Your thoughts become your destiny

Promoter: Prof. Dr. Els. J. M. Van Damme

      Laboratory of Biochemistry and Glycobiology

      Department of Molecular Biotechnology

      Ghent University

Dean: Prof. Dr. ir. Guido Van Huylenbroeck

Rector: Prof. Dr. Anne De Paepe

Bassam Al Atalah

Use of carbohydrate binding proteins to increase stress tolerance of crops

Thesis submitted in fulfillment of the requirements for the degree of Doctor (PhD) in Applied Biological Sciences

Dutch translation

Toepassing van koolhydraatbindende   eiwitten om de stresstolerantie van planten te verhogen

Cover illustration

The left image: rice plant. The upper right image: confocal image collected from living, transiently transformed tobacco BY-2 cells expressing EGFP-Orysata construct. Cell compartments: n, nucleolus; N, nucleus; m, cell membrane; c, cytoplasm; v, vacuole. The lower right image: ribbon diagrams for the EUL domain of OrysaEULS2.

The promoter   **The Author**

Prof. Dr. Els. J. M. Van Damme   Bassam Al Atalah

Members of the examination committee

Prof. Dr. Els. J. M. Van Damme (promoter)

Department of Molecular Biotechnology, Faculty of Bioscience Engineering, Ghent University, Belgium

Prof. Dr. Monica Höfte

Department of Crop Protection, Faculty of Bioscience Engineering, Ghent University, Belgium

Dr. Tina Kyndt

Department of Molecular Biotechnology, Faculty of Bioscience Engineering, Ghent University, Belgium

Prof. Dr. Guy Smagghe

Department of Crop Protection, Faculty of Bioscience Engineering, Ghent University, Belgium

Prof. Dr. Wim Van den Ende

Lab of Molecular Plant Biology, Institute of Botany and Microbiology, KU Leuven, Belgium

Prof. Dr. Kris Verheyen (chairman)

Department of Forest and Water Management, Ghent University, Belgium

Table of Contents

List of abbreviations

AA	amino acid
ABA	abscisic acid
AOX1	alcohol oxidase 1
APA	*Allium porrum* agglutinin
ASA	*Alium sativum* agglutinin
ASAII	*Allium sativum* bulb agglutinin II
ASAL	*Allium sativum* leaf agglutinin
Calsepa	*Calystegia sepium* rhizomes
Con A	Concanavalin A
EEA	*Euonymus europaeus* agglutinin
EGFP	enhanced green fluorescent protein
EUL	*Euonymus* lectin
Fuc	fucose
FW	fresh weight
Gal	galactose
GlcNAc	N-acetylglucosamine
GNA	*Galanthus nivalis* agglutinin
GUS	β-glucuronidase
HCA	Hydrophobic cluster analysis
HHA	*Hippeastrum* hybrid agglutinin
HIV	human immunodeficiency virus
JRL	jacalin-related lectin
LC_{50}	50% lethal concentration
Man	mannose

MS	Murashige and Skoog medium
MW	molecular weight
Orysata	*Oryza sativa* agglutinin (inducible)
OSA	*Oryza sativa* agglutinin (vacuolar)
OSR40	*Oryza sativa* repeats 40 kDa
PHA	*Phaseolus vulgaris* agglutinin
PVDF	polyvinylidene fluoride
RFU	relative fluorescence units
ROS	reactive oxygen species
RSV	respiratory syncytial virus
SE	standard error
TaHfr-1	*Tritium aestivum* Hessian fly responsive 1
TAP	tandem affinity purification
TBS	Tris-buffered saline

Scope

Agricultural plants face a variety of abiotic stresses such as salinity, drought, low and high temperatures as well as biotic stresses caused by insects, bacteria, fungi, viruses and herbivores. Due to the lack of mobility, plants cannot avoid these unfavorable conditions and therefore have developed complex mechanisms to cope with them. One of these mechanisms is the synthesis of protein containing carbohydrate binding domains. During the last decade, a new group of inducible lectins was discovered. Hitherto, little information is known about the possible role of these inducible lectins but the hypothesis has been put forward suggesting a physiological role for the inducible lectins when plants are confronted with stress. Therefore, the general goal of the work described in this thesis is to study the possibility of using some of these inducible lectins for increasing the tolerance of crops against stresses.

Lectins belonging to two families of inducible proteins were selected for this PhD research. The first lectin is called Orysata and was discovered in rice seedlings treated with NaCl and therefore is known as a stress related protein. Orysata belongs to the family of jacalin-related lectins. In addition we focused on OrysaEULS2 and OrysaEULD1A belonging to the family of EUL-related lectins. OrysaEULS2 represents the S-type containing one EUL domain whereas OrysaEULD1A represents the D type with two tandemly arrayed EUL domains.

The first objective of this work was to characterize the proteins under study; Orysata, OrysaEULS2 and OrysaEULD1A. In **chapter 2**, the focus was on the cloning, the expression analysis and the purification for the recombinant proteins using the heterologous system *Pichia pastoris* which allowed us to determine the molecular structure and carbohydrate-binding specificity of the lectins under study.

The second objective for this PhD research was to analyze the expression of the stress-inducible lectins in the plant. **Chapter 3A** describes the promoter activity for OrysaEULS2 and OrysaEULD1A after different abiotic stress treatments using qualitative GUS staining assays in plants during different stages of development. In addition, **chapter 3B** focuses on investigation of the expression level for Orysata as well as the whole set of putative EUL lectins in rice subjected to some important abiotic and biotic stresses.

The third objective of this research was to evaluate the performance of the transgenic lines over-expressing Orysata, OrysaEULS2 and OrysaEULD1A under different abiotic and biotic stresses. To perform this analysis, Orysata, OrysaEULS2 and OrysaEULD1A were expressed ectopically in *Arabidopsis* plants behind the 35S promoter. In **Chapter 4A** the performance of transgenic *Arabidopsis* overexpressing each of Orysata, OrysaEULS2 and OrysaEULD1A was analyzed after abiotic and biotic stresses. In addition, the Orysata sequence was introduced into tobacco. **Chapter 4B** describes the effect of Orysata against biting-chewing and piercing-sucking insects when ectopically expressed in tobacco plants. This effect was investigated by short term experiments on detached tobacco leaves over-expressing Orysata as well as bioassays with an artificial diet containing different concentrations of recombinant Orysata.

Chapter 1

General introduction

1.1 Justifications of this PhD research

1.1.1 Why rice?

In recent years, rice has become a model system to investigate the impact of abiotic and biotic stresses on crops mainly because of the importance of rice for food security, the availability of diverse genetic resources for rice, the existence of the full genome sequence, and the presence of advanced research tools (Paterson et al., 2005; IRGSP 2005; Xu et al., 2006). Rice (*Oryza sativa* L.) is one of the most important food crops in the world supplying 20% of daily calories to more than 3.5 billion people worldwide (Khush 2013; World Rice Statistics, http://www.irri.org). Because rice belongs to the monocot family Gramineae, any progress made in rice can be applied to other members of this family especially maize and wheat which are also very important for food security.

1.1.2 Description of the problem

Abiotic and biotic stresses are the major factors affecting plant growth and crop productivity. Owing to their sessile nature, plants have evolved specific adaptive mechanisms as a response to environmental stresses. Salinity and drought are the most important factors among abiotic stresses. Salinity is one of the most serious factors limiting the productivity of agricultural crops (Munns and Tester 2008). On a world scale, more than 45 million hectares of irrigated land are damaged as a result of high salinity and 1.5 million hectares are taken out of production yearly (Munns and Tester 2008). It is estimated that 6% of the world's total land and 20% of the world's irrigated areas are affected by salinity. Drought is a major environmental stress factor that affects the growth and development of plants. Drought can be chronic in regions with low water availability or random and unpredictable due to changes in the weather conditions. The harmful effect of drought is expected to increase because of the climate change and the limited water sources.

In addition, biotic stresses provoked by fungi, bacteria, viruses, insects,… are also decreasing the yield of crops. For instance, 10 to 15% reduction of rice yield was estimated to be due to diseases caused by bacterial blight (*Xanthomonas oryzae*) and fungal blast (*Magnaporthe oryzae*) (Dai et al., 2007). Based on these facts, the identification of stress tolerance genes

and understanding their functions have become the most urgent tasks in order to enable the search for new methods to increase the yield of crops under unfavorable conditions.

The world population is increasing at an alarming rate and it is expected to reach nine billion by 2050, but unfortunately the food production is limited. It has been proposed that global food production must increase by 70% by 2050 to meet the demand caused by this growing population (Godfray et al., 2010). Global rice demand is estimated to rise from 676 million tons in 2010 to 763 million tons in 2020 and to 852 million tons in 2035. This is an overall increase of 176 million tons in the next 25 years. To reach this high demand, the yield potential of rice has to increase from 10 to 12.3 tons per hectare (Khush 2013).

Conventional breeding approaches have been used to exploit the natural genetic variation for improving rice varieties. However, because of the complexity of stress tolerance traits, conventional approaches are less effective in connecting tolerance traits to the determinant genes that play key roles in the stress response. Therefore, biotechnological approaches have to be used with the aim of increasing the crop tolerance against stresses to avoid the disaster of hunger.

1.2 Plant lectins

1.2.1 Historical overview

Lectins are a class of proteins with a very long scientific history. The lectin story started in 1888 when Stillmark, in his PhD thesis, linked the toxicity of castor bean (*Ricinus communis* L.) to a proteinaceous haemagglutinating factor which was called 'ricin' (Stillmark 1888). The term hemagglutinin was first introduced in 1898 as a common name for all plant proteins that cause clumping of cells (Elfstrand 1898). In 1907, Landsteiner and Raubitschek reported for the first time the presence of nontoxic lectins in the legumes *Phaseolus vulgaris* (bean), *Pisum sativum* (pea), *Lens culinaris* (lentil), and *Vicia sativa* (vetch). Afterwards, certain hemagglutinins showed specificity towards erythrocytes of a particular human blood group within the ABO system (Renkonen 1948; Boyd and Reguera 1949). The discovery of blood group specificity was the direct motive to the introduction of the novel term 'lectin' (from the Latin verb 'legere', which means 'to select') (Boyd and Shapleigh 1954).

The first definition of lectins was based primarily on their sugar specificity and their ability to inhibit the agglutination of cells. According to this concept, lectins were defined as "carbohydrate-binding proteins (or glycoproteins) of non-immune origin that agglutinate cells and/or precipitate glycoconjugates" (Goldstein et al., 1980). Based on this definition, only multivalent carbohydrate-binding proteins can be classified as lectins. In an attempt to broaden the definition, Kocourek and Horejsi (1983) modified the definition as "Lectins are proteins of non-immunoglobulin nature capable of specific recognition and reversible binding to carbohydrate moieties of complex carbohydrates without altering the covalent structure of any of the recognized glycosyl ligands". Afterwards, lectins were redefined as "carbohydrate-binding proteins other than antibodies or enzymes" (Barondes 1988). However, this definition should not exclude all enzymes since some lectins do have an enzymatic part (like type 2 ribosome-inactivating proteins and class I chitinases). At present, the definition which is scientifically accepted describes lectins as "all proteins possessing at least one non-catalytic domain, which binds reversibly to a specific mono- or oligosaccharide" (Peumans and Van Damme 1995). Accordingly, agglutination is no longer considered a criterion to classify a protein as a lectin.

1.2.2 Classification

All known plant lectins can be classified into 12 families based on their overall structure and specificity (Van Damme et al., 2008; 2010; Lannoo and Van Damme 2010). The structure of the carbohydrate domain as well as the carbohydrate specificity for each plant lectin family are summarized in Table 1.1.

Table 1.1 Lectin families based on structural homology of the carbohydrate binding domain

Lectin family	Structure	Specificity	Example	Localization
Agaricus bisporus agglutinin	Homodimer/ β-sandwich	T-antigen	MarpoABA	Nucleocytoplasm
Amaranthins	Homodimer/ β-trefoil	T-antigen and GalNAc	Amaranthin	Nucleocytoplasm
Class V chitinase homologues with lectin activity	Homodimer/ TIM-barrel	High-Man N-glycans	RobpsCRA	Vacuole
Cyanovirins	Homodimer/ Triple-stranded β-sheet and a β-hairpin	High-Man N-glycans	Cyanovirin-N	Vacuole
Lectins with EUL domain(s)	Homodimer/ Structure unknown	Blood group B oligosaccharides and high-Man N-glycans	EEA	Nucleocytoplasm
GNA-related lectins	Different oligomerisation states β-barrel	Man, N-glycans	GNA	Vacuole
			GNA$_{maize}$	Nucleocytoplasm
Jacalin-related lectins	Different oligomerisation states β-prism	Gal/T-antigen specific sub-group	Jacalin	Vacuole
		Man-specific subgroup	Orysata	Nucleocytoplasm
Lectins with hevein domains	Different oligomerisation states Hevein domain	(GlcNAc$_n$)	Hevein	Vacuole
Legume lectins	Different oligomerisation states	Man/Glc, Gal/GalNAc, (GlcNAc$_n$), Fuc, Siaα2,3Gal/GalNAc, complex glycans	Con A	Vacuole

	β-sandwich			
Proteins with LysM domains	Different oligomerisation states

β-α-α-β- structure | $(GlcNAc_n)$ | CEBiP | Vacuole |
| Nictaba-like lectins | Homodimer

Structure unknown | $(GlcNAc_n)$ and N-glycans | Nictaba | Nucleocytoplasm |
| Ricin-B lectins | Different oligomerisation states

β-trefoil | Gal/GalNAC, Siaα2,6Gal/GalNAc | Ricin | Vacuole

Nucleocytoplasm |

1.2.3 Function of plant lectins

1.2.3.1 Vacuolar lectins

To classify a protein as a vacuolar plant lectin, it has to meet several criteria (Van Damme et al., 2004b). First, the vacuolar lectins are constitutively expressed at high concentration (0.1-10% of the total protein) in the seeds and the vegetative tissues suggesting a possible role as storage proteins. Second, the expression level of such lectins is not affected by environmental stresses but it is considered as a part of the developing program of the plant cell. Third, the vacuolar lectins are synthesized on ribosomes attached to the endoplasmic reticulum and afterwards they follow the secretory pathway through the Golgi apparatus. Eventually, they are located in the vacuoles or are secreted into the extracellular spaces. Fourth, these abundant lectins show affinity towards 'foreign' glycans (i.e. not occurring in plants) pointing at their expected role in the defense system. It has been demonstrated that vacuolar lectins such as Concanavalin A (Con A), *Galanthus nivalis* agglutinin and wheat germ agglutinin showed negative effects on the development and survival of phytophagous insects belonging to Lepidoptera, Coleoptera, Diptera and Hemiptera (for details see Michiels et al., 2010; Vandenborre et al., 2011b).

It can be concluded that the vacuolar lectins play a storage role under normal growth conditions but can be recruited for a defense role in case the plant is attacked by predators.

1.2.3.2 Nucleocytoplasmic lectins

Over the last years, evidence has accumulated that plants synthesize well-defined carbohydrate-binding proteins in response to several abiotic and biotic stresses such as drought, high salt, wounding, treatment with plant hormones, or pathogen attack. The discovery of some inducible lectins in rice (Zhang et al., 2000) and tobacco (Chen et al., 2002) led to the development of a novel concept for lectin function. Since these inducible lectins are synthesized only after exposure to stress conditions, the hypothesis was put forward that lectin-mediated protein–carbohydrate interactions in the cytoplasm and the nucleus might play an important role in the stress physiology of the plant cell (Van Damme et al., 2004b; Van Damme et al., 2004a). Taking into consideration that any physiological role of plant lectins most likely will rely on their carbohydrate-binding activity, the discovery of the

novel stress related lectins provides a strong evidence for the importance of protein-carbohydrate interactions in plants. Localization studies demonstrated that these so called "inducible" lectins are located in the cytoplasm and/or the nucleus of the plant cell and therefore they are also called nucleocytoplasmic lectins (Lanno and Van Damme 2010).

