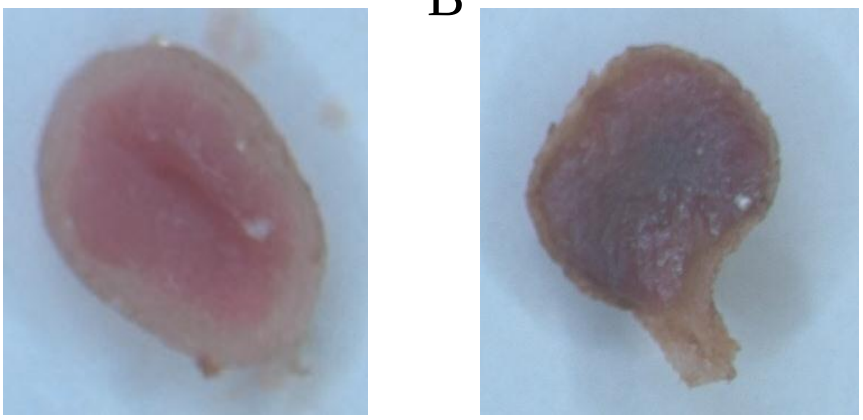

Figure 3.15: Cross section of crown nodules showing the pink colour (a) of active well-watered crown nodule indicating active nitrogen fixation due to the heme cofactor of leghemoglobin and a darkish green colour of a senescing drought treated crown nodule (b) indicating inactive nitrogen fixation due to denatured leghemoglobin.

Table 3.3: Fresh shoot, root and nodule biomass as well as nodule number of crown nodules produced over a period of 10 weeks. Water was withheld completely for 10 days before harvest of drought treated samples. Data represent the mean ± SE for biomass (fresh mass) and nodule number of three individual plants.

Characterization	Well-watered	Drought-stressed
Leaf biomass (g)	10.83 ± 0.61	8.71 ± 1.09*
Shoot biomass (g)	7.00 ± 0.69	3.18 ± 0.62*
Root biomass (g)	13.38 ± 0.61	9.53 ± 2.09*
Crown nodule biomass (g)	0.55 ± 0.17	0.41 ± 0.13
Crown nodule number	17.67 ± 5.13	15.67 ± 4.51

*Significant difference observed for well-watered samples to drought treated samples ($p \leq 0.05$).

3.4.1 Cysteine protease activity under drought

Both papain-like cysteine protease activity, measured as cathepsin L-like activity, and activity of legumain-like cysteine protease activity were measured in lateral and crown nodules either derived from well-watered or drought treated soybean plants. The cysteine protease inhibitor E-64 or the serine protease inhibitors TLCK was further added to enzyme assays to determine the particular contribution of E-64 and TLCK sensitive proteases to total protease activity. Protease activities were measured with either the fluorogenic substrate Z-Phe-Arg-MCA for cathepsin L-like cysteine protease or the substrate Z-Ala–Ala-Asn–MCA, for legumain-like cysteine proteases. E-64 is an irreversible cysteine protease inhibitor (Barrett et al., 1982; Rozman-Pungerčar, 2003) but the compound can also inhibit the serine protease trypsin (Sreedharan *et al*., 1996) whereas TLCK, a serine protease inhibitor, can also inhibit the cysteine protease papain (Wolters, 1969).

Drought treatment affected both total cathepsin L-like and legumain-like protease activity in lateral roots with a 24.7% decrease for cathepsin L-like and a 8.9% decrease in legumain-like protease activity (Table 3.4). A similar result was obtained for crown nodules with a 17.3% decrease due to drought treatment of total cathepsin L-like but a 2.5% increase of legumain-like protease activity (Table 3.4). Even though no significant differences could be seen in the protease activity in legumain-like proteases, cathepsin L-like proteases activity showed a significant ($P \leq 0.05$) decrease in both lateral and crown nodules. Addition of the protease inhibitor E-64 assays further decreased 18.0% total cathepsin L-like activity of lateral nodules derived from well-watered plants but only 0.9% of lateral nodules derived from drought treated plants. This showed a significant ($P \leq 0.05$), decrease in E-64 sensitive proteases activity in lateral nodules. A similar, decrease caused by E-64 addition was also

found for crown nodules derived from drought treated plants (7.1%) when compared to crown nodules from well-watered plants (13.2%). The same effect of E-64 on protease activity was, however, not found when legumain-like protease activity was measured. E-64 addition decreased more legumain-like activity in both types of nodules derived from drought treated plants when compared to nodules derived from well-watered plants (Table 3.4), even though only a significant ($P \leq 0.05$) increase in E-64 sensitive proteases activity were seen in lateral nodules.

Table 3.4: Percentage inhibition of cathepsin-L-like and legumain-like cysteine protease (CP) activities by E-64 in lateral and crown nodule extracts derived from well-watered or droughttreated soybean plants. All assays contained 1 µg protein and protease inhibitors concentrations was 100 µM E-64. Protease activity measured as fluorescence units (FU) min^{-1} µg $protein^{-1}$ and data represent the mean ± SE of three activity measurements.

Treatment	**Cathepsin L-like CP activity Inhibition (%)**		**Legumain-like CP activity Inhibition (%)**	
	Well-watered	Drought	Well-watered	Drought
Lateral nodules				
Non-treated	100 (1856.8±118.3 FU min^{-1})	100 (1397.9 ± 76.8 FU min^{-1})*	100 (164.6 ± 29.4 FU min^{-1})	100 (149.9 ± 4.5 FU min^{-1})
E-64	18.0 ± 7.5	0.9 ± 7.5*	15.7 ± 7.8	35.1 ± 2.5*
Crown nodules				
Non-treated	100 (1894.7 ± 98.3 FU min^{-1})	100 (1566.2± 23.9 FU min^{-1})*	100 (146.5 ± 44.7 FU min^{-1})	100 (150.4 ± 19.2 FU min^{1})
E-64	13.2 ± 12.2	7.1 ± 5.9	41.0 ± 12.2	47.9 ± 6.1

* Significant difference observed from well-watered to drought treated samples for non-treated and E-64 treated nodules.

3.4.2 Cysteine protease and cystatin transcription under drought

Semi-quantitative RT-PCR was used to determine if measured enzyme activities correlate with changes in cysteine protease transcription and which particular member of the cysteine protease gene family is induced or repressed. Also investigated was cystatin transcription under drought. Concentration of total RNA isolated and cDNA synthesized from different tissue samples of plants either exposed to well-watered or drought conditions are outlined in Table 3.5. No increase in the amount of transcripts of papain-like proteases was found in crown and lateral nodules after 10 days of drought exposure (Figure 3.16; Table 3.6). However, transcript amount of all papain-like cysteine proteases, except for Glyma14g40670, was higher after drought treatment in root tips. Drought also increased the transcript amount of Glyma17g05670 in young and older leaves as well as the transcript amount of Glyma10g35100 in younger leaves and of Glyma04g03090 in older leaves (Figure 3.14; Table 3.6).

When transcripts of legumain-like cysteine proteases were measured after drought treatment, Glyma17g14680 transcripts increased. This drought-induced increase was most prominently in both crown and lateral nodules, but also in root tips, when compared to nodules from well-watered controls (Figure 3.17; Table 3.6). Further, Glyma05g04230 transcripts were also higher in both crown and lateral nodules after drought exposure than in nodules of well-watered plants. Papain-like proteases showed a decrease in both activity and transcription in drought stressed plants. However legumain-like cysteine proteases also showed a decrease in protease activity but two legumain-like proteases selected showed a dramatic increase in transcription after drought treated plants.

The amount of transcripts of the cystatin Glyma13g27980 greatly increased in crown nodules, but also in young leaves, due to drought treatment (Figure 3.18; Table 3.6). Such increase in transcripts after drought exposure was further found for cystatin Glyma14g04250 in root tips and cystatin Glyma05g2825 in both crown and lateral nodules but most prominently in both young and old leaves.

Table 3.5: Concentrations of total RNA and cDNA synthesized from different tissue samples of plants subjected either to well-watered (WW) or drought (D) conditions and RNA and cDNA quantified by the Nanodrop technique.

Organ	Treatment	RNA (ng/µl)	cDNA (ng/µl)
Crown nodules	WW	965.9	1247.4
	D	646.7	1159.1
Lateral nodules	WW	1757.3	1393.2
	D	824.6	2480.2
Root tips	WW	350.7	13820.5
	DD	1117.5	1251.4
Young leaves	WW	1317.4	1287.7
	D	1538.6	3188.3
Old leaves	WW	350.7	653.2
	D	1108.6	2499.4

Table 3.6: Transcription of cysteine proteases and cystatins in crown and lateral nodules during development (14 weeks old) and after drought exposure determined by semi-quantitative RT-PCR. The Expression in 4 week old nodules is compared to expression in 8 and 14 week old nodules. Expression in drought treated samples is compared to expression in well-watered samples.↑ indicates up-regulation and ↓ down-regulation of gene expression after 14 weeks or drought.