The localization of the inducible lectins inside the plant cell raises the question about their putative receptors. Carbohydrates attached to proteins can be classified into two main categories: N-glycans, linked to asparagine residues; and O-glycans, linked to the serine, threonine and hydroxyproline residues in the protein chain. With the exception of nucleocytoplasmic O-GlcNAcylation, protein N- and O-glycosylation typically occur in the secretory pathway. Despite the huge progress in glycoconjugate research over the last decades, the nucleocytoplasmic side of glycobiology is still poorly understood. The presence of most of the inducible lectins in the cytoplasm and the nucleus suggests that the binding partners for these proteins might be glycosylated proteins, glycolipids, free sugars or polysaccharides and free glycans.

Higher plants have N-linked glycans ranging from high-mannose to complex N-glycans, with one or two terminal Lewis A antennae (Lerouge et al., 1998; Kimura et al., 2005) or contain other sugars like fucose and xylose (Gomord et al., 2005). The high-mannose-type N-glycans have the same structure in plant and mammalian glycoproteins. In contrast, complex-type N-glycans are structurally different in plants and mammals. For instance, in plants, the proximal N-acetylglucosamine of the core is substituted by an α1,3 fucose rather than an α1,6 fucose in mammals, and the β-mannose of the core is substituted by a bisecting β1,2 xylose in plants, rather than a β1,4 N-acetylglucosamine in mammals. In addition, β1,3 galactose and fucose α1,4-linked to the terminal N-acetylglucosamine of plant N-glycans form Lewis A oligosaccharide structures instead of β1,4 galactose combined with sialic acids in mammals. It is worth mentioning here that Blood group B, Lewis X, Lewis Y and lactosamine structures are well-studied in higher animals (Stanley et el., 2009), bacteria and viruses (Monzavi-Karbassi et al., 2004) but they have never been reported in plants. Free N-glycans in the cytosol presumably originate from the de novo synthesis or N-glycoconjugate degradation (Priem et al., 1993; Nakamura et al., 2008). In 2010, Maeda et al. (2010) analyzed the free N-glycans in the intracellular and the extracellular spaces in a rice cell suspension culture. The intracellular fraction contained mainly high-mannose type N-glycans

with one GlcNAc residue and but also truncated complex type N-glycans whereas complex free N-glycans containing the Lewis A epitope and high-mannose type free N-glycans were characterized from the culture medium. In addition, metabolic glycans, such as cytosolic heteroglycans resulting from the degradation of leaf starch (Fettke et al., 2009; 2011) or other plant polysaccharides could also be a putative target for the lectins in the cytoplasm. Based on these observations, it seems reasonable to expect that the biological activity of the nucleocytoplasmic proteins relies on their binding to cytoplasmic/nuclear receptors. In addition, we cannot exclude the possibility that as a response to external stresses some of the inducible proteins are targeted from the cytosol into the apoplast to interact with glycolipids in the plasma membrane and other sugar containing structures in the cell wall. Specificity studies revealed that Orysata, classified as the first member of the nucleocytoplasmic lectins (see section 1.2.5.4), showed strong affinity towards high-mannose N-glycans (Zhang et al., 2000). Similarly, Nictaba (a nucleocytoplasmic lectin from tobacco) reacts well with GlcNAc oligomers but exhibits a higher affinity for high-mannose N-glycans (Lannoo et al., 2006a). Based on this observation, the cytosolic N-glycosylated glycoproteins might be the glycan-receptors for these lectins. However, most O- and N-glycosylated proteins known thus far follow the secretory pathway for localization in vacuoles and extracellular spaces, but the occurrence of O- and N-glycosylated glycoproteins in the cytosol and in the nucleus of plant cells has been reported (Funakoshi and Suzuki 2009; Schouppe and Van Damme 2011). Furthermore, we cannot exclude the possiblility that the inducible lectins can bind free sugars, polysaccharides, free N-glucans and glycolipids and we can also not exclude that these lectins can act through protein–protein interactions with non-glycosylated proteins.

In conclusion, the nucleocytoplasmic lectins can be characterized by different features. First, they lack the signal peptide sequence and thus are synthesized on free ribosomes in the cytoplasm and finally localize to the nucleus and/or the cytoplasm of the plant cell. Second, the inducible lectins are not constitutively expressed but are synthesized in response to different stimuli. However, even after induction the expression level of these lectins remains low. Third, since they are nucleocytoplasmic proteins, they can interact with plant glycans inside these cell compartments.

It should be mentioned, however, that some plant species accumulate high concentration of proteins that resemble the stress-related lectins. Apparently some of the nucleocytoplasmic lectins have evolved through evolution and plants have used the corresponding genes as templates for the development of lectins with a storage and/or defence related role (Van Damme et al., 2004b). Especially in the family of jacalin-related lectins several examples of highly expressed lectins that locate to the cytoplasm have been reported, such as the banana (*Musa acuminata*) fruit lectin (Koshte et al., 1990; Peumans et al., 2000a), the *Helianthus tuberosus* tuber agglutinin (Van Damme et al., 1999) and the *Calystegia sepium* rhizome agglutinin (Peumans et al., 1997).

1.2.4 Nucleocytoplasmic lectin families

Until now, six families of nucleocytoplasmic lectins have been described (Table 1.1) (Lannoo and Van Damme 2010). These families are: *Agaricus bisporus* agglutinin, amaranthins, lectins with (a) EUL domain(s), GNA-related lectins, jacalin-related lectins and Nictaba-like lectins. In this PhD research, we are focusing on the jacalin-related lectins as well as lectins with (a) EUL domain(s). Therefore, the following sections give an overview of the most important characteristics of these two lectin families.

1.2.4.1 Jacalin-Related Lectins (JRLs)

Classification, structure and localization of JRLs

The family of jacalin-related lectins (JRLs) got its name from jacalin, the trivial name of the galactose-binding lectin isolated from jackfruit (*Artocarpus integrifolia*) seeds (Sastry et al., 1986). According to their specificity, localization and molecular structure, JRLs are subdivided into mannose-specific and galactose-specific lectins (Peumans et al., 2000b). The size and conformation of the carbohydrate-binding site is more extended for the galactose JRLs due to the cleavage of the lectin protomer into two polypeptide chains [~20 amino acid (AA) β and ~130 AA α chains] by the posttranslational excision of a tetrapeptide linker (Bourne et al., 2002; Houlès-Astoul et al., 2002). In contrast, mannose JRLs are composed of either un-cleaved protomers of approximately 150 AA or protomers comprising two to seven tandemly arrayed jacalin domains (Van Damme et al., 2010). Whereas the galactose JRLs are

localized in the vacuoles, the mannose JRLs are localized in the cytoplasm and/or the nucleus of the plant cell (Peumans et al., 2000b).

Inducible JRLs

The first inducible lectin was described in rice seedlings after salt treatment (Claes et al., 1990). This protein, called Orysata, (see **section1.2.5.4**) was classified as a mannose-specific jacalin-related lectin (Zhang et al., 2000; Hirano et al., 2000). Jasmonate-inducible orthologs of Orysata have been identified in *Helianthus tuberosus* (Heltuba) (Nakagawa et al., 2000), *Brassica napus* (Geshi and Brandt 1998) and *Ipomoea batatas* (Ipomoelin) (Imanishi et al., 1997).

Within the Gramineae family, several inducible proteins with a chimeric form containing a jacalin domain fused to an unrelated domain have been identified. One of these proteins is specifically expressed during vernalization in wheat (called VER2) (Yong et al., 2003). Using sequence information of VER2, a rice cDNA encoding mannose-binding jacalin-related lectin (designated OsJAC1) with a disease-response domain in the N-terminal sequence was characterized (Jiang et al., 2006). The expression of VER2 protein in rice (OsJAC1) and wheat was shown to be jasmonate inducible. In addition, TaJA1, *Tritium aestivum* jasmonate 1, from wheat (Wang and Ma 2005) and HvJRP1, *Hordeum vulgare* jacalin-related protein 1, from barley are considered as homologs to the jasmonate inducible proteins from barley collectively called JRP-32, jacalin-related proteins with a molecular mass of 32 kDa, (Ma 2013). TaHfr-1, *Tritium aestivum* Hessian fly responsive 1, is a gene that is induced in wheat upon infestation by larval feeding of the Hessian fly (Williams et al., 2002). The TaHfr-1 protein exhibits mannose-specific activity (Subramanyam et al., 2008). Another structurally similar protein containing a jacalin domain occurs in maize, where it is known as a β-glucosidase-aggregating factor (Molina et al., 2004; Kittur et al., 2007). The wide distribution of JRLs suggests an important role for lectins from this family. It has been reported that overexpression of OsJAC1 in rice suppresses coleoptile and stem elongation, indicating that this lectin is important for rice growth and development (Jiang et al., 2007).

1.2.4.2 Lectins with EUL domain(s)

EEA is a structural unit for a new family

Already in the 1970s it was known that the *Euonymus europaeus* (spindle tree) contains an agglutinin in the fleshy arils surrounding the seeds, referred to as *Euonymus europaeus* agglutinin (EEA) (Petryniak et al., 1977). In 2008, EEA was investigated for the first time at the molecular level, resulting in the cloning of the lectin cDNA sequence (Fouquaert et al., 2008). The EEA cDNA encodes a 17 KDa protein with no signal peptide suggesting that EEA is synthesized on free ribosomes. Furthermore, localization analysis revealed that this lectin is localized in the cytoplasm and nucleus of the plant cell (Van Hove et al., 2011). Sequence comparison revealed that EEA shares high sequence identity (46%) and similarity (62%) with some rice proteins induced by ABA and salt treatments, referred to as OSR40 proteins (*Oryza sativa* repeats 40 kDa) (Moons et al., 1997a). These rice proteins are annotated in the database as "Ricin-B related lectin domain containing proteins" based on the presence in their sequence of two QXW repeats, which are considered typical motifs of the ricin-B domain. But according to the BLASTp search there is no significant sequence similarity with proteins comprising a ricin-B domain and thus it is questionable whether they can be considered as ricin-B family members. Therefore EEA and the OSR40 proteins are now classified in a new family of so-called proteins with *Euonymus* lectin (EUL) domain(s). Since then the EUL domain is considered as the prototype for this new lectin family (Fouquaert et al., 2009a; Fouquaert and Van Damme 2012).

Occurrence and classifications of EULs

Screening on genome and transctiptome levels revealed that proteins with an EUL domain are ubiquitous within the Embryophyta, but are not present in other eukaryotes or in prokaryotes (Fouquaert et al., 2009a). EUL proteins are widely distributed from monocots (rice, maize), to dicots (*Arabidopsis*, tomato) and the lower plants (*Physcomitrella*, *Selaginella*).

Based on the overall domain architecture of all EUL sequences known to date a classification system for this lectin family has been proposed (Fouquaert et al., 2009a). According to this system, the EUL family can be subdivided into two big classes; the members of the S type

comprise a single EUL domain whereas members of the D type containing two tandemly arrayed EUL domains (Fig. 1.1). Based on the occurrence of N- and C- terminal sequences, the S type can be divided into seven subtypes. Type S0 represents all proteins that consist exclusively of EUL domains. Depending on the length of the N-terminal domain three different types of EUL proteins can be classified, being type S1 (short N sequence < 50 AA), type S2 (medium long 50–100 AA) and type S3 (long N-sequence > 100 AA). Similarly based on the length of the C-terminal sequence two additional types are added; type S4 represents proteins consisting of a medium long (50-100 AA) unrelated N-terminal sequence, an EUL domain, and a short (<50) C-terminal extension whereas type S5 represents proteins consisting of a short (<50) unrelated N-terminal sequence, an EUL domain, and a short (<50) or medium long (50-100 AA) C-terminal extension. In addition, type Sv (v stands for vacuole) consists of an EUL domain preceded by a short unrelated N-terminal sequence containing a putative signal peptide.

Similar to the S type, the D type can be subdivided into five different subtypes based on the length of the N-terminal domain preceding the EUL domains and the linker sequence between the two EUL sequences. Type D0 groups proteins consisting of two in tandem arrayed EUL domains separated by a short (<40 AA residues) linker and without N-terminal extension. Type D1 represents all proteins consisting of two in tandem arrayed EUL domains separated by a short (<40 AA residues) linker and preceded by a short (15-35 AA residues) N-terminal extension. Type D2 resembles type D1 but the linker in between the EUL domain is longer (>40 AA residues). Type D3 resembles type D1 but the N-terminal sequence is longer (> 50 AA residues). Finally, type D4 comprises proteins consisting of two in tandem arrayed EUL-related domains (with low sequence similarity to EEA) separated by a short (<40 AA residues) linker and without N-terminal extension.
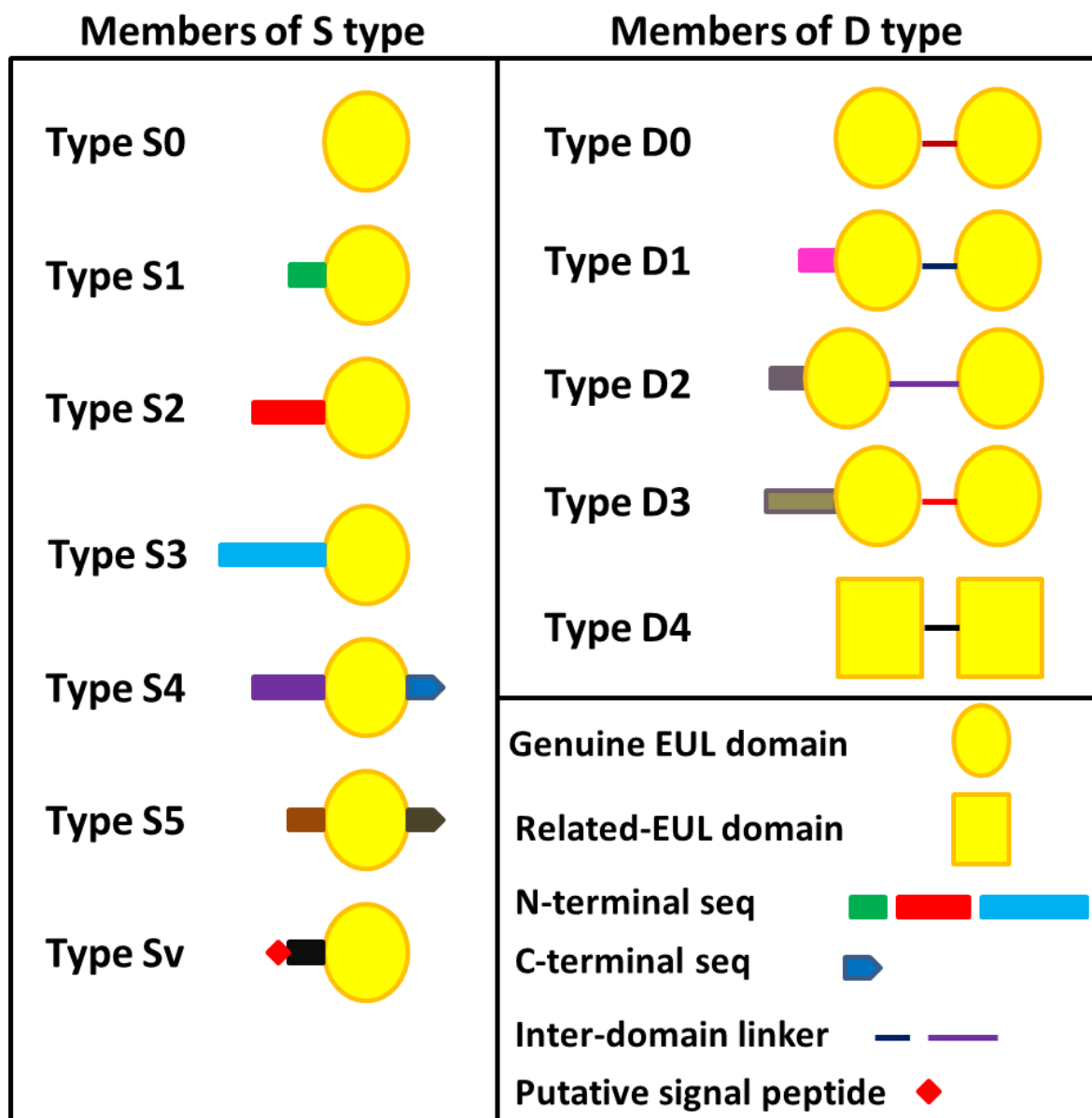
Fig. 1.1 Schematic representation of the 12 types of *Euonymus*-related lectins (EUL). Variability in the N- and C-terminal sequences, inter-domain linkers is indicated by different lengths and colors. Seq: sequence. Adapted from (Fouquaert et al., 2009a).

Inducible EULs

As mentioned above rice EUL proteins are induced by different stresses including salinity and ABA. Therefore, it has been suggested that these EUL proteins are involved in the adaptive response to a hyper-osmotic environment (Moons et al., 1997a). Data generated from microarrays and proteomics approaches revealed that rice EUL proteins are regulated in response to different abiotic and biotic stress conditions (Kawasaki et al., 2001; Dooki et al.,

2006; Ke et al., 2009; Kyndt et al., 2012). *In silico* expression analysis revealed that the EUL type S3 from *Arabidopsis* (being the only type of EUL protein identified in *Arabidopsis*) is up-regulated in response to salinity, osmotic stress and after ABA treatment (Fouquaert et al., 2009a). More evidences have accumulated that EUL proteins are stress related proteins. In response to water stress, the expression of a maize EUL homolog was increased (Riccardi et al., 2004). In addition, it was shown that the EUL proteins from banana contribute towards dehydration tolerance (Carpentier et al., 2007).

It can be concluded, based on above mentioned data that EULs are proteins that can respond to abiotic and biotic stresses. In addition, EULs are synthesized without a signal peptide, and therefore they encode nucleocytoplasmic proteins (except type Sv). Localization analysis performed in tobacco BY2 cells for EEA (Van Hove et al., 2011), in *Arabidopsis* plants for type S3 (from *Arabidopsis*) (Van Hove et al., 2011) and in tobacco as well as in *Arabidopsis* plants for type S2 and type D1 from rice (**chapter 4A**) confirmed the nucleocytoplasmic localization for these EUL proteins. Therefore, it has been suggested that the EUL domain plays a role in stress related physiological processes (Lannoo and Van Damme 2010; Fouquaert and Van Damme 2012).

1.2.5 Rice lectins

1.2.5.1 Lectins with hevein domain(s)

By definition the family of lectins with hevein domain(s) comprises all proteins containing at least one hevein domain. The hevein domain is a small 43 AA protein from the latex of the rubber tree (*Hevea brasiliensis*) (Waljuno et al., 1975). Despite its small size, this domain possesses a fully active chitin-binding site.

The so-called rice lectin (*Oryza sativa* agglutinin, OSA) belongs to the family of lectins with hevein domains. This lectin has been detected in the embryo of several species belonging to the family *Gramineae* including *Oryza sativa* (Peumans and Stinissen 1983). When isolated from rice embryos, it preferentially agglutinated trypsin-treated erythrocytes (Tabary et al., 1984). Inhibition assays revealed that N-acetylglucosamine (GlcNAc) was a potent inhibitor for this lectin (Tsuda 1979). Afterwards, it was shown that OSA (38 kDa) contains four equivalent saccharide-binding sites and confirmed the reaction with GlcNAc (Tabary et al.,

1985). According to the cDNA sequence, OSA is synthesized as a pre-pro-protein consisting of a 28 residue signal peptide, a sequence of 173 AA with four tandemly arrayed hevein domains, and a 26 residue C-terminal pro-peptide with a putative glycosylation site (Wilkins and Raikhel 1989). Biosynthesis studies confirmed the co-translational removal of the signal peptide and glycosylation of the C-terminal propeptide. After removal of the glycosylated C-terminal pro-peptide, the protomers of OSA and some other wild rice species are for 50 to 90% posttranslationally cleaved into two smaller polypeptides of 8 and 10 kDa, respectively. As a result, they contain polypeptides of 18 kDa as well as of 8 and 10 kDa (Van Damme et al., 1998). It is worth mentioning that the insecticidal activity of OSA was tested against the rice brown planthopper (*Nilaparvata lugens*) in an artificial diet assay, but it exhibited no significant antimetabolic effect towards this sap-sucking insect (Powell et al., 1995).

1.2.5.2 GNA-like lectins

The mannose-specific *G. nivalis* agglutinin or GNA was originally isolated from the bulbs of snowdrop (Van Damme et al., 1987). GNA-related lectins are synthesized with an N-terminal signal peptide and a C-terminal pro-peptide (Van Damme et al., 1991; 1998) and thus localize to the vacuoles and/or extracellular spaces. Sequence analysis revealed that GNA orthologs lacking the signal peptide are found in wheat, maize (Van Damme et al., 2004a) and rice. The GNA ortholog from rice (GNA$_{rice}$) can be considered as a naturally occurring nucleocytoplasmic counterpart for the vacuolar GNA. Sequence alignment confirmed that GNA$_{rice}$ corresponds to the mature GNA polypeptide and contains three fully active mannose-binding sites. In contrast to vacuolar GNA, localization analysis revealed that GNA$_{rice}$ was exclusively located in the nucleus and the cytoplasm (Fouquaert et al., 2007). The high sequence identity (37.6%) and similarity (73.4%) with the mature GNA polypeptide suggest that the cytoplasmic and vacuolar GNA homologs from plants are evolutionarily related (Fouquaert et al., 2007).

1.2.5.3 Nictaba-like lectins

Screening of the annotated genome of *Oryza sativa* revealed the presence of eleven types of Nictaba homologs: five types encode proteins consisting of a single Nictaba domain, three types encode chimeric proteins consisting of a Nictaba domain linked to an F-box domain, one type encode a protein kinase domain linked to Nictaba and two types contain multiple

Nictaba domains (investigated by Sofie Van Holle, unpublished data). These proteins have not been investigated in more detail yet.

1.2.5.4 Jacalin-related lectins

In the rice genome, multiple genes encoding jacalin-related proteins have been identified based on genome-wide approaches (Jiang et al., 2010a). *Oryza sativa* agglutinin or Orysata is the jacalin-related lectin which has received most attention. This lectin was detected in rice seedlings after NaCl treatment (Zhang et al., 2000). Sequence analysis revealed that Orysata corresponded to a previously described salt-inducible protein (SalT) (Claes et al., 1990). Orysata is rapidly expressed in roots and sheaths after exposure of whole plants to salt or drought stress, or upon jasmonic acid and abscisic acid (ABA) treatment (Claes et al., 1990; Moons et al., 1997b; Garcia et al., 1998; Hirano et al., 2000; De Souza et al., 2003) but cannot be detected in untreated plants. Interestingly, the lectin is also expressed in excised leaves after infection with an incompatible *Magnaporthe grisea* race (Kim et al., 2003; 2004b; Qin et al., 2003; Liao et al., 2009) as well as during senescence (Lee et al., 2001). Recently, it has been shown that over-expression of Orysata suppressed the growth of invasive hyphae of the fungus *Magnaporthe oryzae*, the rice blast fungus, and subsequently enhanced the resistance of rice plants (Shinjo et al., 2011).

1.2.5.5 Lectins with EUL domain (s)

Analyzing the rice genome resulted in the identification of four types of EUL-related lectins (Table 3A.1). Among the members of the S type, the EUL type S2 (referred to as OrysaEULS2) and the EUL type S3 (OrysaEULS3) were identified. OrysaEULS2 consists of a single EUL domain preceded by medium long (56 AA) unrelated N-terminal domain. OrysaEULS3 resembles the OrysaEULS2 but contains an N-terminal domain with a different and longer sequence (117 AA). Similarly, two members were found in the D type. The EUL type D1 includes two almost identical proteins: OrysaEULD1A and OrysaEULD1B. Both type D1 proteins are composed of two EUL domains separated by a short linker and preceded by an unrelated N-terminal sequence. Similarly, the EUL type D2 (OrysaEULD2) contains two EUL domains, but the linker sequence is longer and the N-terminal domain is different (Fouquaert et al., 2009a).

An extensive screening of the rice expression data reported in literature revealed evidence that the EUL types expressed in rice are responsive to stress treatments, as shown by studies using proteomics analyses with stress-treated rice tissues (Table 1.2).

Table 1.2 Summary of expression data reported for the EUL rice proteins. Up and down indicate the up-regulation and the down-regulation, respectively, of the gene of interest in the rice plants compared to the basal level of this gene in the control plants (non-stressed palnts)

OrysaEULS2				
Tissue	**Expression level**	**Stress**	**Concentration/time**	**Reference**
Shoot	Up	NaCl	150 mM/12 h	(Moons et al., 1997a)
Shoot	Up	ABA	10 to 40 µM/12 h	(Moons et al., 1997a)
OrysaEULS3				
Roots	Up	NaCl	150 mM/48 h	(Cheng et al., 2009)
Panicle	Up	NaCl	50 to 75 mM	(Dooki et al., 2006)
OrysaEULD1A				
Root	Up	NaCl	150 mM/24 h	(Kawasaki et al., 2001)
Root	Up	NaCl	100 mM/3 days	(Moons et al., 1997a)
Scutellar tissue	Up	ABA	1 µM/3 days	(Asakura et al., 2007)
Shoot	R^2	Fungi + bacteria	Bacteria: OD_{600}=1	(Li et al., 2006)
Shoot	Down	Insect	10 insects/seedling/72 h	(Zhang et al., 2004)
Root galls	Down	Root knot nematode	250 nematodes per plant/3 days after inoculation	(Kyndt et al., 2012)
OrysaEULD1B				
Scutellar tissue	Up	ABA	1 µM/3 days	(Asakura et al., 2007)

Root	Up	NaCl	150 mM/2 to 8 h	(Moons et al., 1997a)
Root	Up	ABA	10 to 40 μM/12 h	(Moons et al., 1997a)
Shoot	Up	Fungi + bacteria	Bacteria: OD600=1	(Cheng et al., 2009)
Shoot	Up	Drought	-	(Babu et al., 2002)
shoot	Up	NaCl	140 mM /24 h	(Chao et al., 2005)
Root	Up	NaCl	150 mM/3 h	(Kawasaki et al., 2001)
Root	Up	NaCl	150 mM	(Cheng et al., 2009)
Root	Up	ABA	20 and 100 μM	(Moons et al., 1995)
Transgenic calli	Up	No stress	-	(Takahashi et al., 2005)
Panicle	Exp[1]	No stress	-	(Tang et al., 2005)
OrysaEULD2				
Root	Up	Root rot nematode	400 nematodes per plant/3 days after inoculation	(Kyndt et al., 2012)
Root galls	Up	Root knot nematode	250 nematodes per plant/3 days after inoculation	(Kyndt et al., 2012)
Root galls	Up	Root knot nematode	250 nematodes per plant/7 days after inoculation	(Kyndt et al., 2012)

[1]: expressed under normal conditions. [2]: regulated.

1.3 Stresses

1.3.1 Abiotic stresses

The term abiotic stress is best defined as any non-living factor exerted by the environment on the optimal functioning of plants resulting in altered physiological conditions. The most commonly encountered stress factors are salinity, drought, excess water, temperature extremes (heat, cold and freezing), radiations (high intensity of ultra-violet and visible light), chemicals and pollutants (heavy metals, pesticides, and aerosols), oxidative stress (reactive oxygen species, ozone), wind (sand and dust particles in wind) and nutrient deprivation in soil (Mahajan and Tuteja 2005). Here, we will give an overview about the two most important abiotic stresses with a large negative impact on crop yield and productivity: salinity and drought.

A. Salinity

A saline soil has an electrical conductivity of the saturated paste extract (ECe) above 4dS/m (approximately 40 mM NaCl) (Chinnusamy et al., 2005; Munns 2005). Approximately 20% of irrigated areas (45 Mha) for all cultivated plants are estimated to be affected by salinity (Negrao et al., 2011). This is particularly serious since irrigated areas are responsible for one third of the world's food production. Plants differ in their ability to cope with salinity and according to their tolerance can be divided into two groups: glycophytes and halophytes. Halophytes can generally tolerate salt concentrations higher than 20 ds/m (200 mM NaCl) (Flowers and Colmer 2008) or even higher. However, most plants are glycophytes and cannot tolerate salt concentrations higher than 4dS/m.

Salinity in a certain area increases when the rate of evaporation far exceeds the rate of precipitation. Weathering of rocks also affects salt concentration in the soil. In addition, excessive irrigation and fertilizing soils without proper drainage cause increasing salinity year by year. In addition to the adverse effect of high salt concentrations to plants, salt can alter the basic texture of the soil resulting in decreased soil porosity and consequently reduced soil aeration and water conductance.

Rice is a salt-sensitive crop (Chinnusamy et al., 2005). The threshold for salt stress is 3 dS/m, with a 12% reduction in yield, per dS/m beyond this value.

A.1 Short term response to salinity

Na^+ ions enter the root cells through the symplastic pathway mediated by cation channels/transporters as well as the apoplastic pathway in which Na^+ enters the transpiration stream (Gao et al., 2007). In rice, the apoplastic pathway is more important because rice has limited control of water loss as an aquatic species, and its root anatomical development restricts the symplastic pathway (Ranathunge et al., 2004). In addition the apoplastic leakage was estimated to be 10 times higher in rice compared to wheat (Garcia et al., 1997).

The impact of salinity on plants can be divided into two effects; a rapid osmotic effect and a delayed ionic effect (Munns and Tester 2008). As a consequence of accumulating Na^+ ions in the root region, the osmotic potential is decreased. Due to the difference between the osmotic potential of the cytoplasm inside the root cells and the soil water solution, the osmotic effect reveals direct effects on plant root cells. Osmotic stress decreases the water availability in the soil and subsequently decreases water uptake by the roots resulting in cellular dehydration. This effect is common after all dehydrative stresses (e.g., salinity and drought). Accumulation of Na^+ ions in the plant cell to toxic levels causes the ionic effect. This accumulation happens when the Na^+ exclusion mechanisms are not sufficient to keep the Na^+ ions low inside the cell.

After exposure to high salinity (NaCl >50 mM), rice plants suffer a rapid and temporary drop in stomatal conductance (Moradi and Ismail 2007; Yeo et al., 1991). The same effect was also reported for maize (Neuman 1993), wheat and barley (Passioura and Munns 2000). The influx of salt into roots activates perception and signaling mechanisms that tend to (1) inhibit the entry of further Na^+ into the roots, (2) reduce long-distance Na^+ transport from root to shoot and (3) restore leaf ion homeostasis.

A.2 Long term response to salinity

Plants have developed several mechanisms to cope with high Na^+ concentration inside the cell, in particular osmotic adjustment, exclusion of Na^+ ions and production of antioxidant metabolites.

Osmotic adjustment

As the Na^+ ions accumulate in the apoplastic solution, several mechanisms are triggered to protect the cell from dehydration, protein denaturation and destabilization (Yancey et al., 1982). One of these mechanisms is the accumulation of non-toxic metabolites, known as compatible solutes, in the cytoplasm. These compatible solutes adjust the osmotic potential between the cytosol and the apoplastic solution. Among the compatible solutes different sugars (fructose, glucose and sucrose), complex sugars (trehalose, raffinose and fructans), sugar alcohols (mannitol and glycerol), amino acids and derivatives (proline, glycine-betaine, and proline-betaine) have been suggested to accomplish this function (Bohnert et al., 1995; Flowers and Colmer 2008). In rice, the most remarkably enhanced compatible solute is the amino acid proline. It has been shown that proline accumulated in response to salt stress especially in tolerant rice genotypes, confirming its protective role against hyper-osmotic stress (Demiral and Turkan 2006). On the other hand, the accumulation of compatible solutes in the cytoplasm also helps to balance the ion osmotic pressure in the vacuole, where Na^+ and Cl^- are sequestered.

Exclusion of Na^+ ions

Under normal physiological conditions, plants maintain a relatively high K^+ concentration and a low Na^+ concentration in their cytosol (Binzel et al., 1988). Under salt stress conditions, a high cytosolic K^+/Na^+ ratio is a key feature for maintaining cellular metabolism (Zhu 2003). The cytosolic enzymes in both glycophytes and halophytes are sensitive to salt. For plant cells, the most important way of keeping the cytosolic Na^+ concentration at a low level is to minimize Na^+ influx into the cytosol, and to maximize the Na^+ efflux from the cytosol, either into the apoplast or into the vacuole (Nie et al., 1995; Blumwald 2000; Zhu 2001; Qiu et al., 2004). Although rice is not a good excluder, it still excludes at least 94% of the soil Na^+ from the transpiration stream (Munns 2005), and there is a strong correlation between exclusion capacity and salt tolerance (Lee et al., 2003; Zhu et al., 2004). Once Na^+ enters the cytosol at a toxic level, plant cells can deal with the internal Na^+ by sequestering it either in the apoplast or in the vacuole. Vacuolar compartmentalization is an efficient strategy for plant cells to cope with salinity stress (Tester and Davenport 2003; Fukuda et al., 2004). In rice, salt concentrations in leaves were found to cause different toxicity levels depending on the

genotypes (Flowers et al., 1985; Yeo and Flowers 1983; Yeo et al., 1990). These differences probably reflect differences in excluding excess Na^+ from the cells and/or sequestering the ions in vacuoles.