Accession number	Gene	Crown nodules		Lateral nodules	
		Development	Drought	Development	Drought
Glyma04g04400	Papain-like CP	↓	↓	Unchanged	Unchanged
Glyma17g05670	Papain-like CP	No expression	↓	Unchanged	Unchanged
Glyma10g35100	Papain-like CP	↓	Unchanged	↑	↓
Glyma14g40670	Papain-like CP	Unchanged	Unchanged	Unchanged	↓
Glyma04g03090	Papain-like CP	↓	↓	Unchanged	↓
Glyma17g14680	Legumain-like CP	↑	↑	↑	↑
Glyma05g04230	Legumain-like CP	↑	↑	Unchanged	↑
Glyma14g10620	Legumain-like CP	↑	No expression	No expression	No expression
Glyma15g36180	Cystatin	↓	↓	Unchanged	↓
Glyma14g04250	Cystatin	↑	Unchanged	↑	Unchanged
Glyma13g27980	Cystatin	No expression	↑	↓	↑
Glyma05g28250	Cystatin	↑	↑	Unchanged	↑

Papain-like cysteine proteases

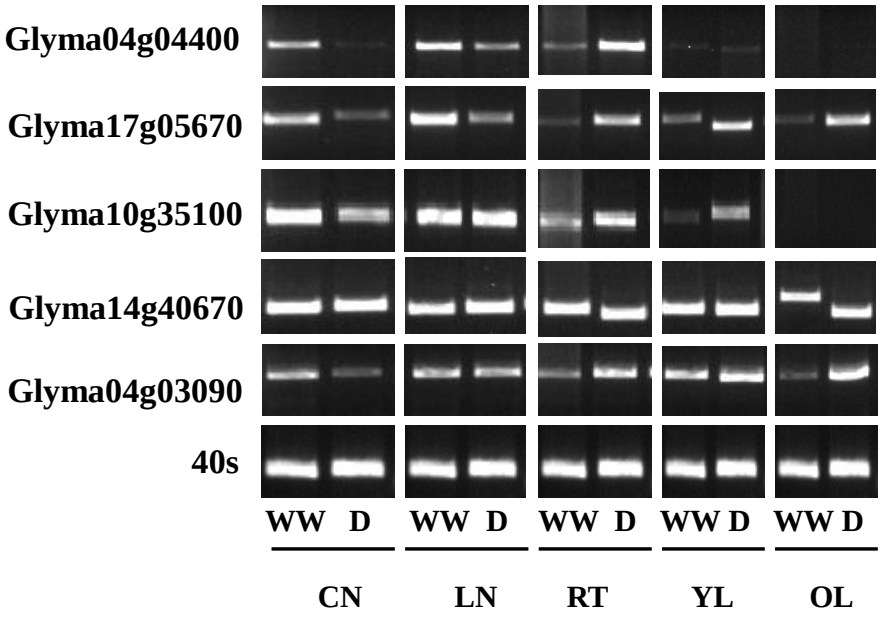

Figure 3.16: Transcription profiles of five papain-like cysteine proteases in soybean tissues that were nine weeks of age either exposed to well-watered (WW) or drought conditions where water was withheld for 10 consecutive days (D), determined by semi-quantitative RT-PCR. Amplification of the 40S gene sequence was used as a control. CN represents crown nodules, LN lateral nodules, RT root tips, YL young leaves, OL old leaves.

Legumain-like cysteine proteases

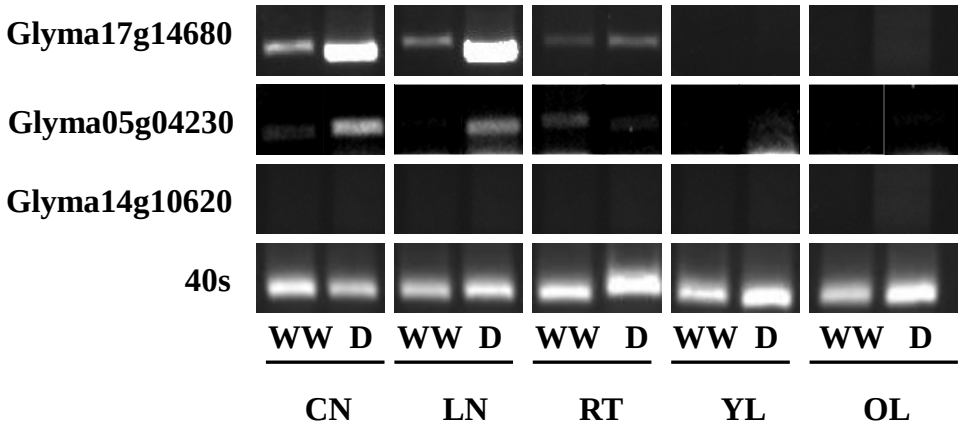

Figure 3.17: Transcription profiles of three legumain-like cysteine proteases in soybean tissues that were nine weeks of age, either exposed to well-watered (WW) or drought conditions where water was withheld for 10 consecutive days (D), determined by semi-quantitative RT-PCR. Amplification of the 40S gene sequence was used as a control. CN represents crown nodules, LN lateral nodules, RT root tips, YL young leaves, OL old leaves.

Cystatins

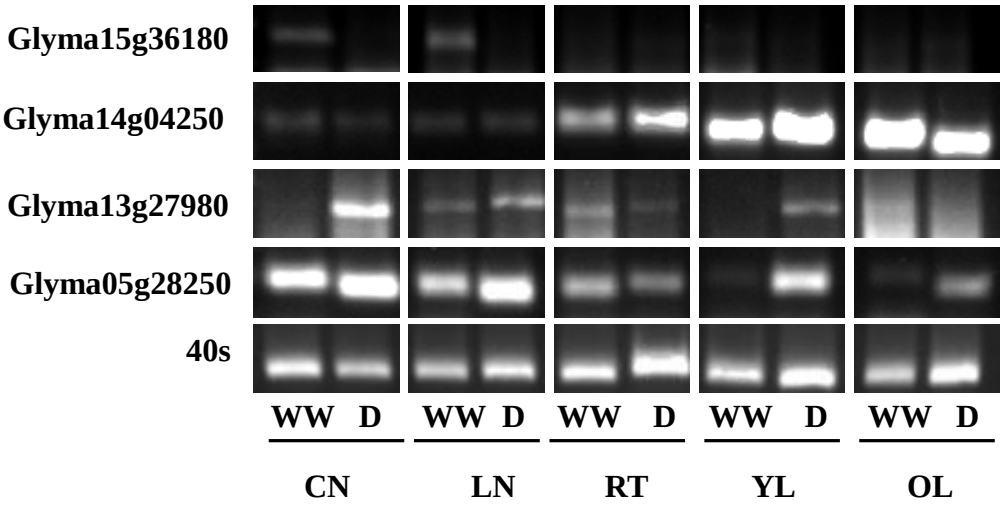

Figure 3.18: Transcription profiles of four cystatins in soybean tissues that were nine weeks of age, either exposed to well-watered (WW) or drought conditions where water was withheld for 10 consecutive days (D), determined by semi-quantitative RT-PCR. Amplification of the 40S gene sequence was used as a control. CN represents crown nodules, LN lateral nodules, RT root tips, YL young leaves, OL old leaves.

4. Discussion

A comprehensive characterization of the cysteine protease-cystatin system in soybean crown nodules during nodule development and after drought treatment was undertaken in this study. Although several researchers have investigated cysteine protease expression during development and after stress treatment in legume nodules, none of the studies investigated specifically papain-like or legumain-like cysteine proteases or investigated in detail their potential inhibitors, the cystatins. This study has therefore catalogued, as a novel aspect, both activity and transcript profiles of these two classes of cysteine proteases and transcription of their potential inhibitors.

Nodule senescence, and in particular the role of cysteine proteases causing protein degradation, has been previously investigated by several research groups with both bacteroids and nodule cells ultimately dying due to proteolytic activity (Vincent *et al*., 1980; Malik *et al*., 1981; Pladys and Vance, 1993; Kardailsky and Brewin, 1996; Vincent and Brewin, 2000). Results of this soybean study also confirmed that cysteine protease activity (cathepsin L-like activity) increases during nodule development, with maximal activity in senescent nodules. Increased protease activity further related to a decrease in nitrogenase activity, but also to a decrease in both fresh crown nodule biomass and number with any increase in both fresh root and shoots biomass levelling off. The crucial role of cysteine proteases in legume nodule senescence was previously demonstrated when anti-sense inhibition of CYP15A, a cysteine protease originally found in wilted pea shoots, delayed the senescence process in *Medicago truncatula* (Sheokand *et al*., 2005). Cytosolic proteins, including leghemoglobin, and degradation of the bacteriod are the first target of proteolytic activity in infected nodules limiting the bacterial symbiosis. In general, degradation of leghemoglobin is visible by a color shift in the nitrogen-fixing zone of nodules from pink to green due to heme-degradation (Roponen, 1970). Further, Chrispeels (1991) previously predicted also an open reading frame

for two cysteine protease genes involved containing a putative vacuolar targeting signal. This possibly directs these cysteine proteases to the bacteroid-containing symbiosome compartment assisting bacteroid protein degradation where, in particular, the peribacteriod membrane stability is affected causing the rupture of the peribacteriod membrane and the elimination of the microbial partner (Pladys and Vance, 1993). In many legumes, nodule senescence coincides with pod filling, (Lawn and Brun, 1974; Bethlenfalvay and Phillips, 1977) with nitrogen fixation and ammonia production affected during nodule senescence (Pfeiffer *et al*., 1983) allowing recovery of nitrogenous compounds from senescing tissues (Pladys and Vance, 1993).

In addition to an increase in cysteine protease activity in crown nodules, serine protease activities were also measured. Serine protease (subtilisin- and trypsin-like) activities increased in crown nodules during development when protease activity was measured with a fluorogenic substrate. An increase in serine protease (subtilisin-like) activity during nodule development has not been previously reported for soybean. Expressed cysteine proteases in soybean nodules also contained DGC-04 sensitive cysteine proteases. DGC-04, with a biotin residue added as a tag (Greenbaum *et al*,. 2000), and is an analogue of E-64 which is an irreversible cysteine protease inhibitor and also able to inhibit trypsin. E-64 forms a tight complex with cysteine proteases, in particular papain-like cysteine proteases, and acts as a pseudo-substrate by entering the active site of the target protease (Sreedharan *et al.* 1996). The biotin residue is conveniently used to detect DCG-04 bound proteases with a streptavidin-based detection system, as done in this study. Activity-based labelling with DCG-04 has identified senescence-associated proteases in *Arabidopsis thaliana* (van der Hoorn *et al.*, 2004). Several DCG-04 tagged protein bands with various sizes (25 kDa and 30 kDa) were identified in this study being expressed in senescent crown nodules. This confirms that

cysteine proteases are abundant during nodule senescence. Although cysteine proteases have further been found being expressed with very low proteolytic activity in younger nodules and further strictly compartmentalized to stop any proteolysis in fixing nodules (Pladys and Rigaud, 1982), a major DCG-04-tagged protein band was found in this study in young soybean nodules with a molecular mass of about 25 kDa. Function of this protease(s) in soybean crown nodules is still unclear and requires future investigation.