Production of antioxidant metabolites

It is well established that many environmental stresses including salinity, drought and high temperature cause excessive production of reactive oxygen species (ROS) in the plant cell (Ramamurthy and Nina 2003; Bienert et al., 2006; Hong et al., 2009). This accumulation will cause oxidative damage to different cellular components including membrane lipids, proteins and nucleic acids (Haliwell 1986; McCord 2000; Mittler 2002; Van Breusegem and Dat 2006). Salt-induced ROS are predominantly represented by H_2O_2, both outside and inside the cell (Yang et al., 2007; Pang and Wang 2008). Extracellular ROS production depends on the activity of NADPH oxidases that contribute to salt stress tolerance in *Arabidopsis thaliana* (Ma et al., 2012). In rice, the production of NADPH oxidase–dependent H_2O_2 emerges within several minutes of salt stress (Hong et al., 2009). To scavenge ROS, plants produce enzymatic scavengers like superoxide dismutase, ascorbate peroxidase, glutathione peroxidases, glutathione S-transferase and catalase (Mittler 2002). In addition the plants up-regulate low molecular mass antioxidants such as ascorbate, tocopherol, and carotenoids (Allen 1995). The harmonized activities of the multiple forms of these enzymes in different cell compartments, achieve a balance between the rate of synthesis and removal of ROS, and maintain hydrogen peroxide at the levels required only for cell signaling (Munns and Tester 2008).

A.3 Salinity effect on whole plant level

When the tolerance mechanisms of the plant cell are not sufficiently high to exclude salt from the transpiration stream, leaves will accumulate salt to toxic levels causing their death. Because new leaf growth is supported through the export of carbon from mature leaves, the fate of the plant depends on the balance between the rate at which mature leaves die and young leaves are produced (Munns and Tester 2008). The ultimate goal of any crop plant under stress is to complete the life cycle by reaching successful reproduction. In rice, if the stress is severe (NaCl > 100 mM), plants die before maturity but if the stress is less severe (NaCl < 50 mM), delayed panicle initiation and flowering have been observed (Grattan et al.,

2002). Na$^+$ accumulation in the panicle has a negative impact on some yield parameters such as tillering, spikelet number and grain weight (Khatun and Flowers 1995). At plant level, an improved Na$^+$ partitioning between older tissues/ leaves may protect the developing panicles from excessive Na$^+$ accumulation (Mohammadi-Nejad et al., 2010).

B. Drought

Drought is a meteorological term and is commonly defined as a period without significant rainfall. Generally drought stress occurs when the available water in the soil is reduced and atmospheric conditions cause continuous loss of water by transpiration or evaporation. A definition of drought generally accepted by plant breeders is "a shortfall of water availability sufficient to cause loss in yield" (Price 2002), or "a period of no rainfall or irrigation that affects crop growth (Fukai and Cooper 1995).

Rice is particularly susceptible to drought (Cabuslay et al., 2002; Inthapan and Fukai 1988; Lafitte et al., 2007) due to its small root system, rapid stomatal closure and leaf senescence during mild water stress (Hirasawa 1999). Most high-yielding rice cultivars developed for irrigated conditions are highly susceptible to drought stress as well (Lafitte et al., 2007).

B.1 Drought effect on stomata

Stomatal pores are located in the epidermis of plant leaves. They control the uptake of CO_2 required for photosynthesis and the water loss during transpiration (Schroeder et al., 2001a; Hetherington and Woodward 2003). The first response of all plants to drought stress is the closure of their stomata to prevent the transpirational water loss (MansWeld and Atkinson 1990). Closure of stomata can be hydropassive resulting from direct evaporation of water of the guard cells or hydroactive which is based on metabolic involvement. This hydroactive closure mediated by ABA is a common mechanism that plants use to cope with drought stress (Schroeder et al., 2001b; Zhu 2002; Fan et al., 2004). Environmental conditions that increase the rate of transpiration also result in an increase in the pH of the leaf sap, which can promote ABA accumulation and lead to a reduction of stomatal conductance (Wilkinson and Davies 2002; Davies et al., 2002). ABA promotes the efflux of K$^+$ ions from the guard cells, which results in the loss of turgor pressure leading to stomatal closure. The fact that ABA can act as a long distance communication signal between water deficit roots and leaves,

inducing the closure of stomata was proposed almost three decades ago (Blackman and Davies 1985).

B.2 Drought effect on photosynthesis

Drought causes stomatal closure and decreases the chlorophyll content (Lauriano et al., 2004; Lawlor and Cornic 2002). Subsequently, the rate of photosynthesis is affected and severe drought conditions can decrease the activity of Rubisco which in turn results in limited photosynthesis (Bota et al., 2004). Decreased CO_2 diffusion from the atmosphere to the site of carboxylation is generally considered as the main cause for decreased photosynthesis under moderate drought stress (Chaves et al., 2009; Flexas et al., 2004; Grassi and Magnani 2005). However, it has been also reported that despite a decrease in photosynthetic pigments, the photosynthesis light phase remained unaffected under mild drought stress (Vandoorne et al., 2012). In all cases, plants respond to drought stress by the optimization of CO_2 gain through stomatal aperture while minimizing water loss (Price et al., 2002). The activity of the photosynthetic electron chain is controlled by the availability of CO_2 in the plant and photosystem II activity often declines in parallel under drought conditions (Loreto et al., 1995). Decline in intracellular CO_2 levels results in the over-reduction of components within the electron transport chain and the electrons get transferred to oxygen at photosystem I. This generates ROS including superoxide, hydrogen peroxide and hydroxyl radicals, and plants need to scavenge ROS species as they might lead to photooxidation. ROS accumulation will cause lipid peroxidation in the cell membrane and a series of physiological and biochemical changes associated with direct or indirect oxidative stress in plants and finally influence yield and seed quality (Jonaliza et al., 2004; Chaves and Oliveira 2004). Plants can eliminate the harmful effect of ROS by the synergistic action of protective enzymes and the accumulation of non-enzyme substances (Tang et al., 2004; Tian et al., 2005).

Since photosynthesis is inhibited by drought, the grain filling process becomes dependent on stem reserve utilization (Blum 2005). Stem reserve is a major resource providing carbohydrates and nitrogen for grain filling when the transient photosynthetic source is inhibited by stress. Numerous studies have reported that the stem reserve mobilization capacity is related to yield under drought stress in wheat (Blum 2005). In rice, the same

mechanism was reported to maintain grain yield under drought stress at the grain filling stage (Yang et al., 2001a; 2001b; Yang et al., 2002). This drought tolerance mechanism is stimulated by a decrease in gibberellic acid concentration and an increase in ABA concentration.

B.3 Role of osmolytes in response to drought

As a plant detects a water-deficit stress, it accumulates a variety of osmotically active compounds such as amino acids, sugars and ions inside its cells, resulting in a lowering of the cell osmotic potential (Zhang et al., 1999). As a consequence of this accumulation, the water present in intercellular spaces can enter the cell (Nguyen et al., 1997). This so-called osmotic adjustment was proposed as a potential factor that could enable plants to maintain turgor and survive better at low water status. Osmotic adjustment could contribute to drought tolerance via osmolyte accumulation in the roots that would maintain or increase root development into deeper soil layers, thereby increasing available water for crop use (Manschadi et al., 2006; Sharp et al., 2004).

B.4 Drought resistance in crops

Rice, like other crops, can tolerate drought stress using three different strategies: drought escape, drought avoidance, or drought tolerance (Zhang 2007). A proper timing of the lifecycle, resulting in the completion of the most sensitive developmental stages while water is abundant, is considered to be a drought escape strategy (Price et al., 2002, Zhang 2007). Avoiding water-deficit stress with a root system capable of extracting water from deep soil layers, or by reducing evapotranspiration without affecting yields, is considered as drought avoidance (Price et al., 2002, Zhang 2007). Mechanisms such as osmotic adjustment whereby a plant maintains cell turgor pressure under reduced soil water potential are categorized as drought tolerance mechanisms (Price et al., 2002, Zhang 2007). Drought avoidance mechanisms can be expressed even in the absence of stress and are then considered constitutive. Drought tolerance mechanisms are the result of a response triggered by drought stress itself and are therefore considered adaptative (Hazen et al., 2005). When the stress is terminal and predictable, drought escape through the use of shorter duration varieties is often the preferable method of improving yield potential.

Drought avoidance and tolerance mechanisms are required in situations where the timing of drought is mostly unpredictable (Pantuwan et al., 2002).

B.5 Drought effect on whole plant level

Drought stress is characterized by diminished leaf water potential and turgor loss, closure of stomata and decrease in cell enlargement and growth. Severe water stress may result in the arrest of photosynthesis, disturbance of metabolism and finally the death of the plant (Jaleel et al., 2008).

There are three basic drought patterns affecting rice production: early, intermittent and late drought stresses (Fukai and Cooper 1995). Early droughts often result in delayed sowing or transplanting. Yield reductions from early droughts (occurring during vegetative growth, after establishment but before maximum tillering) are often minimal, and result from a reduction in tiller numbers (Boonjung and Fukai 1996; Jongdee et al., 2006). Intermittent or continuous droughts (occurring between the tillering and flowering stages), may greatly reduce yields despite no apparent drought symptoms (such as e.g. leaf rolling), mainly as a result of reduced leaf expansion and photosynthesis (Fukai and Cooper 1995). When drought occurs during later growing stages (following panicle initiation and especially during flowering), it can cause loss of pollen fertility, spikelet death or abortion of newly formed seeds (Liu et al., 2006). In addition, drought delays the development of the rice plant (Puckridge and O'Toole 1980) and strongly affects morphology (O'Toole and Baldia 1982; O'Toole and Cruz 1980; O'Toole and Moya 1981) as well as physiological processes like transpiration, photosynthesis, respiration and translocation of assimilates to the grain (Turner 1986; Fukai and Cooper 1995). The leaf and root phenology of rice cultivars are known to influence their vegetative response to water deficit (Lafitte et al., 2004).

1.3.2 Biotic stresses

The most common biotic stress factors that plants face are pathogens (viruses, bacteria, and fungi), insects, plant-parasitic nematodes, herbivores and rodents (Mahajan and Tuteja 2005). The discussion below will be limited to the insects and pathogens used in this PhD research.

A. Insects

Insects account for a vast majority of animal species on earth. They can be found in almost all terrestrial and freshwater habitats, from the driest deserts to freshwater ponds, from the canopy of a tropical rainforest to the arctic areas. They affect many aspects of our lives, despite their small size. Moreover, crop yield losses due to the attack of various classes of insects are a problem worldwide.

A. 1 Lepidoptera

Lepidoptera is an order of insects that includes moths and three super-families of butterflies, skipper butterflies, and moth butterflies. The order Lepidoptera constitutes one of the four largest groups of insects, in terms of their diversity. About 180,000 species have been described (Biodiversity Institute of Ontario 2006), but many more remain undiscovered. The total number of species is probably between 300,000 and 500,000 (Kristensen et al., 2007). The larvae of these insects cause a large destruction to agricultural crops.

The beet armyworm, *Spodoptera exigua* (Fig. 1.2a) is a cosmopolitan polyphagous insect. The pest has a wide host range of plants such as vegetables and field crops (Senthil-Nathan et al., 2008). Armyworm larvae are difficult to control with insecticides once an infestation is well established (Wang et al., 2006).

A.2 Hemiptera

The group of Hemiptera is the largest and most heterogeneous order of exopterygotes. Aphids belong to the superfamily Aphidoidea. Aphids are small, soft-bodied insects with long, slender mouthparts with which they pierce stems, leaves, and other tender plant parts to suck out plant fluids. Almost every plant has one or more aphid species which occasionally feed on it. Aphids cause a direct economic damage on crops and indirect effect by transmitting viral diseases (Hogenhout et al., 2008).

The peach-potato aphid, *Myzus persicae* (or green peach aphid) (Fig. 1.2b) is known to transmit over 100 phytopathogenic viruses among 50 different plant families. Many of its hosts include major crops (e.g. sugar beet, beans, brassicas, potatoes, citrus) and on a world-wide scale this species is considered as the most important aphid pest (Mackauer and Way 1976).

The pea aphid, *Acyrthosiphon pisum* (Fig. 1.2c) is a non-host-alternating aphid that is associated with a wide range of legume species including many important forage and vegetable crops (Van Emden and Harrington 2007). Pea aphids consume the nutritive fluid transported in the phloem of their host plant. This is done directly by inserting their stylet into the phloem tissue. Pea aphids are particularly important as potential migratory pests because they have a wide host range and a parthenogenetic reproduction (Losey and Eubanks 2000).
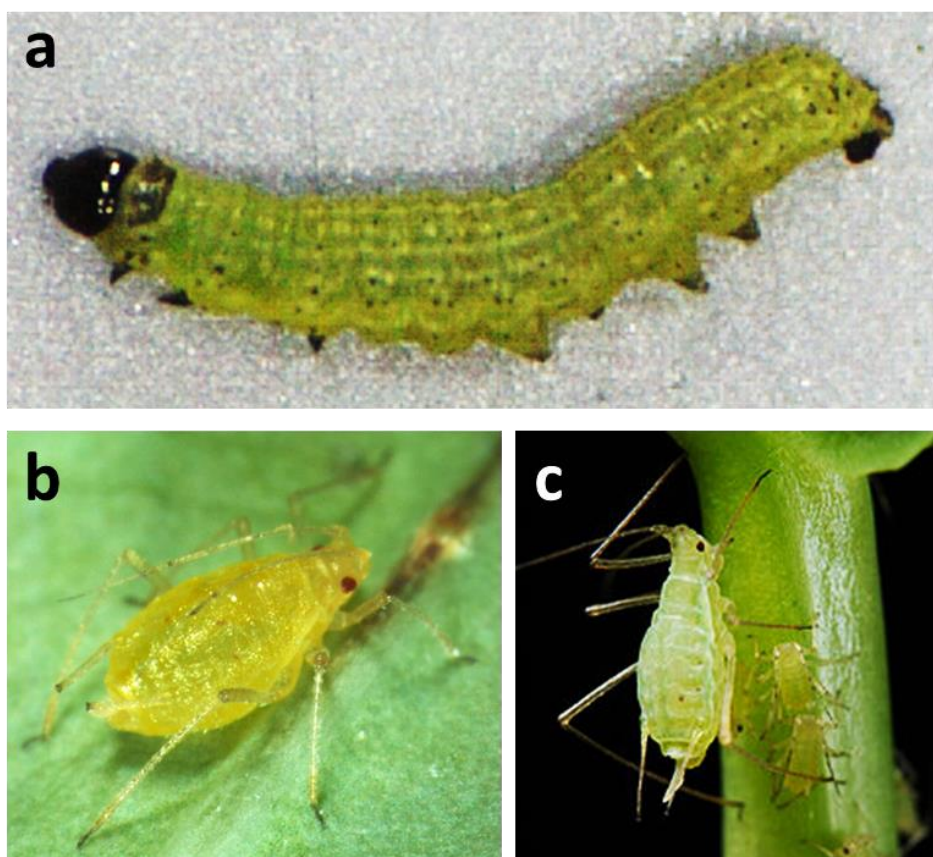
Fig. 1.2 Pictures showing (a): the second larval stage of beet armyworm, *Spodoptera exigua*. (b) the peach-potato aphid, *Myzus persicae*. (c): the pea aphid, *Acyrthosiphon pisum*.

B. Pathogens

Because of the public availability of the genome sequences of rice (Sasaki et al., 2005) and its two major pathogens, *Magnaporthe oryzae* (Dean et al., 2005) and *Xanthomonas oryzae* pv. *oryzae* (Lee et al., 2005), the rice blast and bacterial blight pathosystems have become the

genetic models for understanding host-pathogen interactions and co-evolution for cereals (Nino-Liu et al., 2006).