As part of a larger study, RNAseq was applied in our group as an alternative state-of-the-art method to study transcription of both papain and legumain-like cysteine proteases and cystatins in developing crown nodules. With the release of the assembly of the soybean genome it is now possible to identify all cysteine proteases and cystatins present in the soybean genome. A search carried out of the soybean genome in our research group revealed that there are up to 52 putative cysteine protease and 19 putative cystatin sequences in the soybean genome (Vorster *et al*,. 2013 –accepted). Five papain-like, three legumain-like cysteine proteases and five cystatins were identified being expressed during nodule development. In this soybean study, transcript amounts of three legumain-like cysteine proteases (Glyma17g14680, Glyma05g04230 and Glyma14g10620) increased during crown nodule development with Glyma17g14680 having the highest transcription of all cysteine proteases tested in senescent crown nodules. In contrast, none of the transcripts of the papain-like cysteine proteases tested increased when crown nodules aged. This is in contrast to the findings of Lievens *et al*. (2004) where a nodule specific papain-like protease *As*NODF32 isolated from *A. sinicus* (a winter growing Chinese milk vetch) is activated during nodule senescence with nodule lifespan prolonged and nitrogen fixation increased when *As*NODF32 was silenced. However, none of the two types of cysteine proteases and cystatins were uniquely transcribed in crown nodules or in any other type of tissue investigated (lateral

nodules, root tips, younger and older leaves or pods). Some of the cysteine proteases and cystatins were also not transcribed at all in these tissues, in particular in soybean pods. A microarray analysis of cDNA libraries of soybean nodules at different stages of nodulation showed that the transcript level increases in mature nodules and senescent nodules and a cysteine protease clone DD15 encoding a protein involved in remobilisation occurring during natural or stress-induced senescence shows high expression in senescent nodules (Alesandrini *et al*., 2003). Van der Velde et al. (2006) undertaking a comprehensive transcriptomics study to understand the onset of nodule senescence in the legume *Medicago truncatula* also reported that the most abundantly expressed genes in the nodule senescence dataset are four cysteine protease genes that are highly homologous to one of the prominent markers of leaf senescence, *sag12* (Lohman *et al.*, 1994). In this soybean study, detected legumain-like proteases, or VPEs, are cysteine protease processing vacuolar-based proteins during seed germination and leaf senescence (Müntz and Shutov, 2002). Certain types of VPEs, known as caspase-like proteases, are further very likely involved in PCD (Hatsugai *et al.*, 2004; Rojo *et al.*, 2004). They might also play a regulatory role during nodule senescence by activating other proteases through posttranslational maturation.

Nodule senescence is triggered prematurely by various types of stress (Gogorcena *et al.*, 1997; Gonzalez *et al.*, 1998; Matamoros *et al.*, 1999) with induced premature nodule senescence developing much more rapidly than developmental senescence involving oxidative stress (Puppo *et al*., 2005). Up-regulation of cysteine proteases under drought stress has been reported by Jones and Mullet (1995). The cysteine protease *Ps*Cys15a in *Pisum sativum* is up-regulated during the symbiosis with *Rhizobium* in pea nodules and is also drought inducible, but un-responsive to abscisic acid. In addition, Khanna-Chopra *et al*. (1999) found that in the legume cowpea, not watered for 7 days, papain-like polypeptides

increased which could be detected by Western blotting using a polyclonal antiserum raised against papain. Interestingly, in this soybean study the activity of E-64-inhibited proteases decreased in both crown and lateral nodules indicating that E-64-inhibited proteases are sensitive to drought stress which was further directly related to lack of up-regulation of any papain-like proteases under drought. In contrast, when transcripts of legumain-like cysteine proteases were measured in this study after drought treatment of crown nodules, Glyma17g14680 transcription (legumain-like cysteine protease) was strongly up-regulated comparable to up-regulation of this protease during nodule development. This clearly indicates an important role of legumains in nodule senescence.

No previous research endeavors have focused on investigating cystatin expression in soybean nodules after drought treatment. Drought induces methyl jasmonate signalling which is involved in diverse developmental processes, such as seed germination, root growth, fertility, fruit ripening and senescence. Jasmonate also induces cystatin production (Cheong and Choi, 2003; Munne´-Bosch and Penuelas, 2003; Dutt et al. 2011). Some cystatins investigated in this study were not up-regulated after drought treatment including cystatin Glyma15g36180. This cystatin was also not expressed at all in any of the nodules or tissues investigated after exposure to drought. Further, cystatin Glyma14g04250 was also not affected by drought treatment in both crown and lateral nodules. In this study, however, transcripts of the cystatin Glyma14g04250 and Glyma05g28250 (crown and lateral nodules) greatly increased in nodules. Glyma05g28250 was also strongly up-regulated after drought stress in young soybean leaves and its specific function in drought stress should be further investigated, in particular if endogenous cysteine protease activity during drought is controlled by this cystatin.

Overall, in this soybean study new information has been obtained about expression of two types of cysteine proteases (papain-like and legumain-like) during nodule development and after drought exposure. Despite previous reports on cysteine protease expression in legume nodules, the current knowledge about the function and expression in particular of the cysteine protease cystatin system is still scanty. Providing more detailed knowledge including expression profiles of the components involved in the system, as done in this soybean study, is crucial to advance the understanding of system regulation and function. For the first time this study has included information about expression of cystatins. Since RNAseq allows direct isolation of full-length cDNA sequences production of soybean nodules with decreased cysteine protease and cystatin activities is now possible to study the specific function of these components during nodule development and under stress or to develop them into a biomarker for pro-longed nodule life. Extending the period of active nitrogen fixation by delaying the nodule senescence process might have a beneficial effect on soybean yield and seed quality. Existence of soybean varieties with delayed nodule senescence and a longer active nitrogen fixation period already points to a genetic control of the onset of senescence, suggesting a molecular basis to modify timing and/or progression of developmental nodule senescence (Espinosa-Victoria et al., 2000).

Future actions

Some future actions proposed as an outcome of this soybean study include:

(1) A first immediate aspect to be investigated should focus on confirmation of the semi-quantitative RT-PCR results by real-time quantitative PCR to determine more precisely transcript amounts during nodule development and after stress. This might also include a

repeat of the development study to include time points at 16 and 18 weeks for transcript analysis to encompass the period of late nodule senescence.

(2) A second action might include screening various soybean cultivars for expression of cysteine proteases and cystatins with the aim to possibly develop a protease or cystatin-based biomarker for pro-longed nodule life or stress tolerance.

(3) Due to RNAseq data available, a third action might focus on using RNAi technology and transferring selected cysteine protease or cystatin coding sequences to *Rhizobium* cells for infection of soybean roots for protease and cystatin down-regulation.

(4) Since papain-like cysteine transcription was not directly related to cathepsin L-like activities and inhibitor tagging activities, a fourth action might focus on investigating if the highly transcribed legumain-like cysteine protease in both senescent and drought treated nodules is responsible for processing and activation of already accumulated papain-like cysteine proteases.

(5) Other aspects to be investigated might include applying, in addition to drought, a different type of stress, such as cold stress, known to induce plant senescence. Isolated cystatin sequences might also be used to transform *Arabidopsis* as a model plant to investigate whether these cystatins provide any protection against senescence or abiotic stress.

(6) The impact of reduced carbohydrates supply to the nodules should be investigated as it is still unclear if nodule senescence, induced by drought stress is caused by low energy availability to the nodules or weather the drought induced senescence is caused by disrupting the membranes and lowering oxygen availability to the nodules.

5. Appendix 1

5. Appendix 1

5.1 Papain-like proteases sequences

Glyma04g04400

ATGGGTACTAACAGATCTTTAATGGCCACGATCCTCATCGTGTTCTTCACGGTGTTGGCAGTGTCATCGGCGTTG
GACATGTCGATAATCTCGTATGACAGAAGCCACGCGGACAAGTCAGGGTGGAAGAGTGACGAGGAGGTAATGAGT
ATATACGAGGAGTGGCTTGTGAAGCATGGGAAGGTGTATAACGCCGTCGAAGAGAAGGAGAAGAGGTTTCAGATC
TTCAAAGACAACCTCAACTTTATCGAAGAACACAATGCTGTGAACCGGACATACAAGGTTGGACTGAACCGGTTC
AGTGATCTTAGCAACGAGGAGTACCGGTCAAAATACCTGGGAACCAAGATTGATCCCAGCAGAATGATGGCCAGA
CCAAGTAGACGTTACTCTCCGCGCGTGGCTGACAATTTGCCGGAATCTGTTGATTGGAGGAAAGAAGGTGCTGTT
GTCAGAGTCAAAAACCAATCAGAATGCGAGGGCTGCTGGGCATTCTCAGCAATTGCCGCAGTTGAAGGAATTAAC
AAGATAGTAACAGGCAATCTAACTGCACTGTCAGAGCAAGAACTGTTGGACTGTGATAGAACCGTCAATGCAGGA
TGTAGTGGCGGACTTGTGGACTATGCCTTTGAGTTCATTATCAACAATGGTGGCATTGACACTGAAGAGGATTAC
CCCTTTCAAGGTGCTGATGGTATCTGTGATCAATACAAGATAAATGCTAGAGCTGTTACAATTGATGGTTACGAA
CGTGTTCCTGCTTATGATGAATTAGCCTTGAAAAAGGCAGTAGCAAATCAACCCGTGAGCGTTGCCATTGAAGCA
TATGGCAAAGAGTTTCAATTATATGAATCTGGTATATTCACAGGAACATGTGGGACATCAATAGACCATGGTGTT
ACAGCTGTTGGGTACGGAACAGAAAATGGAATTGACTATTGGATCGTCAAAAATTCATGGGGTGAGAATTGGGGA
GAGGCAGGTTATGTAAGAATGGAACGTAATATAGCAGAAGACACTGCAGGAAAGTGTGGAATTGCGATTCTGACA
TTGTATCCCATAAAGATTGGCCAAAATCCCTCAAACCCAGATAACTCCAGCTAA