B.1 *Xanthomonas oryzae*

Bacterial leaf blight of rice, caused by the gram-negative bacterium *Xanthomonas oryzae* pv *oryzae*, is responsible for huge economic losses. This pathogen is classified as a biotroph (Parbery 1996). Bacterial leaf blight is known to occur in all rice growing areas and is exceptionally severe in Asia. It is reported to reduce the annual production by as much as 60% in India and 50% in Japan (http://www.knowledgebank.irri.org). The symptoms in adult plants appear as water-soaked yellowish stripes on leaf blades or starting at leaf tips which increase in length and width killing the infected leaves. Infected plants produce sterile and empty panicles and in severe cases the plant wilts and dies.

B.2 *Magnaporthe oryzae*

Rice blast, caused by *Magnaporthe oryzae*, is one of the most devastating diseases in rice growing regions worldwide, causing 11-15% yield loss annually (Baker et al., 1997). This pathogen is classified as a hemi-biotroph (Parbery 1996). Within the plant cell, the fungus faces two different fates. In an incompatible interaction, resistance gene products recognize the corresponding avirulence gene products from the invading pathogen and invoke a series of defense responses to restrict pathogen growth (Ahn et al., 2005). In a compatible interaction, the host plant mobilizes defense responses much later, resulting in visible blast lesions (Tucker and Talbot 2001). The blast fungus infects rice plants at all stages and in all tissues, such as leaves, stems, nodes, panicles and roots (Wilson and Talbot 2009; Sesma and Osbourn 2004).

B.3 *Pseudomonas syringae*

The plant-pathogenic species *Pseudomonas syringae* is a complex taxon consisting of strains with a range of different phenotypic, biochemical, and genetic characteristics, as well as a different host range. These gram negative bacteria infect a wide variety of plants and can cause necrotic symptoms in leaves, stems and fruits (Rico and Preston 2008). The bacterial pathogen *Pseudomonas syringae* pv. tomato (PstDC3000) is virulent to *Arabidopsis* plants (Whalen et al., 1991).

C. Glycosylation patterns

Glycosylation is a post-translational modification that strongly influences protein folding, secretion, cellular localization, and biological activity (Helenius and Aeb 2004). It occurs by the attachment of a glycan to a protein either at an asparagine residue (N-X-S/T motif, where X denotes any amino acid except proline) (Bause 1983), termed N-glycosylation, or at a hydroxylysine, hydroxyproline, serine, or threonine, called O-glycosylation (Carraway and Hull 1989). Glycan structures are very diverse because of the branched structure, differences in glycosidic linkages, heterogeneity in glycan profile for a specific glycoprotein and multiple modifications that may occur on specific sugar monomers.

C.1 Insects

In insects, the N-linked glycan structures are mainly of the high-mannose or paucimannose type but small amounts of complex glycans have been identified. These hybrid and complex structures are present in small amounts and often only for a specific developmental stage of the insect (Rendic et al., 2008; Tiemeyer et al., 2008). For instance, sialylated glycan structures have been reported in *Drosophila* embryos. The presence of these complex glycan structures in the central nervous system (Aoki et al., 2007; Koles et al., 2007) suggests a crucial role of these carbohydrate structures in the insect.

C.2 Pathogens

In fungi, both N- and O-glycosylation patterns have been reported. The N-glycans are most likely of the high-mannose type (Jacobs et al., 2009). O-Mannosylation is commonly found in fungi (Strahl-Bolsinger et al., 1999).

Although bacteria are prokaryotes, it is now established that they possess both N-linked and O-linked glycosylation pathways that display many commonalities with their eukaryotic and archaeal counterparts (Nothaft and Szymanski et al., 2010).

1.3.3 ABA as a stress hormone

ABA is a plant stress hormone and one of the foremost important signaling molecules in plants, which plays versatile functions in regulating many developmental processes and adaptive stress processes (Santner et al., 2009; Cutler et al., 2010). ABA is accumulated

rapidly in response to abiotic stresses and therefore it can mediate many stress responses that help the plant to survive. The first prerequisite to perform this job is that ABA production should be rapidly triggered by the stress to avoid any inhibition of plant growth and functions. The second prerequisite is that ABA should be rapidly degraded and deactivated once the stress is relieved allowing normal plant growth and functions to resume.

Indeed, ABA is an important signal for triggering plant responses to adverse environmental conditions during vegetative growth (Leung and Giraudat 1998; Nambara and Marion-Poll 2005). ABA coordinates many stress responses, including the immediate stomatal closure, the osmolyte accumulation and the induction of the synthesis of stress related proteins, such as late embryogenesis abundant and heat shock proteins and reactive oxygen scavengers. However, whereas many abiotic-stress-inducible genes are controlled by ABA, some are not, which indicates that both ABA-dependent and ABA-independent regulatory systems are involved in stress-responsive gene expression (Bray et al., 2000; Zhu 2002). Moreover, it has recently been shown that ABA applied exogenously at 100 µM concentration prior to and during the salt-stress period induced salt tolerance in rice (Sripinyowanich et al., 2013).

The role of ABA in plant innate immunity is still poorly understood. Recently, ABA has emerged as a key signaling molecule in plant-pathogen interactions (Asselbergh et al., 2008). In this respect, both positive and negative effects of ABA on disease resistance have been reported but its effect as a negative regulator of immunity is predominant (Cao et al., 2011; Asselbergh et al., 2008). In rice, exogenous ABA can suppress basal immunity to both *X. oryzae* pv. *oryzae* and *M. oryzae*. When the infection by these pathogens is established, extensive reprogramming of ABA-response and -biosynthesis genes takes place, suggesting that these pathogens influence the rice ABA pathway (Jiang et al., 2010b; Koga et al., 2004, Yazawa et al., 2012; Xu et al., 2013). Interestingly, ABA also antagonizes defense against the migratory root rot nematode *Hirschmanniella oryzae* (Nahar et al., 2012). In contrast, ABA enhanced the basal resistance against the rice brown spot pathogen *Cochliobolus miyabeanus* by preventing the fungus from hijacking the ethylene pathway (De Vleesschauwer et al., 2010). Therefore, ABA appears to play an ambiguous role in the rice

immune signaling network, acting either as a positive or as a negative regulator of disease resistance.

1.4 Lectins are putative candidates to improve stress tolerance of crops

Plants have evolved protective mechanisms that allow them to successfully resist unfavorable conditions. Current engineering strategies to enhance the tolerance of important crops rely on the transfer of one or several genes that are either involved in signaling and regulatory pathways, or that encode enzymes present in pathways leading to the synthesis of functional and structural protectants, such as osmolytes and antioxidants, or that encode stress-tolerance conferring proteins (Wang et al., 2003; Vinocur and Altman 2005; Valliyodan and Nguyen 2006; Sreenivasulu et al., 2007; Kathuria et al., 2007). Among the most important stress-tolerance conferring proteins are the plant lectins.

In higher plants, there is a group of proteins, called lectin receptor-like kinases (LecRLKs). These proteins play a role in stress tolerance of the plant as well as in the growth development. The basic structure of LecRLK protein comprises of a lectin and a kinase domain, which are interconnected by transmembrane region. Based on a genome wide analysis, 173 genes of different types of LecRLKs have been identified within the rice genome (Vaid et al., 2012). Multiple lectins have been overexpressed to improve the tolerance of crops against abiotic and biotic stresses. Recently, it has been shown that the overexpression of these LecRLKs confers tolerance towards stresses e.g. the overexpression of a plasma membrane-localized lectin-like protein kinase 1 conferred salinity tolerance of transgenic *Arabidopsis* plants (Huang et al., 2013).

Furthermore, the insecticidal activity of plant lectins against insects of the orders Coleoptera, Diptera, Lepidoptera, and Hemiptera has been studied (Macedo et al., 2003; Oliveira et al., 2011; Vandenborre et al., 2011b). Several plant lectins have been used to increase the tolerance of the rice plant against the most important biotic factors affecting rice yield. Expression of some mannose binding lectins like GNA and *Allium sativum* leaf lectin (ASAL) in rice plants enhanced the tolerance against some sap-sucking insects on rice, such as brown planthopper (*Nilaparvata lugens*) and green leafhopper (*Nephotettix*

nigropictus) (Rao et al., 1998; Saha et al., 2006). Moreover, expressing the garlic leaf lectin ASAL in transgenic rice plants also significantly reduced the infection incidence of rice tungro diseases; tungro bacilliform virus and rice tungro spherical virus caused by co-infection of green leafhopper (Saha et al., 2006). In addition, transgenic rice plants overexpression the Pi-d2, a receptor-like kinase protein with mannose specific binding lectin domain, confer race-specific resistance to the *M. oryzae* strain, ZB15 (Chen et al., 2006).

The discovery of the inducible plant lectins opens the door for the possibility to exploit these proteins in increasing the tolerance of crops towards stresses.

Chapter 2

Characterization of the lectins under study expressed in *Pichia pastoris*

Chapter 2A

Expression analysis of the nucleocytoplasmic lectin'Orysata' from rice in *Pichia pastoris*

Redrafted from:

Al Atalah B, Fouquaert E, Vanderschaeghe D, Proost P, Balzarini J, Smith DF, Rougé P, Lasanajak Y, Callewaert N, Van Damme EJM (2011) Expression analysis of the nucleocytoplasmic lectin "Orysata" from rice in *Pichia pastoris*. FEBS J 278: 2064–2079

2A.1 Abstract

The *Oryza sativa* lectin, abbreviated Orysata is a mannose-specific, jacalin-related lectin expressed in rice plants after exposure to certain stress conditions. Expression of a fusion construct containing the rice lectin sequence linked to enhanced green fluorescent protein (EGFP) in BY2 tobacco cells revealed that Orysata is located in the nucleus and the cytoplasm of the plant cell, indicating that it belongs to the class of nucleocytoplasmic jacalin-related lectins. Since the expression level of Orysata in rice tissues is very low the lectin was expressed in the methylotrophic yeast *Pichia pastoris* with the *Saccharomyces* α-factor sequence to direct the recombinant protein into the secretory pathway and express the protein into the medium. Approximately 12 mg of recombinant lectin was purified per liter medium. SDS–PAGE and Western blot analysis showed that the recombinant lectin exists in two molecular forms. Far Western blot analysis revealed that the 23 kDa lectin polypeptide contains an N-glycan which is absent in the 18.5 kDa polypeptide. Characterization of the glycans present in the recombinant Orysata revealed high-mannose structures, Man9-11 glycans being the most abundant. Glycan array analysis showed that Orysata interacts with high-mannose as well as with more complex N-glycan structures. Orysata has potent anti-human immunodeficiency virus and anti-respiratory syncytial virus activity in cell culture compared to other jacalin-related lectins.

2A.2 Introduction

The family of jacalin-related lectins (JRLs) groups all proteins that possess one or more domains equivalent to 'jacalin', a galactose-binding protein from jack fruit (*Artocarpus integrifolia*) seeds (Sastry et al., 1986). In the last decade many JRLs have been identified which resulted in a subdivision of this family into two groups: the galactose binding and the mannose binding lectins. In contrast to the galactose-binding JRLs that are synthesized on the endoplasmic reticulum and follow the secretory pathway to accumulate in protein storage vacuoles, the mannose-binding JRLs are synthesized and located in the cytoplasm (Peumans et al., 2000b).

The very first inducible lectin to be purified and characterized was a mannose-specific JRL from NaCl-treated rice seedlings, called *Oryza sativa* agglutinin or Orysata (Zhang et al., 2000). Sequence analysis revealed that Orysata corresponded to a previously described salt-inducible protein (SalT) (Claes et al., 1990) and can be classified in the group of JRLs. Orysata cannot be detected in untreated plants but is rapidly expressed in roots and sheaths after exposure of whole plants to salt or drought stress, or upon jasmonic acid and ABA treatment (Claes et al., 1990; De Souza et al., 2003; Moons et al., 1995). Interestingly, the lectin is also expressed in excised leaves after infection with an incompatible *Magnaporthe grisea* strain (Kim et al., 2003; Qin et al., 2003) as well as during senescence (Lee et al., 2001). Since Orysata is expressed at very low levels in certain plant tissues and only after exposure to stress, the purification of the lectin is cumbersome and requires huge amounts of plant material.

In the last decades the methylotrophic yeast *Pichia pastoris* has become the leading yeast vehicle for the production of a broad range of proteins (Cereghino and Cregg 2000). Heterologous protein expression in *Pichia* is controlled by the alcohol oxidase 1 (AOX1) promoter. Expression of the AOX1 gene is tightly regulated and induced by methanol to high levels (Elias et al., 1985; Hartner and Glieder 2006). A variety of lectins were among the proteins reported to be successfully expressed in *P. pastoris*. For example, Raemaekers et al. (1999) described the successful expression of the legume lectin *Phaseolus vulgaris* agglutinin (PHA) and the GNA-related lectin from snowdrop (*Galanthus nivalis* agglutinin) in *P. pastoris*. A glucose-mannose-binding legume lectin from the seeds of *Canavalia brasiliensis*, a

homolog of the classical vacuolar Con A was also expressed by the yeast *P. pastoris* (Bezerra et al., 2006). Oliveira et al. (2008) described the expression of the JRL from breadfruit seeds (*Artocarpus incisa*) in *Pichia*. In 2007 the first nucleocytoplasmic lectin from tobacco (Nictaba) related to the Cucurbitaceae lectins was expressed and purified from *P. pastoris* (Lannoo et al., 2007a). More recently, the first nucleocytoplasmic GNA homolog from plants (GNA$_{maize}$) was expressed in *P. pastoris* (Fouquaert et al., 2009b).

In this paper we describe the heterologous expression of Orysata, a JRL from rice. Based on a detailed analysis of its sequence, this lectin was predicted to locate to the nucleocytoplasmic compartment of plant cells, as shown by expression of a fusion protein in tobacco cells. Furthermore, the successful expression of the His-tagged Orysata in the yeast *P. pastoris* allowed purifying sufficient amounts of the lectin to study in detail the molecular structure of the protein, its carbohydrate-binding specificity and its antiviral activity. Interestingly, antiviral assays showed that Orysata is active against human immunodeficiency virus (HIV) as well as respiratory syncytial virus (RSV), indicating that the lectin may qualify as a microbicide agent.

2A.3 Materials and methods

2A.3.1 Construction of the EGFP-fusion vector for expression analysis in tobacco cells

The coding sequence for Orysata (Genbank accession number CB632549) was amplified by PCR using the cDNA clone encoding Orysata as a template. The primers for amplification of Orysata were ORY-f1 (5'AAAAAGCAGGCTTCACGCTGGTGAAGATTGGCCTG3') and ORY-r1 (5' AGAAAGCTGGGTGTCAAGGGTGGACGTAGATGCC3'). The PCR program was as follows: 5 min 94°C, 25 cycles (15 sec 94°C, 30 sec 65°C, 24 sec 72°C), 5 min 72°C. PCR was performed in a 50 µl reaction volume containing 40 ng DNA template, 10x DNA polymerase buffer, 10 mM dNTPs, 5 µM of each primer and 0.625 U Platinum Pfx DNA Polymerase (Invitrogen) using an AmplitronIIR Thermolyne apparatus (Dubuque, Iowa, USA). The PCR product was 1:10 diluted and used as a template in an additional PCR, using attB-primers EVD 2 (5'GGGGACAAGTTTGTACAAAAAAGCAGGCT3') and EVD 4 (5'GGGGACCACTTTGTACAAGAAA GCTGGGT3') in order to complete the attB recombination sites. The reaction mixture was as

described for previous PCR. The cycle conditions were as follows: 2 min at 94°C, 5 cycles each consisting of 15 sec at 94°C, 30 sec at 50°C, 30 sec at 72°C, 20 cycles with 15 sec at 94°C, 30 sec at 55°C, 30 sec at 72°C, and a final incubation of 5 min at 72°C. Subsequently, the BP reaction was performed using the pDONR221 vector (Invitrogen). After sequencing of the resulting entry clone, the LR reaction was done with the pK7WGF2 destination vector (Karimi et al., 2002) to fuse Orysata sequence C-terminally to EGFP. Overexpression of EGFP alone was achieved using the pK7WG2 destination vector (Karimi et al., 2002).