Glyma17g05670

ATGGCACGCTTGTCACTGCTCATTTTCGCGTTCTGCGCCGTCGCCGTCGCCGTCGCCGTCGCGGGCTCCTCCTTC
GATGATGCCAACCCTATCCGTTTAGCGTCGGACTTGGAGTCGCAGGTGCTGGACGTGATCGGACAGAGCCGCCAC
GCTCTCTCCTTCGCCCGCTTCGCTCGCCGCCACGGCAAGCGCTACCGTTCCGTCGACGAGATCCGCAACCGCTTC
CGGATCTTCTCTGATAATCTTAAACTCATCAGATCCACCAACAGGAGGTCTCTCACCTACACGCTCGGCGTCAAT
CATTTTGCTGACTGGACTTGGGAGGAGTTCACCAGACACAAGCTCGGCGCTCCTCAGAATTGCTCTGCCACGCTC
AAAGGCAACCACAGGCTCACCGATGCTGTTCTTCCTGACGAGAAAGACTGGAGAAAAGAAGGTATAGTCAGCCAA
GTTAAAGATCAAGGCAACTGCGGATCTTGCTGGACATTCAGCACAACTGGTGCATTGGAGGCAGCTTACGCACAG
GCCTTTGGGAAGAATATCAGTCTTTCTGAGCAGCAGCTAGTGGACTGTGCCGGTGCTTTCAATAACTTTGGCTGT
AATGGTGGTTTGCCATCCCAAGCCTTTGAATACATTAAATACAATGGTGGCCTCGACACAGAAGAAGCATATCCC
TACACCGGAAAAGATGGTGTCTGCAAATTTACAGCTAAAAACGTTGCTGTTCGAGTCATTGACTCTATCAATATC
ACTTTGGGTGCTGAGGATGAATTGAAACAAGCGGTTGCTTTTGTTCGGCCAGTTAGTGTGGCATTTGAGGTAGCG
AAGGACTTCCGATTCTACAATAATGGAGTTTACACTAGTACCATTTGTGGTAGCACGCCCATGGATGTAAATCAT
GCTGTTCTTGCTGTTGGGTATGGAGTTGAAGATGGTGTTCCATATTGGATCATTAAAAATTCATGGGGAAGCAAC
TGGGGTGACAATGGTTACTTCAAGATGGAATTAGGGAAGAATATGTGCGGTGTTGCAACTTGTGCATCTTATCCT
GTTGTGGCTTAA

Glyma10g35100

ATGGAGGCAAAGCGAGGCCATGCCCTCATGTGCTTGGCACGTGTTTCACTCTTCCTCTGCGCTCTAACCCTATCC
GCTGCTCACGGAAGCACCACGGTTCAAGACATCGCGAGAAAGTTGAAACTCGGGGACAACGAGCTTCTGCGCACG
GAGAAGAAGTTCAAGGTTTTCATGGAGAATTACGGGAGGAGCTACTCCACAGAGGAGGAGTATTTGCGCCGTTTG
GGGATTTTCGCCCAGAACATGGTCAGAGCGGCGGAGCATCAGGCGCTGGACCCCACCGCCGTCCACGGCGTCACG
CAGTTCTCCGATCTCACCGAGGACGAGTTTGAGAAGCTCTACACTGGCGTCAATGGCGGCTTCCCGTCATCGAAT
AATGCCGCCGGCGGTATTGCGCCGCCGTTGGAGGTGGACGGGCTGCCGGAGAATTTTGATTGGAGAGAGAAAGGA
GCTGTCACCGAAGTTAAGTTACAGGGTAGATGTGGATCATGCTGGGCTTTCAGCACCACAGGATCTATAGAAGGA
GCCAATTTTCTTGCCACCGGAAAGCTCGTCAGTCTTAGTGAACAACAGCTCCTTGACTGTGACAACAAGTGTGAC
ATAACAGAAAAGACTTCATGTGACAATGGGTGTAATGGAGGTCTTATGACAAATGCCTACAATTATTTGCTCGAG

TCTGGTGGGTTGGAGGAGGAGTCTTCATATCCCTACACTGGGGAGAGAGGTGAATGCAAGTTTGATCCAGAGAAA
ATAGCAGTTAAGATCACAAATTTCACCAATATCCCTGCTGATGAGAACCAAATTGCCGCATATTTAGTAAAAAAT
GGTCCACTTGCTATGGGCGTGAATGCAATTTTCATGCAAACATACATTGGTGGTGTCTCATGCCCTCTCATATGT
AGCAAGAAAAGGCTGAACCATGGTGTGTTACTTGTGGGATATGGTGCAAAAGGCTTCTCAATTCTTAGGCTTGGC
AACAAACCATACTGGATCATCAAGAATTCGTGGGGGGAGAAATGGGGAGAAGATGGCTATTACAAGCTCTGCAGA
GGGCATGGCATGTGTGGAATAAACACTATGGTTTCTGCTGCAATGGTGCCTCAACCACAAACAACACCTACAAAG
AATTATGCTTCTTATTAG

Glyma14g40670

ATGGCTCGCTACTCCTTCCTCTTGGCCCTCCTCCTCCTCCTCTTCGCCGCCGTGTCCAGCGCCGCCGACGACATC
CTGATCCGTCAGGTGGTGCCGGATGCCGTCGGCGAAGCGGCGGAGAAGGAGGAGGACCACCTGCTGAACGCGGAG
CACCACTTCGCGAGCTTCAAGGCGAAGTTCGGGAAGAAGTACGCGACCAAGGAGGAGCACGACCGTCGCTTCGGC
GTCTTCAAGTCCAACCTCCGCAGAGCGAGGCTGCACGCCAAGCTCGACCCCTCCGCCGTCCACGGCGTCACCAAA
TTCTCCGATCTCACTCCGGCCGAGTTCCGCCGCCAATTCCTGGGCTTCAAGCCGCTGCGCTTGCCGGCGAACGCT
CAGAAGGCACCCATTCTCCCCACCAAGGATCTCCCAAAGGATTTCGATTGGCGAGACAAAGGAGCCGTCACTAAC
GTCAAGGACCAAGGGGCGTGTGGTTCGTGCTGGTCTTTCAGTACCACGGGAGCGTTGGAAGGAGCTCATTATCTG
GCCACTGGTGAGCTTGTGAGCCTTAGTGAGCAACAGCTTGTGGATTGCGACCACGTGTGTGATCCAGAAGAATAT
GGAGCGTGTGACTCGGGCTGTAATGGTGGGTTGATGAACAATGCATTTGAGTACATACTTCAGTCTGGGGGAGTA
CAGAAAGAAAAGGACTATCCATACACGGGGAGAGATGGCACCTGCAAATTTGACAAAACCAAAGTTGCAGCTACT
GTATCTAATTACAGTGTGGTTTCCCTTGATGAAGATCAAATTGCTGCAAACCTAGTGAAGAATGGCCCTCTTGCA
GTTGGTATCAATGCAGTTTTTATGCAGACATATATAGGTGGCGTGTCGTGCCCGTACATCTGTGGGAAGCATTTG
GATCATGGTGTTCTTATAGTGGGTTATGGTGAAGGTGCATATGCTCCCATTCGTTTTAAGAACAAGCCTTACTGG
ATCATTAAGAATTCATGGGGTGAGAGCTGGGGAGAGAATGGATATTACAAGATTTGCAGAGGTCGAAATGTGTGT
GGAGTAGATTCCATGGTCTCAACTGTAGCTGCTATATATCCATCCAGCCATTAA

Glyma04g03090

ATGACTCGTTTATATTCCCTTTCTCTTTTACAATTCCTGTCTCTGATTCTCCTTTTCACACTTTTCTTCTTGTCT
GCCTCTGACACCTCCGAACTGTTCGAAAAGTGGTGTAAGGAACACAGCAAAACATACTCTTCCGAGGAAGAAAAA
CTGTACAGGCTCAAAGTGTTCGAGGACAACTATGCCTTTGTCGCCCAACACAACCAAAACGCCAATAATAACAAT
AATAATTCAAGCTACACTCTTTCCCTCAACGCTTTCGCCGATCTCACCCACCATGAATTCAAAACCACTCGTCTC
GGTCTTCCCCTTACCTTGCTTCGATTCAAACGTCCCCAGAATCAACAATCCCGTGACCTTCTGCACATTCCTTCT
CAGATCGATTGGAGACAGAGCGGCGCAGTTACACCCGTCAAAGACCAAGCAAGCTGTGGTGCATGTTGGGCGTTC
TCAGCTACTGGTGCTATTGAAGGCATAAATAAGATTGTCACAGGGTCTCTTGTAAGCCTTTCTGAACAAGAGCTG
ATAGATTGTGATACATCCTACAACAGTGGTTGTGGAGGTGGACTCATGGATTTTGCATACCAATTTGTCATCGAC
AACAAAGGGATTGACACTGAGGATGATTATCCGTATCAAGCTCGGCAACGGTCCTGCAGTAAGGACAAGTTAAAA
AGGCGTGCTGTTACAATTGAGGATTATGTTGACGTGCCTCCGAGTGAGGAAGAGATACTAAAGGCCGTTGCATCA
CAACCTGTAAGCGTAGGCATATGTGGTAGTGAAAGAGAATTCCAGTTGTACTCAAAGGGTATCTTCACTGGGCCA
TGTTCAACTTTTTTGGATCACGCAGTATTGATAGTAGGATATGGTTCAGAAAATGGAGTTGATTATTGGATTGTA
AAGAATTCATGGGGAAAATATTGGGGAATGAATGGCTACATACACATGATACGAAATAGTGGCAATTCCAAAGGA
ATATGCGGAATTAACACATTAGCTTCATATCCTGTCAAGACCAAGCCTAATCCTCCTATCCCTCCTCCCCCAGGT
CCTGTAAGATGTAATTTGTTTACTCACTGTTCAGAAGGAGAAACATGTTGCTGTGCCAAGAGCTTTCTTGGAATA
TGCTTTTCTTGGAAATGCTGTGGGTTAACTTCTGCCGTGTGCTGCAAAGACAAGCGCCACTGTTGCCCCCAAGAT
TATCCGATTTGTGATACAAGAAGAGGTCAATGTTTAAAGAGAACTGCAAATGGGACGACAACAATCACTTCTGAG
AACCAAGATTTTTCCCACAAGTCAAGAGGTTGGAAATCTCAGTGA

5.2 Legumain-like proteases sequences

Glyma17g14680.1

GTGAGTGACCGAGTGAGTTTGTTTTTCTCAGCTGATATATATGGCGCTTGATCGCTCCATTATAAGCAAAACGAC
GTGGTACAGCGTCGTATTATGGATGATGGTGGTGCTGGTGAGAGTGCACGGTGCAGCCGCGAGGCCGAACCGGAA
GGAGTGGGACTCAGTCATAAAGTTACCGACTGAACCGGTGGATGCTGACTCGGATGAAGTGGGAACACGATGGGC
GGTTCTCGTGGCTGGTTCAAACGG**CTACGGAAACTACAGGCAT**CAAGCAGATGTGTGCCATGCGTACCAGTTGCT
GATAAAAGGTGGACTAAAAGAAGAGAACATAGTGGTGTTTATGTACGATGACATAGCTACCAACGAGTTGAATCC
TAGACATGGAGTCATCATCAACCACCCTGAGGGAGAAGATCTGTATGCTGGTGTTCCTAAGGATTACACCGGTG**A**
TAATGTGACGACGGAGAACCTCTTTGCTGTTATTCTTGGAGACAAGAGTAAATTGAAGGGAGGAAGTGGCAAAGT
GATCAACAGCAAACCCGAGGACAGAATATTTATATACTACTCTGATCATGGAGGTCCTGGAATACTTGGGATGCC
AAACATGCCATACCTTTATGCCATGGATTTTATTGATGTCTTGAAGAAGAAACATGCATCTGGAAGTTACAAGGA
GATGGTTATATACGTGGAAGCTTGTGAAAGTGGGAGCGTGTTTGAGGGTATAATGCCTAAGGATCTGAATATTTA
TGTCACAACTGCATCAAATGCACAAGAGAATAGTTGGGGGACTTATTGTCCTGGAATGGATCCTTCTCCACCTCC
AGAGTACATCACTTGCCTAGGGGATTTGTACAGCGTTGCTTGGATGGAAGATAGTGAGGCTCACAATCTAAAAAG
GGAATCCGTGAAACAACAATACAAATCGGTAAAGCAACGGACTTCAAATTTCAACAACTATGCGATGGGTTCTCA
TGTGATGCAATATGGTGATACCAACATCACAGCTGAAAAGCTTTATTTATACCAAGGTTTTGATCCTGCCACTGT
GAACTTCCCTCCACAAAACGGCAGGCTAGAAACTAAAATGGAAGTTGTTAACCAAAGAGATGCAGAACTTTTCTT
ATTGTGGCAAATGTATCAGAGATCAAACCATCAGTCAGAAAATAAGACAGACATCCTCAAACAAATTGCGGAGAC
AGTGAAGCATAGGAAACACATAGATGGTAGCGTGGAATTGATTGGAGTTTTACTGTATGGACCAGGAAAAGGTTC
TTCTGTTCTACAATCCGTGAGGGCTCCTGGTTCGTCCCTTGTTGATGACTGGACATGCCTAAAATCAATGGTTCG
GGTGTTTGAAACTCACTGTGGGACACTGACTCAGTATGGCATGAAACACATGCGAGCATTCGCCAACATTTGCAA
CAGTGGCGTTTCTGAGGCCTCCATGGAAGAGGCTTGTTTGGCAGCCTGTGAAGGCTACAATGCTGGGCTATTCCA
TCCATCAAACAGAGGCTACAGTGCTTGATTTTGGGTTTTGTACACAAAAGCTTTAAAGCCCGGTTGATGATGTAA
TATTTCTCTATTGCATTCTGCCTACTGGTTTCTGCTGCTTGTGTCAAATTTTCTCTAAACTAGAGTAGCCCAATA
GCATACGTGTTATGTGCATTGGTCATGTATACAAGTGTAATACTAATACCTTCCTACATAATATAAGATTAGTTA
GTTTACTTGTC

Glyma05g04230

TGTTGCTGTCGAGCTGATATGGCGGTTGATCGCTCCCTTACGAGGTGCTGTAGCCTCGTACTGTGGTCGTGGATG
TTGCTGAGGATGATGATGGCGCAGGGTGCAGCCGCGAGGGCCAACCGGAAGGAGTGGGACTCGGTCATAAAGTTA
CCGGCTGAACCGGTCGATGCTGACTCGGATCATGAAGTGGGAACACGATGGGCGGTTCTTGTGGCTGGTTCAAAC
GGCTATGGAAACTACAGGCATCAAGCAGATGTGTGCCATGCGTACCAGTTGCTGATAAAAGGTGGGCTAAAAGAA
GAGAACATAGTGGTGTTTATGTACGATGACATAGCTACAGACGAGTTAAATCCCAGACCTGGAGTCATCATCAAC
CACCCTGAGGGACAAGATGTGTATGCTGGTGTTCCTAAGGATTACACCGGTGAGAATGTGACGGCCCAGAACCTC
TTTGCCGTTATTCTTGGAGACAAGAATAAAGTGAAGGGAGGAAGTGGCAAAGTGATCAATAGCAAACCTGAGGAC
AGAATATTTATATACTACTCTGATCATGGAGGTCCGGGAGTTCTTGGGATGCCAAACATGCCATACCTTTATGCT
ATGGACTTTATTGAAGTCTTGAAGAAGAAACATGCATCTGGAGGTTACAAGAAGATGGTCATATACGTGGAAGCT
TGTGAAAGTGGGAACCATGTTTTGAAGGGTATAATGCCTAAGGATCTGCAGATTTATGTCACAACTGCATCAAAT
GCACAAGAGAATAGTTGGGGAACTTATTGTCCTGGAATGGATCCTTCTCCACCTCCAGAGTACATCACTTGCCTA
GGGGATTTGTACAGTGTTGCTTGGATGGAAGATAGTGAGACTCATAATCTAAAAAGGGAGTCCGTGAAACAACAA
TACAAATCGGTAAAGCAACGGACTTCAAATTTCAACAACTATGCGATGGGTTCTCATGTGATGCAATACGGTGAC
ACAAACATCACAGCTGAAAAGCTTTATTTATACCAAGGTTTTGATCCTGCCGCTGTGAACTTCCCTCCACAGAAC
GGAAGGCTAGAAACTAAAATGGAAGTTGTTAACCAAAGAGATGCAGAACTTTTCTTCATGTGGCAAATGTATCAG
AGATCAAACCATCAGCCAGAAAAGAAGACAGACATCCTCAAACAGATAGCGGAGACAGTGAAGCATAGGAAACAC
ATAGATGGTAGCGTGGAATTGATTGGAGTTTTATTGTATGGACCAGGAAAAGGTTCTTCTGTTCTACAATCCATG
AGGGCTCCTGGTCTTGCCCTTGTTGATGACTGGACATGCCTAAAATCAATGGTTCGGGTGTTTGAGACTCACTGT
GGGACACTGACTCAGTATGGCATGAAACACATGCGAGCATTTGCCAACATTTGCAACAGCGGTGTTTCTGAGGCC
TCCATGGAAGAGGTTTGTGTGGCAGCTTGTGAAGGCTACGATTCTGGGCTATTACATCCATCAAACAAAGGCTAT
AGTGCTTGATTTTGGGTTTTGTACACAGCTTAAAAACCCGGTTGATGATGTAATACTTCTCTATTGCATTCTCCC
TACTGGTTTCTGCTGCATGTGTCAAATTTTCTCTAAACTAGAGTAGCCCAATAGCATACGTGTTATGAGCATTGG
TCATGTATATAAGTGTAATAGTAATATCTTTTACATATTATAAGATCAGTTAGTTTGGTTTACTA

Glyma14g10620

TCTCGAGTTTTTTTTTTTTTTTTTTTTTTTTTGAGTGTTGTAAAGGGTGTACCCCAAGCATAATTGAGTCCCAA
TATTTTACAAAAATAAATACAAGTTGTCATTAATATTAAACAATAATTTTGGCATCATAAATTCTTCCGGAATAA
TTATACAAAAGTGACCAGGGAAAACCCGATCTGCCAGTGGCCCACTGAATAAAAACACACATCATATTACTATGT
GTTGTCATACATGCATTGTAATGAGTGGATATTCCAAAGCAATTGGTCCCTTTATCCCTCCTTAATCCTTATGCC
GTAAATGGCACCATCCCTTGTAAATTGTAATTGCTTCTTCAGAAATAGAATTATACAAAGGTAGTTGCAATATCT
ATTTGCCAAACTATAAATGTAAAGACCCTTCTTCCCAATCCTAGTAGTGTCCCAATTGTATTTACAAAGGTGGGA
CAAGATATGTATATACAAGTACAACCATGTAACATAATGTAACAATAATCATACTTTGGTCTTTACTAGTGCACA
TTACANGGAGTTATTATGCACTGAAACCCCTATGCAGAGAACTCCAGGGATTGGAAGGAATACTGACACAAGCTG
GTGCTGAAGCCTCAGCCATTGATCATTCTTTATTCCTACATGGCAGATGTTGGCCAAGGACCT