Tobacco BY-2 cells were transiently transformed with the EGFP-fusion construct by means of particle bombardement and the expression was analyzed by confocal laser microscopy as described by Fouquaert et al. (2007).

2A.3.2 Expression of Orysata in *P. pastoris*

The EasySelect *Pichia* Expression Kit from Invitrogen was used to clone and express Orysata in the *P. pastoris* strain X-33 (Invitrogen, Carlsbad, CA USA). To achieve secretion of the recombinant protein into the culture medium, the E. coli/*P. pastoris* shuttle vector pPICZαB containing an α–mating sequence from *Sacharomyces cerevisiae* was used. This vector contains a polyhistidine tag located downstream from the multiple cloning site. The coding sequence for Orysata was amplified by PCR starting from the Bluescript vector containing the cDNA encoding Orysata (Genbank accession number CB632549) using primers EVD 519 (5'GGCGGACTGCAGCAATGACGCTGGTGAAGATTGGCCTGT3') and EVD 518 (5'CCCGCTTTCTAG AATAGGGTGGACGTAGATGCCAATTGCG3'). The PCR conditions were: 2 min denaturation at 94°C, 25 cycles of 15 sec 94°C, 30 sec 55°C, 1 min 72°C, ending with additional 5 min elongation at 72°C. The amplified Orysata sequence was cloned as a PstI/XbaI fragment in the shuttle vector pPICZαB and transformed in *E. coli* Top10F cells using heat shock transformation. Afterwards, *E. coli* transformants were selected on LB agar plates containing zeocin (25 µg/ml). The plasmids were purified using the E.Z.N.A. Plasmid Mini kit I (Omega Bio-Tek, Georgia, USA). Finally, the sequence of the fusion construct was verified by sequencing using 5' and 3' AOX1 specific primers (forward EVD 21, 5'GACTGGTTC CAATTGACAAGC3' and reverse EVD 22, 5'GCAAATGGCATTCTGACATCC3'), carried out by LGC Genomics GmbH, Berlin, Germany.

2A.3.3 *Pichia* transformation and expression analysis on a small scale

The plasmid DNA from *E. coli* cells was purified and linearized using the restriction enzyme SacI (Fermentas, St Leon-Rot, Germany) with overnight incubation at 37°C. After linearization, 10 μg of the expression vector was transformed into the *Pichia* strain X-33 via electroporation (Bio-Rad, Hercules CA, USA) using the following pulse settings: 25 μF, 1.5 kV and 200 Ω. Transformants were selected on YPDS plates (1% yeast extract, 2% peptone, 2% dextrose, 1 M sorbitol, 2% agar) containing 100 μg/ml zeocin. Genomic DNA was extracted from *Pichia* transformants as reported before (Schroder et al., 2007). The integration of the Orysata sequence in the chromosomal AOX1 locus of *P. pastoris* was confirmed by PCR using the AOX1 primers EVD 21 and EVD 22, and the following parameters: 2 min 95°C, 30 cycles of 1 min 95°C, 1 min 55°C, 1 min 72°C, ending with an elongation step of 7 min at 72°C. For expression analysis, several colonies were inoculated in 5 ml BMGY medium, i.e. 1% yeast extract, 2% peptone, 1.34 % yeast nitrogen base with ammonium sulfate and without amino acids, 4 x 10-5% biotin, 100 mM potassium phosphate (pH 6.0) and 1% glycerol, and grown at 30°C in a shaker incubator at 200 rpm for 24 h. Afterwards, *Pichia* cells were washed with sterilized water and transferred to the BMMY medium (BMGY medium supplemented with 1% of methanol instead of 1% of glycerol). Induction of the culture was achieved by adding 100% methanol (2% final concentration) for three successive days once in the morning and once in the evening. Protein profiles in the medium and the cell pellet were compared. Proteins in the culture medium were analyzed after trichloroacetic acid precipitation (10% final concentration) by SDS-PAGE and Western blot analysis.

2A.3.4 Large scale culture and purification of Orysata

Transformed *P. pastoris* X-33 colonies were inoculated into 5 ml BMGY medium and grown for 24 h at 30°C in a rotary shaker at 200 rpm. Afterwards, cultures were transferred to 50 ml BMGY in 250 ml Erlenmeyer flasks and allowed to grow until the culture reached an optical density between 2 and 6 at 595 nm (16 to 24 h). *Pichia* cells were washed with sterilized water and re-suspended in 200 ml of BMMY medium. The culture was allowed to grow for 72 h in a 500 ml Erlenmeyer flask under the same conditions as before. Every 24 h, 100% methanol was added to the culture twice a day as indicated above (2% final concentration). After three days of methanol induction, the culture was centrifuged for 10 min at 3,000 g

and the supernatant was brought to 80% ammonium sulphate for protein precipitation and stored at 4°C until use. Five 200 mL cultures were pooled for one purification of recombinant Orysata. Purification of the lectin was achieved in three chromatographic steps. After precipitating the protein by centrifugation for 15 min at 5,000 g the resulting pellet was re-suspended in 150 ml 20 mM 1,3 diaminopropane. After overnight dialysis against 20 mM 1,3 diaminopropane, the supernatant was loaded on a Q Fast Flow column (GE Healthcare, Uppsala, Sweden) equilibrated with 20 mM 1,3 diaminopropane. After washing the column with 20 mM 1,3 diaminopropane, elution of the bound proteins was achieved using 100 mM Tris-HCl (pH 8.7) containing 0.5 M NaCl. Subsequently, the eluted fractions were pooled and imidazole was added to a final concentration of 25 mM. The protein sample eluted from the Q Fast Flow column was applied on a Ni-Sepharose column (GE Healthcare) equilibrated with start buffer (0.1 M Tris pH 7.0 containing 0.5 M NaCl and 25 mM imidazole) to purify the His-tagged protein. After washing the Ni-Sepharose column using the start buffer, proteins were eluted using the elution buffer (0.1 M Tris pH 7.0 containing 0.5 M NaCl and 250 mM imidazole). Finally, fractions eluted from Ni-Sepharose were diluted 5 times with phosphate buffered saline (1X PBS: 1.5 mM KH_2PO4, 10 mM Na_2HPO4, 3 mM KCl, 140 mM NaCl, pH 7.4) and applied on a mannose-Sepharose 4B column equilibrated with PBS. After washing the column with PBS, the lectin fraction was eluted using 20 mM 1,3 diaminopropane. The purity of the protein samples was verified by SDS-PAGE and/or western blot analysis after each purification step.

2A.3.5 N-terminal sequence analysis

A sample from the affinity purified Orysata was analyzed by SDS-PAGE, electroblotted onto a ProBlot[TM] polyvinylidene difluoride membrane (Applied Biosystems, Foster City, CA, USA) and visualized by staining with 1:1 mix of Coomassie Brilliant Blue and methanol. Bands of interest were excised from the membrane and the N-terminal sequence determined by Edman degradation on a capillary Procise 491cLC protein sequencer without alkylation of cysteines (Applied Biosystems).

2A.3.6 Agglutination assay

To examine the lectin activity, an agglutination assay was performed using trypsin-treated rabbit red blood cells (BioMérieux, Marcy l'Etoile, France). Therefore 10 µl of the purified protein (165 µg/ml), 10 µl of 1 M ammonium sulfate and 30 µl of trypsinized erythrocytes were mixed in a glass tube. The negative control contained 10 µl PBS, 10 µl 1 M ammonium sulfate and 30 µl trypsinized erythrocytes. After a few minutes agglutination was observed as clumping of the cells at the bottom of the glass tube. Samples that yielded no visible agglutination activity after incubation for 1 h were regarded as lectin negative. Dilution series of the lectin were analyzed to determine its agglutination titer.

2A.3.7 Carbohydrate inhibition test

Several carbohydrates (mannose, trehalose, glucose, galactose, N-acetylglucosamine (GlcNAc) or methyl mannopyranoside, at 0.5 M) and glycoproteins (ovomucoid, asialomucin or thyroglobulin, at 10 mg/ml) were used to test the carbohydrate specificity of the recombinant Orysata. Therefore 10 µl of the purified lectin were mixed with 10 µl of aliquots of a serially twofold diluted carbohydrate or glycoprotein solution. After incubation for 10 min at room temperature, 30 µl trypsin-treated erythrocytes were added. Agglutination activity was assessed visually after incubation for 1h at room temperature.

2A.3.8 Glycan array screening

The microarrays are printed as described previously (Blixt et al., 2004) and version 4.2 with 511 glycan targets was used for the analyses reported here (https://www.functionalglycomics.org/static/consortium/resources/resourcecoreh8.shtml). The printed glycan array contains a library of natural and synthetic glycan sequences representing major glycan structures of glycoproteins and glycolipids. Recombinant Orysata containing a His tag was purified from *P. pastoris* and detected using a fluorescent-labeled anti-His monoclonal antibody (Qiagen, Valencia, CA). The lectin was diluted to desired concentrations in binding buffer (Tris-bufferred saline containing 10 mM $CaCl_2$, 10 mM $MgCl_2$, 1% BSA, 0.05% Tween 20) and 70 µl of the lectin solution was applied to separate microarray slides. After 60 min incubation under a cover slip in a humidified chamber at room temperature, the cover slip was gently removed in a solution of Tris-buffered saline

containing 0.05% Tween 20 and washed by gently dipping the slides 4 times in successive washes of Tris-buffered saline containing 0.05% Tween 20, and Tris-buffered saline. To detect bound lectin, the labeled anti-His antibody (70 µl at one µg/ml in binding buffer) was applied to the slide under a cover slip. After removal of the coverslip and gentle washing of the slide as described above, the slide was finally washed in deionized water and spun in a slide centrifuge for approximately 15 sec to dry. The slide was immediately scanned in a PerkinElmer ProScanArray MicroArray Scanner using an excitation wavelength of 488 nm and ImaGene software (BioDiscovery, Inc., El Segundo, CA) to quantify fluorescence. The data are reported as average Relative Fluorescence Units (RFU) of six replicates for each glycan presented on the array after removing the highest and lowest values. The results for Orysata were compared to the glycan array data obtained for the mannose-binding JRLs purified from *Calystegia sepium* rhizomes (Calsepa), and *Morus nigra* bark (Morniga M) (Nakamura-Tsuruta et al., 2008).

2A.3.9 Antiviral assays

Human lymphocyte CEM cells (5×10^5 cells per ml) were suspended in fresh culture medium [RPMI-1640 (Gibco, Paisley, UK), supplemented with 10% foetal calf serum, 2 mM L-glutamine and 0.075% $NaHCO_3$] and exposed to HIV-1(IIIB) (provided by R.C. Gallo at that time at the NIH, Bethesda, MD) or HIV-2(ROD) (provided by L. Montagnier at that time at the Pasteur Institute, Paris, France) at 100 x the CCID50 per ml of cell suspension. Then, 100 µl of the infected cell suspension was transferred to 200 µl-microplate wells, mixed with 100 µl of the appropriate dilutions of the test compounds, and further incubated at 37°C. After 4 days, giant (syncytium) cell formation was recorded microscopically in the CEM cell cultures, and the number of giant cells was estimated as the percentage of the number of giant cells present in the non-treated virus-infected cell cultures (~ 50 to 100 giant cells in one microscopic field when examined at a microscopic magnitude of 100 x). The 50% effective concentration (EC50) corresponds to the compound concentration required to prevent syncytium formation by 50%. The 50% cytostatic concentration (CC50) corresponds to the compound concentration required to inhibit CEM cell proliferation by 50%. In the co-cultivation assays, 5×10^4 persistently HIV-1-infected human lymphocyte HUT-78 cells (designated HUT-78/HIV-1 (IIIB) were mixed with 5×10^4 human lymphocyte SupT1 cells,

along with appropriate concentrations of the test compounds. After 24-36 h, marked syncytium formation was reached in the control cell cultures, and the number of syncytia was determined under the microscope. The anti-respiratory syncytial virus (RSV strain Long) assay was based on inhibition of virus-induced cytopathicity in human cervix carcinoma HeLa cell cultures. Confluent cell cultures were inoculated with 100 CCID50 of virus (1 CCID50 being the virus dose to infect 50% of the cell cultures) in the presence of varying concentrations of the test compounds. Viral cytopathicity was recorded as soon as it reached completion in the control virus-infected cell cultures that were not treated with the test compounds.

2A.3.10 Molecular modeling and docking

Homology modeling of Orysata was performed on a Silicon Graphics O2 10000 workstation, using the programs InsightII, Homology and Discover (Accelrys, San Diego CA, USA). The atomic coordinates of banana lectin complexed to mannose (code 1X1V) (Singh et al., 2005) were taken from the RCSB Protein Data Bank (Berman et al., 2000) and used to build the three-dimensional model of Orysata. The amino acid sequence alignment was performed with CLUSTAL-X (Thompson et al., 1997) and the Hydrophobic Cluster Analysis (Gaboriaud et al., 1987) plot was generated by the mobile server at (http://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py?form=HCA) to recognize the structurally conserved regions common to Orysata and banana lectin. Steric conflicts resulting from the replacement or the insertion of some residues in the modeled lectin were corrected during the model building procedure using the rotamer library (Ponder et al., 1987) and the search algorithm implemented in the Homology program (Mas et al., 1992) to maintain proper side-chain orientation. Energy minimization and relaxation of the loop regions were carried out by several cycles of steepest descent using Discover3. After correction of the geometry of the loops using the minimize option of TurboFrodo, a final energy minimization step was performed by 150 cycles of steepest descent using Discover 3, keeping constrained the amino acid residues forming the carbohydrate-binding site. The program TurboFrodo (Bio-Graphics, Marseille, France) was used to draw the Ramachandran plots (Ramachandran et al., 1968) and perform the superimposition of the models. PROCHECK (Laskowski et al., 1993) was used to check the stereochemical quality of the three-dimensional model: 82.8%

of the residues were assigned to the favorable regions of the Ramachandran plot (84.6% for banana lectin), except for three residues Ser20, Glu61 and Tyr105, which occur in the non-allowed region of the plot. Using ANOLEA (Melo et al., 1997a) to evaluate the model, only 7 residues over 146 (vs. 3 over 137 for the banana lectin 1X1V used as a template) exhibited an energy over the threshold value

The docking of methyl mannose (MeMan) into the carbohydrate-binding sites of Orysata and other JRLs was performed with the program InsightII (Accelrys, San Diego CA, USA). The lowest apparent binding energy (Ebind expressed in kcal.mol-1) compatible with the hydrogen bonds (considering Van de Waals interactions and strong [2.5 Å < dist(D-A) < 3.1 Å and 120° < ang(D-H-A)] and weak [2.5 Å < dist(D-A) < 3.5 Å and 105° < ang(D-H-A) < 120°] hydrogen bonds; with D: donor, A: acceptor and H: hydrogen) found in the Man-banana lectin complex (RCSB PDB code 1X1V) (Singh et al., 2005) was calculated using the forcefield of Discover3 and used to anchor the pyranose ring of the sugars into the binding sites of the lectin. The position of mannose observed in the Man-banana lectin complex was used as the starting position to anchor mannose in the carbohydrate-binding sites of Orysata. Mannose (Man) was similarly docked into the saccharide-binding site of Calsepa (RCSB PDB code 1OUW) (Bourne et al., 2004). Cartoons showing the docking of Man/MeMan in the mannose-binding sites of the lectins were drawn with PyMol (http://www.pymol.org).