5.3 Cystatin sequences

Glyma14g04250

ATAATCAATTATTTTATAATCTCGATGGCAATATTTATACTGAAATATATGGATGGCAAACATGCACAACCCCAC
ACCGTTGTAGAGTTAATAGGTACATTTATTCAGTTCATACATATTATTACTCTTGCGGCCAAGCTAGTCACTTCC
ATAACCTAGCTATGTGCTAGACCGTCACCGAAAGAGGATTAACAGCAACAGTGATGGGCTTGAATTCCTGAACCT
CCTTGATGTTCAACCATGGTTTCTCCAACACCTTAGTTTCATAAACCTTTTTCGTCACACCATCATTTGCCTCCA
AAGTGATATAGTACAAGGTGCCAGAAACCACTTGTTTCTTTGCACTAATCACCCTCACAAACTCCAGAACTGCAT
TCTCTTTTTTGTTGTGTTCATCAACAGCAAAGCGAGCGAGATTTTCGATATCGATGCTGTTCTGTGCTCCGGTGA
TGTCGGTAAAGCCACCCAGTGCTGCCATTTGCCTCGTGC

Glyma05g28250

AATTGTAAACATTTATTTCTCTCTTTCATATTTTTTATTAGAAGGATTAAAGCATACTAAACCAATTAAGCATGT
ATTATGAGTTAAAAAAGTAATTGAAACGGAACCGAAATTGATCGTATTTCACTGCGTGGAAGGAGCGAAATTGAA
AGCTGCTTGGAACGAAGCCACGGCTTCACCACCACAACGGATTCGAATATTCTGGAATCTCCACCGTCGCCACCC
TGCGTGGCCGATATCTTCATGTAGTACTTGATCCCAGACACCACTTGTTGTTGCGCCTCCACCACTTCCACGAAC
GTGACTTGCTCCTCCTCCTTCTGCGCCTGCCTCAGCATCCGGTTATGCTCCTCCACCGAGAACCGCCCTAGATCC
TGCACCTTCTTGTTCGCCTTCACGTCAGGGATCTTCGACCTTCCCCCGACCAATCCCCCGTACGACGCNGTACAC
GCGATGCACGCCNNGATCGTCAGGATCGCCAGTATCACCGCCGGTGACCTTATCAACG

Glyma15g36180

ATGAGAGCTCTTACTTCTTCTTCTTCTACTTTTATTCCTAAGAGATATTCTTTTTTTTTTTTTCTTTCTATTCTT
TTTGCTCTTAGATCTTCTTCTGGAGGATGTTCTGAATATCATCATCATCATGCTCCTATGGCTACTATTGGAGGA
CTTAGAGATTCTCAAGGATCTCAAAATTCTGTTCAAACTGAAGCTCTTGCTAGATTTGCTGTTGATGAACATAAT
AAGAAGCAAAATTCTCTTCTTGAATTTTCTAGAGTTGTTAGAACTCAAGAACAAGTTGTTGCTGGAACTCTTCAT
CATCTTACTCTTGAAGCTATTGAAGCTGGAGAAAAGAAGCTTTATGAAGCTAAGGTTTGGGTTAAGCCTTGGCTT
AATTTTAAGGAACTTCAAGAATTTAAGCCTGCTGGAGATGTTCCTTCTTTTACTTCTGCTGATCTTGGAGTTAAG
AAGGATGGACATCAACCTGGATGGCAATCTGTTCCTACTCATGATCCTCAAGTTCAAGATGCTGCTAATCATGCT
ATTAAGACTATTCAACAAAGATCTAATTCTCTTGTTCCTTATGAACTTCATGAAGTTGCTGATGCTAAGGCTGAA
GTTATTGATGATTTTGCTAAGTTTAATCTTCTTCTTAAGGTTAAGAGAGGACAAAAGGAAGAAAAGTTTAAGGTT
GAAGTTCATAAGAATAATCAAGGAGGATTTCATCTTAATCAAATGGAACAAGATCATTCT

Glyma13g27980

ATGAAATTGGAGTGCCTTGTTGTTTTGGCCTTTGTTCTCTTTGCTTCTGCAGTTGCTAGGAAGGAAATTGACCCC
AAAGGTTGGATCCCCATCGTTAAAATCAACGATCCTGCCCTAATAGAGATCGCTAAATTTTCTGTGACAGCATTC
AATGAGTGGGCTGGTTTGACTTTCAAGTTGCAGAAGCTGCTAAAAGCCGAAATGCAGGGCCAAGTTGTAAATGGG
ATCTACTTTCGACTCATCGTCGCCGCCGGCGATGATCAGTCCTCCTCCAACAATTACCAAGCCATTGTCTGGAAT
TTGCCGTGGACGGAAAATTACATGAACCTTATTTCCTTTATCCCTGTTCATGATGCACCTCCTAGTTACTATTAA

6. Literature cited

Alesandrini, F., Frendo, P., Puppo, A. and Hérouart, D. (2003). Isolation of a molecular marker of soybean nodule senescence. Plant Physiology and Biochemistry. 41: 727-732.

Alonso, J.M. and Granell, A. (1995). A putative vacuolar processing protease is regulated by ethylene and also during fruit ripening in Citrus fruit. Plant Physiology. 109: 541-547.

Andersson, B. and Aro, E.M. (1997). Proteolytic activities and proteases of plant chloroplasts. Physiologia Plantarum. 100: 780-793.

Asp, T., Bowra, S., Borg, S. and Holm, P.B. (2004). Molecular cloning, functional expression in *Escherichia coli* and enzymatic characterisation of a cysteine protease from white clover *Trifolium repens*. Biochimica et Biophysica Acta (BBA)-Proteins & Proteomics. 1699: 111-122.

Barrett, AJ. (1986). The classes of proteolytic enzymes. Edited by Dalling, M.J. Plant Proteolytic Enzymes: 1-16. CRC Press, Boca Raton.

Barrett, A.J., Kembhavi, A.A., Brown, M.A., Kirschke, H., Knight, C.G., Tamai, M. and Hanada, K. (1982). L-trans-epoxysuccinyl-leucylamido(4-guanidino) butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. Biochemical Journal. 201: 189–198.

Beers, E.P., Woffenden, B.J. and Zhao, C. (2000). Plant proteolytic enzymes: possible roles during programmed cell death. Plant Molecular Biology. 44: 399-415.

Benchabane, M., Schluter, U., Vorster, J., Goulet, M.C. and Michaud, D. (2010). Plant cystatins. Biochimie. 92: 1657-1666.

Bethlenfalvay, G.J. and Phillips, D.A. (1977). Ontogenetic interactions between photosynthesis and symbiotic nitrogen fixation in legumes. Plant Physiology. 60: 419-421.

Beyene, G., Foyer, C.H. and Kunert, K.J. (2006). Two new cysteine proteinases with specific expression patterns in mature and senescent tobacco (*Nicotiana tabacum* L.) leaves. Journal of Experimental Botany. 57: 1431-1443.

Bohnert, H.J. and Sheveleva, E. (1998). Plant stress adaptations—making metabolism move. Current Opinion in Plant Biology. 1: 267-274.

Boller, T. and Wiemken, A. (1986). Dynamics of vacuolar compartmentation. Annual Review of Plant Physiology. 37: 137-164.

Bottari, A., Capocchi, A., Galleschi, L., Jopova, A. and Saviozzi, F. (1996). Asparaginyl endopeptidase during maturation and germination of durum wheat. Physiologia Plantarum. 97: 475-480.

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

Bray, E.A. (1997). Plant responses to water deficit. Trends in Plant Science. 2: 48-54.

Cheong, J.J. and Choi, Y.D. (2003). Methyl jasmonate as a vital substance in plants. Trends in Genetics. 19: 409-413.

Chrispeels, M.J. and Raikhel, N.V. (1992). Short peptide domains target proteins to plant vacuoles. Cell. 68: 613-616.

Clement, M., Lambert, A., Herouart, D. and Boncompagni, E. (2008). Identification of new up-regulated genes under drought stress in soybean nodules. Gene 426: 15-22.

Colebatch, G., Desbrosses, G., Ott, T., Krusell, L., Montanari, O., Kloska, S., Kopka, J. and Udvardi, M.K. (2004). Global changes in transcription orchestrate metabolic differentiation during symbiotic nitrogen fixation in *Lotus japonicus*. The Plant Journal. 39: 487-512.

Coupe, S.A., Sinclair, B.K., Watson, L.M., Heyes, J.A. and Eason, J.R. (2003). Identification of dehydration-responsive cysteine proteases during post-harvest senescence of broccoli florets. Journal of Experimental Botany. 54: 1045.

Devienne-Barret, F., Justes, E., Machet, J. and Mary, B. (2000). Integrated control of nitrate uptake by crop growth rate and soil nitrate availability under field conditions. Annals of Botany. 86: 995-1005.

Diop, N.N., Kidric, M., Repellin, A., Gareil, M., d'Arcy-Lameta, A., Pham Thi, A.T. and Zuily-Fodil, Y. (2004). A multicystatin is induced by drought-stress in cowpea (*Vigna unguiculata* (L.) Walp.) leaves. Federation of European Biochemical Societies Letters. 577: 545-550.

Downing, W.L., Mauxion, F., Fauvarque, M.O., Reviron, M.P., Vienne, D., Vartanian, N. and Giraudat, J. (1992). A *Brassica napus* transcript encoding a protein related to the Künitz protease inhibitor family accumulates upon water stress in leaves, not in seeds. The Plant Journal. 2: 685-693.

Dutt, S., Pandey, D. and Kumar, A. (2011). Jasmonate signal induced expression of cystatin genes for providing resistance against Karnal bunt in wheat. Plant Signaling & Behavior. 6: 821-830.

Drake, R., John, I., Farrell, A., Cooper, W., Schuch, W. and Grierson, D. (1996). Isolation and analysis of cDNAs encoding tomato cysteine proteases expressed during leaf senescence. Plant Molecular Biology. 30: 755-767.

Espinosa-Victoria, D., Graham, P.H. and Vance, C.P. (2000). Host variation in traits associated with crown nodule senescence in soybean. Crop Science. 40: 103-109.

Estelle, M. (2001). Proteases and cellular regulation in plants. Current Opinion in Plant Biology. 4: 254-260.

Fedorova, E. and Brown, S. (2007). Cytochemistry of proteolytic activity and pH status of vacuoles in *Medicago truncatula* root nodules. Russian Journal of Plant Physiology. 54: 25-31.

Food and fertilization technology Centre (FFTC). (2012): http://www.agnet.org/library.php?func=view&id=20110803182212&type_id=1

Gogorcena, Y., Gordon, A.J., Escuredo, P.R., Minchin, F.R., Witty, J.F., Moran, J.F. and Becana, M. (1997). N_2 fixation, carbon metabolism, and oxidative damage in nodules of dark-stressed common bean plants. Plant Physiology. 113: 1193-1201.

Gonza´lez, EM., Aparicio-Tejo, PM., Gordon, AJ., Minchin, FR., Royuela, M. and Arrese-Igor, C. (1998). Water-deficit effects on carbon and nitrogen metabolism of pea nodules. Journal Experimental Botany. 49: 1705–1714.

Greenbaum, D., Medzihradszky, K.F., Burlingame, A. and Bogyo, M. (2000). Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. Chemistry & Biology. 7: 569-581.

Groten, K., Dutilleul, C., van Heerden, P.D.R., Vanacker, H., Bernard, S., Finkemeier, I., Dietz, K.J. and Foyer, C.H. (2006). Redox regulation of peroxiredoxin and proteinases by ascorbate and thiols during pea root nodule senescence. Federation of European Biochemical Societies Letters. 580: 1269-1276.

Grudkowska, M. and Zagdanska, B. (2004). Multifunctional role of plant cysteine proteinases. Acta Biochimica Polonica. 51: 609-624.

Gruis, D.F., Selinger, D.A., Curran, J.M. and Jung, R. (2002). Redundant proteolytic mechanisms process seed storage proteins in the absence of seed-type members of the vacuolar processing enzyme family of cysteine proteases. The Plant Cell. 14: 2863-2882.

Hara-Nishimura, I., Hatsugai, N., Nakaune, S., Kuroyanagi, M. and Nishimura, M. (2005). Vacuolar processing enzyme: an executor of plant cell death. Current Opinion in Plant Biology. 8: 404-408.

Hatsugai, N., Kuroyanagi, M., Yamada, K., Meshi, T., Tsuda, S., Kondo, M., Nishimura, M. and Hara-Nishimura, I. (2004). A plant vacuolar protease, VPE, mediates virus-induced hypersensitive cell death. Science Signalling. 305: 855-858.

Hong, J.K., Hwang, J.E., Lim, C.J., Yang, K., Jin, Z.L., Kim, C.Y., Koo, J.C., Chung, W.S., Lee, K.O. and Lee, S.Y. (2007). Over-expression of Chinese cabbage phytocystatin 1 retards seed germination in *Arabidopsis*. Plant Science. 172: 556-563.

Huffaker, R. (1990). Proteolytic activity during senescence of plants. New Phytologist. 116: 199-231.

Ishii, S.I. (1994). Legumain: Asparaginyl endopeptidase. Methods in Enzymology. 244: 604-615.

Johnson, R., Narvaez, J., An, G. and Ryan, C. (1989). Expression of proteinase inhibitors I and II in transgenic tobacco plants: effects on natural defence against *Manduca sexta* larvae. Proceedings of the National Academy of Sciences (USA). 86: 9871-9875.

Jones, J.T. and Mullet, J.E. (1995). A salt-and dehydration-inducible pea gene, Cyp15a, encodes a cell-wall protein with sequence similarity to cysteine proteases. Plant Molecular Biology. 28: 1055-1065.

Kameli, A. and Lösel, D.M. (1995). Contribution of carbohydrates and other solutes to osmotic adjustment in wheat leaves under water stress. Journal of Plant Physiology. 145: 363-366.

Kardailsky, I.V. and Brewin, N.J. (1996). Expression of cysteine protease genes in pea nodule development and senescence. Molecular Plant-Microbe Interactions. 9: 689-695.

Keyser, H.H. and Li, F. (1992). Potential for increasing biological nitrogen fixation in soybean. Plant and Soil. 141: 119-135.

Khanna-Chopra, R., Srivalli, B. and Ahlawat, Y. (1999). Drought induces many forms of cysteine proteases not observed during natural senescence. Biochemical and Biophysical Research Communications. 255: 324-327.

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

Lawlor, D. and Cornic, G. (2002). Photosynthetic carbon assimilation and associated metabolism in relation to water deficits in higher plants. Plant, Cell & Environment. 25: 275-294.

Lawn, R. and Brun, W.A. (1974). Symbiotic nitrogen fixation in soybeans. I. Effect of photosynthetic source-sink manipulations. Crop Science. 14: 11-16.

Lea, P. and Miflin, B. (1980). Transport and metabolism of asparagine and other nitrogen compounds within the plant. The Biochemistry of Plant. 5: 569-607.

Lee. H., Hur, C.G., On, C.J., Kim, B.H.,Park, S.Y. and An, C.S. (2004). Analysis of the root nodule-enhanced transcriptome in soybean. Molecules and Cells. 18: 53-62

Li, Y., Zhou, L., Li, Y., Chen, D., Tan, X., Lei, L. and Zhou, J. (2008). A nodule-specific plant cysteine proteinase, *As*NODF32, is involved in nodule senescence and nitrogen fixation activity of the green manure legume *Astragalus sinicus*. New Phytologist. 180: 185-192.

Lievens, S., Goormachtig, S. and Holsters, M. (2004). Nodule-enhanced protease inhibitor gene: emerging patterns of gene expression in nodule development on *Sesbania rostrata*. Journal of Experimental Botany. 55: 89-97.

Lohman, K.N., Gan, S., John, M.C. and Amasino, R.M. (1994). Molecular analysis of natural leaf senescence in *Arabidopsis thaliana*. Physiologia Plantarum. 92: 322-328.

Malik, N.S.A., Pfeiffer, N., Williams, D. and Wagner, F.W. (1981). Peptidohydrolases of soybean root nodules: Identification, separation, and partial characterization of enzymes from bacteroid-free extracts. Plant Physiology. 68: 386-392.

Manen, J.F., Simon, P., Van Slooten, J.C., Osteras, M., Frutiger, S. and Hughes, G.J. (1991). A nodulin specifically expressed in senescent nodules of winged bean is a protease inhibitor. The Plant Cell. 3: 259-270.

Martínez, D.E., Bartoli, C.G., Grbic, V. and Guiamet, J.J. (2007). Vacuolar cysteine proteases of wheat (*Triticum aestivum L.*) are common to leaf senescence induced by different factors. Journal of Experimental Botany. 58: 1099-1107.

Martínez, M., Cambra, I., González-Melendi, P., Santamaría, M.E. and Díaz, I. (2012). C1A cysteine-proteases and their inhibitors in plants. Physiologia Plantarum. 145: 85-94.

Martinez, M. and Diaz, I. (2008). The origin and evolution of plant cystatins and their target cysteine proteinases indicate a complex functional relationship. BMC Evolutionary Biology. 8, : 98-210.

Massonneau, A., Condamine, P., Wisniewski, J.P., Zivy, M. and Rogowsky, P.M. (2005). Maize cystatins respond to developmental cues, cold stress and drought. Biochimica et Biophysica Acta-Gene Structure and Expression. 1729: 186-199.

Matamoros, M.A., Baird, L.M., Escuredo, P.R., Dalton, D.A., Minchin, F.R., Iturbe-Ormaetxe, I., Rubio, M.C., Moran, J.F., Gordon, A.J. and Becana, M. (1999). Stress-induced legume root nodule senescence. Physiological, biochemical, and structural alterations. Plant Physiology. 121: 97-112.

Michaud, D. and Asselin, A. (2005). Application to plant proteins of gel eletrophoretic methods. Journal of Chromatography A. 698: 263-279.

Munne-Bosch, S. and Penuelas, J. (2003). Photo-and antioxidative protection, and a role for salicylic acid during drought and recovery in field-grown *Phillyrea angustifolia* plants. Planta. 217: 758-766.

Müntz, K. and Shutov, A.D. (2002). Legumains and their functions in plants. Trends in Plant Science. 7: 340-344.

Nagata, K., Kudo, N., Abe, K., Arai, S. and Tanokura, M. (2000). Three-dimensional solution structure of oryzacystatin-I, a cysteine proteinase inhibitor of the rice, *Oryza sativa* L. japonica. Biochemistry. 39: 14753-14760.

Naito, Y., Fujie, M., Usami, S., Murooka, Y. and Yamada, T. (2000). The involvement of a cysteine proteinase in the nodule development in Chinese milk vetch infected with *Mesorhizobium huakuii* subsp. rengei. Plant Physiology. 124: 1087-1096.

Noodén, L.D., Guiamét, J.J. and John, I. (1997). Senescence mechanisms. Physiologia Plantarum. 101: 746-753.