2A.3.11 Analytical methods

The protein content was estimated using the Coomassie (Bradford) Protein Assay Kit (Thermo Fischer Scientific, Rockford, IL USA), based on the Bradford dye-binding procedure (Bradford 1976). SDS-PAGE was performed using 15% polyacrylamide gels under reducing conditions as described by Laemmli (Laemmli 1970). Proteins were visualised by staining with Coomassie Brilliant Blue R-250. For western blot analysis, samples separated by SDS-PAGE were electrotransferred to 0.45 μm polyvinylidene fluoride (PVDF) membranes (BiotraceTM PVDF, PALL, Gelman Laboratory, Ann Arbor, MI USA). After blocking the membranes in Tris-Buffered Saline (TBS: 10 mM Tris, 150 mM NaCl and 0.1% (v/v) Triton X-100, pH 7.6) containing 5% (w/v) milk powder, blots were incubated for 1 h with a mouse monoclonal anti-His (C-terminal) antibody (Invitrogen), diluted 1/5000 in TBS. The secondary antibody was a 1/1000 diluted rabbit anti-mouse IgG labelled with horse radish peroxidase

(Dako Cytomation, Glostrup, Denmark). Immunodetection was achieved by a colorimetric assay using 3,3'-diaminobenzidine tetrahydrochloride (Sigma-Aldrich, St. Louis Missouri, USA) as a substrate. For far western blot analysis the blot was incubated with purified Nictaba (1 µg/ml, diluted in Tris-HCl pH 7.6) for one hour prior to incubation with the primary antibody against Nictaba, the secondary antibody and the detection buffer. All washes and incubations were conducted at room temperature with gentle shaking. The N-glycans of the purified Orysata (16 µg) were released using the on-membrane deglycosylation method as described earlier (Laroy et al., 2006). Briefly, the sample was incubated for 1 h at 50°C in denaturing buffer (360 mM Tris-HCl, pH 8.6 containing 8 M urea and 3.2 mM EDTA) and subsequently loaded on a 96-well Multiscreen-ImmobilonP plate containing a PVDF membrane (Millipore). Then, the bound proteins were reduced and carboxymethylated using dithiothreitol and iodeacetic acid, respectively. Next, the N-glycans were released using PNGaseF (in the negative control we omitted the enzyme). After labeling the N-glycans with 8-aminopyrene-1,3,6-trisulphonic acid, the excess of label was removed by size-exclusion chromatography using Sephadex G-10. The samples were finally reconstituted in 10 µl of ultrapure water and 10 µl of a 1:10 dilution was analyzed by capillary electrophoresis on an ABI 3130 DNA sequencer (Applied Biosytems, CA) as described (Laroy et al., 2006). To identify the structures, exoglycosidase digests were performed overnight at 37°C by adding 66 ng of *Trichoderma reseii* α-1,2-mannosidase (Maras et al., 2000) or 20 mU of jack bean α-mannosidase (Sigma) to 1.5 µl of sample in a total reaction volume of 3 µl containing 5 mM NH_4Ac, pH 5.

2A.4 Results

2A.4.1 Orysata is located in the cytoplasmic/nuclear compartment

Analysis of the amino acid sequence of Orysata (Genbank accession number CB632549) using the SignalP 3.0 tool (http://www.cbs.dtu.dk/services/SignalP) indicated the absence of a signal peptide, suggesting that the corresponding rice protein is synthesized on free polysomes. Furthermore the PSORT program (http://psort.nibb.ac.jp) predicted a subcellular localization of Orysata in the cytoplasmic compartment of the plant cell. The localization of

Orysata was corroborated by expression of an EGFP-fusion construct for the lectin in tobacco cells. Therefore the lectin sequence was fused in-frame to the C-terminus of EGFP and the fusion protein was transiently expressed in tobacco Bright Yellow-2 (BY-2) cells. Confocal microscopy of EGFP-Orysata at different time points after particle bombardment revealed that the rice lectin is located in the nucleus and the cytoplasm of the plant cell. No fluorescence emission was seen in the nucleolus or the vacuole. A very similar distribution pattern was observed at different time points after transformation and fluorescence was detectable until approximately 80 h after transformation (Fig. 2A.1).

A construct for the native 27 kDa EGFP under the control of the 35S promoter was used as a control. Expression of this protein in tobacco cells yielded an even distribution of the fluorescence pattern over the cytoplasm and the nucleoplasm, including the nucleolus (Fig. 2A.1).
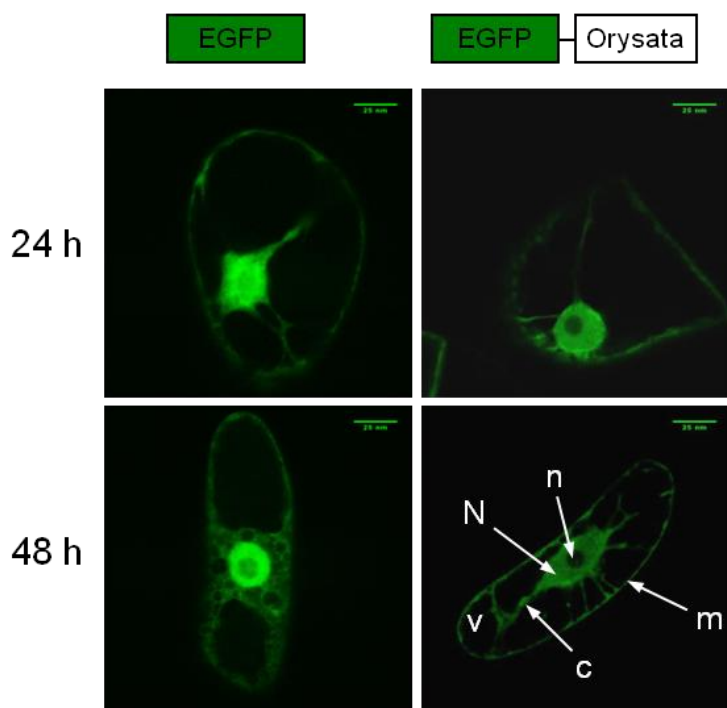

Fig. 2A.1 Confocal images collected from living, transiently transformed tobacco BY-2 cells expressing free EGFP and EGFP-Orysata. Expression of EGFP-Orysata or EGFP was analyzed at different time points after transformation. Scale bars represent 25 nm. Cell compartments: n, nucleolus; N, nucleus; m, cell membrane; c, cytoplasm; v, vacuole

2A.4.2 Purification and characterization of recombinant Orysata expressed in *P. pastoris*

Cloning of the coding sequence of Orysata into the *E. coli/P. pastoris* shuttle vector pPICZαB yielded a fusion construct whereby the Orysata sequence was linked to a C-myc epitope and a C-terminal histidine tag (Fig. 2A.2). The fusion protein was successfully expressed in the *Pichia* strain X-33. Because of the presence of the α–mating sequence from *Sacharomyces cerevisiae* at the N-terminus of the construct, the recombinant Orysata was secreted into the medium. Transformed *Pichia* colonies that yielded a positive result after analysis of the total protein by SDS-PAGE and subsequent Western blot analysis were grown in one-liter cultures. Afterwards the recombinant Orysata was purified from the medium using a combination of ion exchange chromatography, metal affinity chromatography on a Ni–Sepharose column and affinity chromatography on a mannose-Sepharose 4B column. Starting from a one-liter culture approximately 12 mg of recombinant protein was obtained.

SDS-PAGE analysis of the purified Orysata from *Pichia* revealed two bands of approximately 18.5 and 23 kDa (Fig. 2A.3A). A very similar result was obtained after Western blot analysis and detection of the recombinant proteins using a monoclonal antibody directed against the polyhistidine tag (Fig. 2A.3B). The deduced molecular mass of the lower band is in good agreement with the calculated molecular mass of Orysata fused to the c-myc epitope and the polyhistidine tag (18.46 kDa).

a

```
 1
M RFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPFSNSTN
                          86 87  89   92            100
NGLLFINTTIASIAAKEEGVSLEKRE A  EA  AAM TLVKIGLW  GGNGGSAQDISVP
PKKLLGVTIYSSDAIRSIAFNYIGVDGQEYAIGPWGGGEGTSTEIKLGSSEHIKEISGTHG
PVYDLADIVTYLKIVTSANNTYEAGVPNGKEFSIPLQDSGHVVGFFGRSGTLIDAIGIYVH
  236             254        259
P  ILEQKLISEEDLNSAVDH  HHHHH  *
```

b

```
MornigaM  MAGTSTNTQTTGTSQTVEVGLWGGPGGNAWDD-GSYTGIREINLSHG----DAIGAFSVI 55
Calespa   ----------MAVPMDTISGPWGNNGGNFWSF-RPVNKINQIVISYGGGGNNPIALTFSS 49
Orysata   -------------MTLVKIGLWGGNGGSAQDISVPPKKLLGVTIYSS---DAIRSIAFNY 44


MornigaM  YDLNGQPFTGPTHPGNEPSFKTVKITLDFPNEFLVSVSGYTGVLPRLATGKDVIRSLTFK 115
Calespa   TKADGSKDTITVGGGGPDSITGTEMVNIGTDEYLTGISGTFGIY----LDNNVLRSITFT 105
Orysata   IGVDGQEYAIGPWGGGEGTSTEIKLG---SSEHIKEISGTHGPVY---DLADIVTYLKIV 98


MornigaM  TNKKTYGPYGKEEGTPFSLP-IENGLIVGFKGRSGFVVDAIGVHLSL 161
Calespa   TNLKAHGPYGQKVGTPFSSANVVGNEIVGFLGRSGYYVDAIGTYNRH 152
Orysata   TSANNTYEAGVPNGKEFSIPLQDSGHVVGFFGRSGTLIDAIGIYVHP 145
```

Fig. 2A.2 (a) Sequence of recombinant Orysata expressed in *Pichia*, preceded by an N-terminal signal peptide (residues 1 to 89) necessary for secretion and a C-terminal tag containing a c-myc epitope and a (His)6 tag (residues 254 to 259). The cleavage sites for the signal peptide are indicated (Kex2 protease site at position 86 and Ste 13 protease sites at positions 87 and 89). The N-terminal sequence of recombinant Orysata determined by Edman degradation is underlined. The putative N-glycosylation site is shown in bold. (b) Sequence alignment for the three mannose-binding JRLs from *Oryza sativa*, *Calystegia sepium* and *Morus nigra*. Identical residues are shown in white with a black background and similar residues are boxed. The amino acid residues forming the monosaccharide-binding site are shown in red

N-terminal sequence analysis of both polypeptides yielded an identical sequence EAEAAAMTLVKIGLW. Since the six N-terminal amino acid residues in this sequence correspond to the yeast α-mating sequence it can be concluded that part of the signal peptide sequence was not cleaved properly (Fig. 2A.2). Detailed analysis of the amino acid sequence for Orysata revealed the presence of a putative glycosylation site NNT (Fig. 2A.2). Far western blot analysis whereby the blotted proteins were incubated with the N-glycan binding lectin Nictaba (Lannoo et al., 2007a) revealed interaction of Nictaba with the Orysata polypeptide of appr. 23 kDa, suggesting that this polypeptide is glycosylated (Fig. 2A.3c).

Indeed, only one polypeptide band of 18.5 kDa remains after removing the N-glycans of Orysata using peptide N-glycosidase F (PNGaseF) treatment (Fig. 2A.3d). Subsequent N-glycan analysis (Fig. 2A.4) revealed that the carbohydrate structures are high-mannose (Man9-11) glycans which are typically produced by wild-type *P. pastoris* (Jacobs et al., 2009). Molecular modeling of the mature Orysata sequence with an N-glycan at the position of the putative N-glycosylation site revealed that the glycan is located at the opposite side of the carbohydrate-binding site and hence is unlikely to interfere with the carbohydrate-binding properties of the lectin (Results not shown).
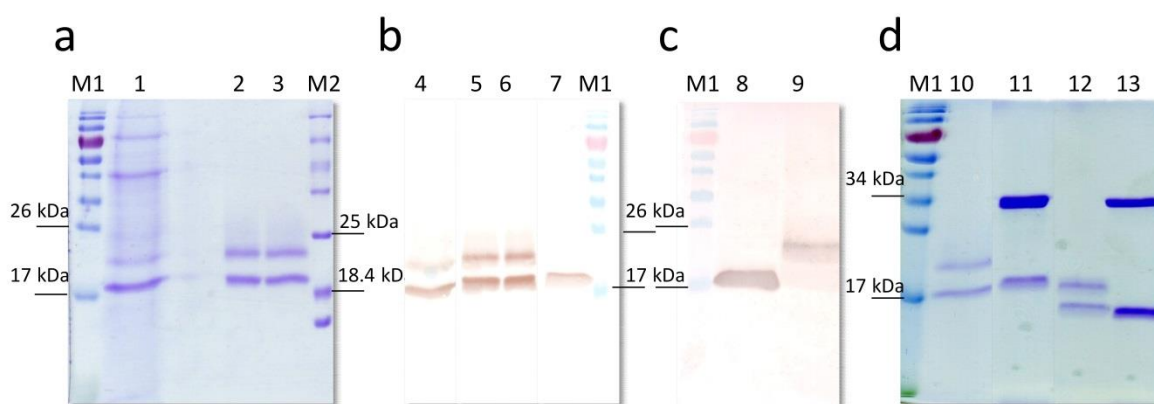
Fig. 2A.3 Crude protein extract from the medium of *Pichia* cell culture and purified Orysata were analyzed by SDS-PAGE (a), Western blot analysis with a monoclonal anti-His antibody (b), Far western blot analysis using Nictaba (1 µg/ml) (c) and PNGase F treatment (d). Samples are loaded as follows: lane M1: protein ladder (increasing MW: 10, 17, 26, 34, 43, 55, 72, 95, 130, 170 kDa), lane M2: unstained protein ladder (increasing MW: 14.4, 18.4, 25, 35, 45, 66.2, 116 kDa) (Fermentas, St. Leon-Rot, Germany), lanes 1 and 4: crude extract from *Pichia* cells expressing Orysata (15 µg), lanes 2 and 5: purified recombinant Orysata (2.5 µg) analyzed in the presence of mercaptoethanol and lanes 3 and 6: purified recombinant Orysata (2.5 µg) analyzed in the absence of mercaptoethanol. Lanes 7 and 8: positive controls (Nictaba), lane 9: recombinant Orysata (2.5 µg), lane 10: pure Orysata (2.5 µg), lane 11: pure Orysata (2.5 µg) digested with PNGase F (3.8 IUB mU), lane 12: positive control RNase B (2.5 µg) and lane 13: RNase B (2.5 µg) digested with PNGase F (3.8 IUB mU)
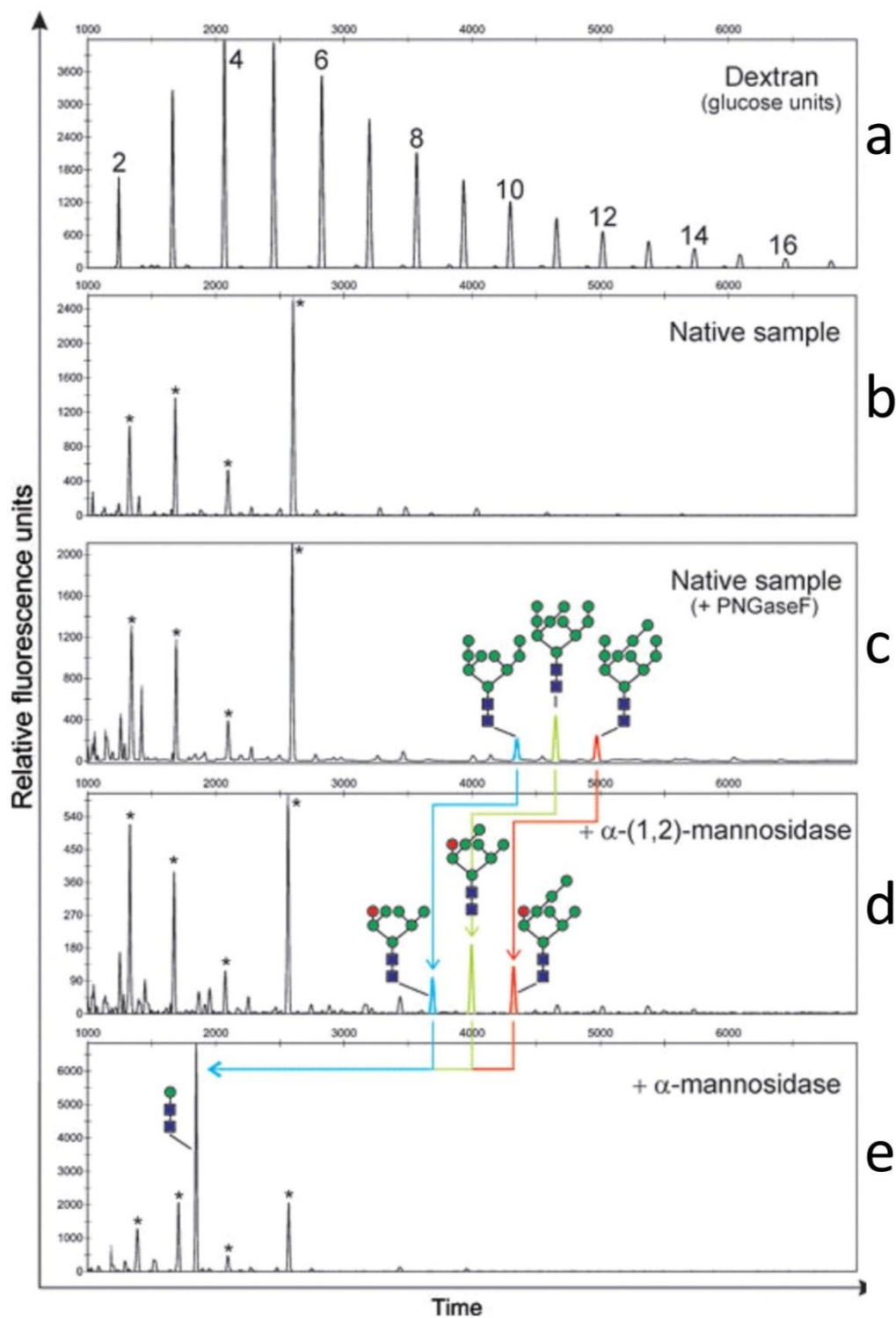
Fig. 2A.4 Identification of the N-glycans present on recombinant Orysata. N-glycans were released using PNGaseF (c) and to identify aspecific peaks (*) we also omitted the enzyme as a negative control (b). Alpha-1,2-mannosidase (d) and a broad-specific α-mannosidase (e) were added to the PNGaseF treated Orysata to identify the N-glycan structures. The result of a malto-dextrose reference is also given (a). Sugar code used: green circles indicate mannose residues, red circles are α-1,2-mannoses that cannot be cleaved by the α-(1,2)-mannosidase due to steric hindrance. Blue squares indicate N-acetylglucosamine residues and yellow circles indicate galactose residues as suggested by the Consortium for Functional Glycomics