Pfeiffer, N.E., Torres, C.M. and Wagner, F.W. (1983). Proteolytic activity in soybean root nodules: activity in host cell cytosol and bacteroids throughout physiological development and senescence. Plant Physiology. 71: 797-802.

Pimratch, S., Jogloy, S., Vorasoot, N., Toomsan, B., Patanothai, A. and Holbrook, C. (2008). Relationship between biomass production and nitrogen fixation under drought-stress conditions in Peanut genotypes with different levels of drought resistance. Journal of Agronomy and Crop Science. 194: 15-25.

Pinheiro, C., Chaves, M.M. and Ricardo, C.P. (2001). Alterations in carbon and nitrogen metabolism induced by water deficit in the stems and leaves of *Lupinus albus* L. Journal of Experimental Botany. 52: 1063-1070.

Pladys, D. and Rigaud, J. (1982). Acid protease and leghemoglobin digestion in French-bean nodules. Zeischrift Pflanzenphysiologie 108: 163-169

Pladys, D. and Vance, C.P. (1993). Proteolysis during development and senescence of effective and plant gene-controlled ineffective alfalfa nodules. Plant Physiology. 103: 379-384.

Prins, A., van Heerden, P.D.R., Olmos, E., Kunert, K.J. and Foyer, C.H. (2008). Cysteine proteinases regulate chloroplast protein content and composition in tobacco leaves: a model

for dynamic interactions with ribulose-1, 5-bisphosphate carboxylase/oxygenase (Rubisco) vesicular bodies. Journal of Experimental Botany. 59: 1935-1950.

Puppo, A., Groten, K., Bastian, F., Carzaniga, R., Soussi, M., Lucas, M.M., De Felipe, M.R., Harrison, J., Vanacker, H. and Foyer, C.H. (2005). Legume nodule senescence: roles for redox and hormone signalling in the orchestration of the natural aging process. New Phytologist. 165: 683-701.

Richardson, M. (1977). The proteinase inhibitors of plants and micro-organisms. Phytochemistry. 16: 159-169.

Roberts, I.N., Caputo, C., Criado, M.V. and Funk, C. (2012). Senescence-associated proteases in plants. Physiologia Plantarum. 145: 130–139.

Rojo, E., Martin, R., Carter, C., Zouhar, J., Pan, S., Plotnikova, J., Jin, H., Paneque, M., Sanchez-Serrano, JJ. and Baker, B., (2004). VPE gamma exhibits a caspase-like activity that contributes to defence against pathogens. Current Biology. 14: 1897–1906.

Roponen, I. (2006). The effect of darkness on the leghemoglobin content and amino acid levels in the root nodules of pea plants. Physiologia Plantarum. 23: 452-460.

Rozman-Pungerčar, J., Kopitar-Jerala, N., Bogyo, M., Turk, D., Vasiljeva, O., Sttefe1 I., Vandenabeele I., Brömme, D., Puizdar, V., Fonović, M., Trstenjak-Prebanda M. and Dolenc, I. (2003). Inhibition of papain-like cysteine proteases and legumain by caspase-specific inhibitors: when reaction mechanism is more important than specificity. Cell Death Differentiation. 10: 881-888.

Sajid, M., McKerrow, J.H., Hansell, E., Mathieu, M.A., Lucas, K.D., Hsieh, I., Greenbaum, D., Bogyo, M., Salter, J.P. and Lim, K.C. (2003). Functional expression and characterization

of *Schistosoma mansoni* cathepsin B and its trans-activation by an endogenous asparaginyl endopeptidase. Molecular and Biochemical Parasitology. 131: 65-75.

Salas, C.E., Gomes, M.T.R., Hernandez, M. and Lopes, M.T.P. (2008). Plant cysteine proteinases: Evaluation of the pharmacological activity. Phytochemistry. 69: 2263-2269.

Salvesen, G. and Nagase, H. (1989). Inhibition of proteolytic enzymes. Proteolytic enzymes: A practical approach. 83-104.

Schaller, A. (2004). A cut above the rest: the regulatory function of plant proteases. Planta. 220: 183-197.

Scheible, W.R., Lauerer, M., Schulze, E.D., Caboche, M. and Stitt, M. (1997). Accumulation of nitrate in the shoot acts as a signal to regulate shoot-root allocation in tobacco. The Plant Journal. 11: 671-691.

Schultze, M. and Kondorosi, A. (1998). Regulation of symbiotic root nodule development. Annual Review of Genetics. 32: 33-57.

Serraj, R., Sinclair, T.R. and Purcell, L.C. (1999). Symbiotic N_2 fixation response to drought. Journal of Experimental Botany. 50: 143-155.

Sheokand, S., Dahiya, P., Vincent, J. and Brewin, N. (2005). Modified expression of cysteine protease affects seed germination, vegetative growth and nodule development in transgenic lines of *Medicago truncatula*. Plant Science. 169: 966-975.

Simova-Stoilova, L., Vaseva, I., Grigorova, B., Demirevska, K. and Feller, U. (2010). Proteolytic activity and cysteine protease expression in wheat leaves under severe soil drought and recovery. Plant Physiology and Biochemistry. 48: 200-206.

Smart, C.M., Hosken, S.E., Thomas, H., Greaves, J.A., Blair, B.G. and Schuch, W. (1995). The timing of maize leaf senescence and characterisation of senescence-related cDNAs. Physiologia Plantarum. 93: 673-682.

Solomon, M., Belenghi, B., Delledonne, M., Menachem, E. and Levine, A. (1999). The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. The Plant Cell. 11: 431-443.

Sprent, J.I. (1972). The effects of water stress on nitrogen-fixing root nodules. New Phytologist. 71: 603-611.

Sreedharan, S.K., Verma, C., Caves, L.S., Brocklehurst, S.M., Gharbia, S.E., Shah, H.N. and Brocklehurst, K. (1996). Demonstration that 1-trans-epoxysuccinyl-L-leucylamido-(4-guanidino) butane (E-64) is one of the most effective low Mr inhibitors of trypsin-catalysed hydrolysis. Characterization by kinetic analysis and by energy minimization and molecular dynamics simulation of the E-64-beta-trypsin complex. Biochemical Journal. 316: 777-786.

Stougaard, J. (2000). Regulators and regulation of legume root nodule development. Plant Physiology. 124: 531-540.

Sugawara, H., Shibuya, K., Yoshioka, T., Hashiba, T. and Satoh, S. (2002). Is a cysteine proteinase inhibitor involved in the regulation of petal wilting in senescing carnation (*Dianthus caryophyllus* L.) flowers? Journal of Experimental Botany. 53: 407-413.

Swaraj, K. and Bishnoi, N. (1996). Physiological and biochemical basis of nodule senescence in legumes: a review. Plant physiology and Biochemistry. 23: 105-116.

Swaraj, K., Dhandi, S. and Sheokand, S. (1995). Relationship between defence mechanism against activated oxygen species and nodule functioning with progress in plant and nodule development in *Cajanus cajan* L. Millsp. Plant Science. 112: 65-74.

Timmers, A.C.J., Soupène, E., Auriac, M.C., de Billy, F., Vasse, J., Boistard, P. and Truchet, G. (2000). Saprophytic intracellular *rhizobia* in alfalfa nodules. Molecular Plant-Microbe Interactions. 13: 1204-1213.

Turk, V. and Bode, W. (1991). The cystatins: protein inhibitors of cysteine proteinases. Federation of European Biochemical Societies Letters. 285: 213-219.

Turner, G.L. and Gibson, A.H. (1980). Measurement of nitrogen fixation by indirect means. In Methods for Evaluating Biological Nitrogen Fixation. Edited by F. J. Bergersen. Chichester: Wiley: 111 - 138.

Valueva, T. and Mosolov, V. (1999). Protein inhibitors of proteinases in seeds: Classification, distribution, structure, and properties. Russian Journal of Plant Physiology. 46: 307-321.

Van Der Hoorn, R.A.L., Leeuwenburgh, M.A., Bogyo, M., Joosten, M.H.A.J. and Peck, S.C. (2004). Activity profiling of papain-like cysteine proteases in plants. Plant Physiology. 135: 1170-1178.

Van de Velde, W., Guerra, J.C.P., De Keyser, A., De Rycke, R., Rombauts, S., Maunoury, N., Mergaert, P., Kondorosi, E., Holsters, M. And Goormachtig, S. (2006). Aging in legume symbiosis. A molecular view on nodule senescence in *Medicago truncatula*. Plant Physiology. 141: 711-720.

Vincent, J.L. and Brewin, N.J. (2000). Immunolocalization of a cysteine protease in vacuoles, vesicles, and symbiosomes of pea nodule cells. Plant Physiology. 123: 521-530.

Vincent, J., Newton, W. and Orme-Johnson, W. (1980). Factors controlling the legume-*Rhizobium* symbiosis. Nitrogen fixation. Volume 2: symbiotic associations and cyanobacteria: 103-129.

Vorster, B.J., Schlüter, U., du Plessis. M., van Wyk, S., Makgopa, M.E., Ncube, I., Quain, M.D., Kuner, K. and Foyer, C.H. (2013). The cysteine protease – cysteine protease inhibitor system in soybean nodule development. Agronomy –Accepted.

Walsh, T.A. and Strickland, J.A. (1993). Proteolysis of the 85-kilodalton crystalline cysteine proteinase inhibitor from potato releases functional cystatin domains. Plant Physiology. 103: 1227-1234.

Wiederanders, B. (2003). Structure-function relationships in class CA1 cysteine peptidase propeptides. Acta Biochimica Polonica. 50: 691-713.

Wolters, B.C. (1969). Kinetics of inhibition of papain by TLCK and TPCK in the presence of BAEE as substrate. Federation of European Biochemical Societies Letters. 3: 143-145

Zahran, H.H. (1999). *Rhizobium*-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. Microbiology and Molecular Biology Reviews. 63: 968-989.