2A.4.3 Agglutination activity and carbohydrate-binding properties of recombinant Orysata

To study the biological activity of the recombinant lectin expressed in *Pichia*, the recombinant Orysata was tested for agglutination activity towards rabbit erythrocytes. Agglutination was observed after adding the red blood cells to the purified lectin, the minimal concentration of lectin necessary to obtain agglutination activity being 5 μg/ml whereas it was 0.12 μg/ml for the native Orysata (Zhang et al., 2000). Preliminary carbohydrate inhibition assays revealed that the agglutination activity of the recombinant Orysata was similar to that of the native lectin in that the agglutination of rabbit erythrocytes was inhibited by mannose, methyl α-mannopyranoside and trehalose (Table 2A.1). Several glycoproteins also inhibited the agglutination activity of recombinant Orysata, though at higher concentration than required for inhibition of the native lectin.

More detailed carbohydrate-binding studies were performed using a screening of the lectin on a glycan array (Table 2A.2). The carbohydrate-binding properties of recombinant Orysata were investigated on glycan array v4.2, and compared to the sugar binding specificities of two other mannose-binding JRLs from *Calystegia sepium* and *Morus nigra*, further referred to as Calsepa and Morniga M, respectively (Fig. 2A.2b). On a first sight all three JRLs show similar interaction patterns with the glycan array (Fig. 2A.5). All lectins react with both high-mannose and complex N-glycans. However, more detailed analyses of the glycan array data show that Orysata and Morniga M show a higher reactivity towards high-mannose N-glycans, compared to Calsepa, which interacts primarily with galactosylated and sialylated biantennary complex N-glycans.

Table 2A.1 Comparison of the carbohydrate-binding specificities of native and recombinant Orysata. IC50 is the inhibitory concentration required to give a 50% inhibition of the agglutination of trypsin-treated rabbit erythrocytes at a lectin concentration of 12 µg/ml. The results for native Orysata are taken from (Zhang et al., 2000)

	IC_{50}	
	Native Orysata	Recombinant Orysata
Sugar		
Mannose	12 mM	50 mM
Trehalose	12 mM	25 mM
Methyl α-mannopyranoside	12 mM	25 mM
Glycoprotein		
Thyroglobulin	2 µg/ml	60 µg/ml
Ovomucoid	8 µg/ml	250 µg/ml
Asialomucin	250 µg/ml	500 µg/ml

Table 2A.2 Comparative analysis of glycan array results for Orysata, Morniga M and Calsepa. The glycan with the highest relative fluorescence unit (RFU) is assigned a value of 100. The rank is the percentile ranking. [a]: no reactivity

Glycan no	Structure	Orysata 25 µg/ml		Morniga M 50 µg/ml		Calsepa 50 µg/ml	
		RFU	Rank	RFU	Rank	RFU	Rank
360	Galα1-3Galβ1-4GlcNAcβ1-2Manα1-3(Galα1-3Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp20	42939	100	29317	76	18912	86
212	Manα1-6(Manα1-3)Manα1-6(Manα1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	41305	96	31139	81	6814	31
342	Manα1-3(Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAc-Sp12	34647	81	33653	87	10507	48
321	Galβ1-3GlcNAcβ1-2Manα1-3(Galβ1-3GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp19	34083	79	28240	73	12119	55
56	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3(Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp13	32258	75	33609	87	19389	88
361	Manα1-3(Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	31759	74	35422	92	11095	51
305	GlcNAcβ1-2Manα1-3(Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	30801	72	30973	80	7586	35
399	Galα1-4Galβ1-3GlcNAcβ1-2Manα1-3(Galα1-4Galβ1-3GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp19	29008	68	25848	67	5930	27
358	Fucα1-2Galβ1-4GlcNAcβ1-2Manα1-3(Fucα1-2Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp20	28743	67	19812	51	8588	39
316	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3(Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	28510	66	33022	86	14593	67
51	GlcNAcβ1-2Manα1-3(GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	27612	64	29277	76	6775	31
346	Galβ1-4GlcNAcβ1-2Manα1-3Manβ1-4GlcNAcβ1-4GlcNAc-Sp12	27579	64	37958	98	13309	61
458	Galβ1-4GlcNAcβ1-6(Galβ1-4GlcNAcβ1-2)Manα1-6(Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp19	27178	63	30338	79	11613	53
53	Galβ1-4GlcNAcβ1-2Manα1-3(Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	26984	63	31648	82	13724	63
393	Galβ1-4GlcNAcβ1-2Manα1-3(GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAc-Sp12	26515	62	24029	62	11719	53
52	GlcNAcβ1-2Manα1-3(GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp13	26286	61	38115	99	15111	69
345	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3Manβ1-4GlcNAcβ1-4GlcNAc-Sp12	25287	59	33568	87	18242	83

323	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3(Neu5Acα2-3Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	25059	58	32351	84	15692	72
49	Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	24991	58	38600	100	12609	58
343	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAc-Sp12	24979	58	29082	75	12118	55
317	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3(GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	24343	57	24806	64	12033	55
418	GlcNAcβ1-2Manα1-3(GlcNAcβ1-2(GlcNAcβ1-6)Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp19	23801	55	23280	60	a	a
425	Galβ1-3GlcNAcβ1-2Manα1-3(Galβ1-3GlcNAcβ1-2(Galβ1-3GlcNAcβ1-6)Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp19	23714	55	16526	43	5874	27
315	Neu5Acα2-3Galβ1-4GlcNAcβ1-2Manα1-3(Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	23432	55	24349	63	1325	6
368	Galα1-3(Fucα1-2)Galβ1-4GlcNAcβ1-2Manα1-3(Galα1-3(Fucα1-2)Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp20	23094	54	30841	80	5745	26
50	Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp13	21861	51	34978	91	21918	100
213	Manα1-6(Manα1-3)Manα1-6(Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	21621	50	26316	68	7179	33
477	Manα1-6(Manα1-3)Manβ1-4GlcNAcβ1-4(Fucα1-6)GlcNAcβ-Sp19	21471	50	28412	74	2475	11
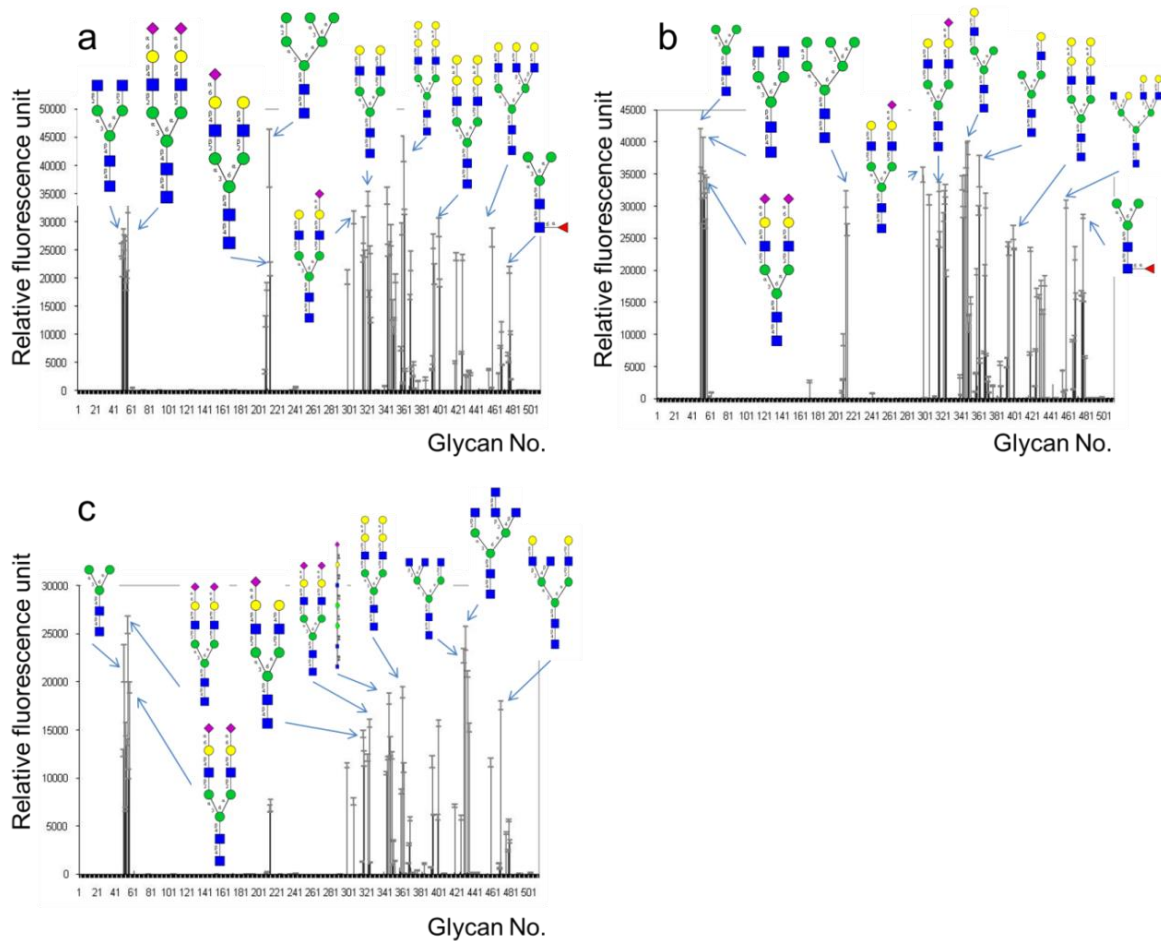
Fig. 2A.5 Comparative analysis of binding of recombinant Orysata, Morniga M and Calsepa on the glycan array. Panels a-c show interaction of recombinant Orysata (25 µg/ml), Morniga M (50 µg/ml) and Calsepa (50 µg/ml), respectively. The complete primary data set for each protein is available on the website of the Consortium for Functional Glycomics (www.functionalglycomics.org). Sugar code used: Green circles indicate mannose residues, yellow circles indicate galactose residues, blue squares indicate N-acetylglucosamine residues, purple diamonds indicate sialic acid residues and red triangles indicate fucose

2A.4.4 Antiviral activity of recombinant Orysata, compared to Calsepa and Morniga M

The three JRLs were evaluated for their antiviral activity against HIV-1(IIIB) and HIV-2(ROD) in CEM cell cultures (Table 2A.3). The $\alpha1,3/\alpha1,6$-mannose-specific *Hippeastrum* hybrid agglutinin (HHA) was included as a control since it showed significant activity against HIV (Balzarini 2006) and other viruses. Orysata efficiently

suppressed HIV infection at a 50% effective concentration of 1.7 to 5.6 μg/ml, corresponding to a concentration which is ~ 10-fold higher than required for HHA. In contrast, Calsepa was marginally inhibitory against HIV-1 (EC$_{50}$: ≥100 μg/ml). Morniga M could not be evaluated at compound concentrations higher than 4 μg/ml due to cytotoxicity in the cell cultures at a concentration of ≥ 20 μg/ml.

The lectins have also been investigated for their inhibitory activity against syncytia formation between persistently HIV-1(III$_B$)-infected HUT-78/HIV-1 cells and uninfected Sup T1 cells. The three lectins prevented giant cell formation at 18 to 38 μg/ml by 50%. This concentration proved to be 10- to 20-fold higher than required for HHA under similar experimental conditions (Table 2A.3). Interestingly, when exposed to RSV-infected HeLa cell cultures, Orysata and Calsepa (EC$_{50}$: 1.6-2.1 μg/ml) but not Morniga M and HHA (EC$_{50}$: ≥ 20 μg/ml) efficiently inhibited viral infection.

Table 2A.3 Inhibitory activity of the lectins against HIV-1 and HIV-2 in human T-lymphocyte (CEM) cell cultures and against syncytium formation between HUT-78 ∕ HIV-1 and Sup T1 cells. EC50 is the effective concentration or the concentration required to protect CEM cells against the cytopathogenicity of HIV by 50% or to prevent syncytia formation in co-cultures of persistently HIV-1-infected HUT-78 cells and uninfected Sup T1 lymphocyte cells

Compound	EC$_{50}$ (μg/ml)		
	HIV-1(III$_B$)	HIV-2(ROD)	HUT-78/HIV-1 + Sup T1
Orysata	1.7 ± 0.14	5.6 ± 3.7	38 ± 6.7
Calsepa	≥100	>100	26 ± 10
MornigaM	>4	>4	18± 4.0
HHA	0.17 ± 0.021	0.49 ± 0.47	1.7 ± 0.8

2A.4.5 Molecular modeling of carbohydrate-binding sites

Although the three Man-specific JRLs Orysata, Morniga M and Calsepa accommodate both Man and MeMan in a very similar way (Fig. 2A.6 panels a, d and g), they display a rather different affinity towards more complex saccharides as shown from the reported glycan array experiments (Table 2A.2) and the anti-HIV activity (Table 2A.3). In this respect, Orysata resembles Morniga M, since both lectins predominantly interact with high-mannose N-glycans, whereas Calsepa exhibits a higher affinity for

complex N-glycans. These discrepancies most probably depend on differences in the shape and size of their carbohydrate-binding cavities. The carbohydrate-binding cavity of Man-specific JRLs (Calsepa, Morniga M, Orysata) consists of three loops L1, L2 and L3 containing two conserved Gly (L1) and Asp (L3) residues and two other variable residues (Thr134 and Leu135 in Orysata, Phe150 and Val151 in Calsepa, Tyr141 and Tyr142 in Morniga M) that also belong to loop L3 (Fig. 2A6 panels c, f and i). Depending on the bulkiness of loop L2, the carbohydrate-binding cavity of the lectins exhibits considerable differences in shape and size (Raval et al., 2004; Van Damme et al., 2007a). Orysata and Calsepa exhibit a crescent-shaped binding-cavity largely open at both extremities, and thus can accommodate extended oligosaccharide chains (Fig. 2A.6 panels b and e). The binding site of Morniga M possesses a totally different shape due to the bulkiness of loop L2 which closes up the cavity at one extremity and considerably decreases its size (Fig. 2A.6e). However, the carbohydrate-binding cavity of Morniga M remains largely open at the opposite extremity which should allow α3-O-linked saccharides to interact with the lectin but prevent the correct accommodation of α1-O-linked saccharides.
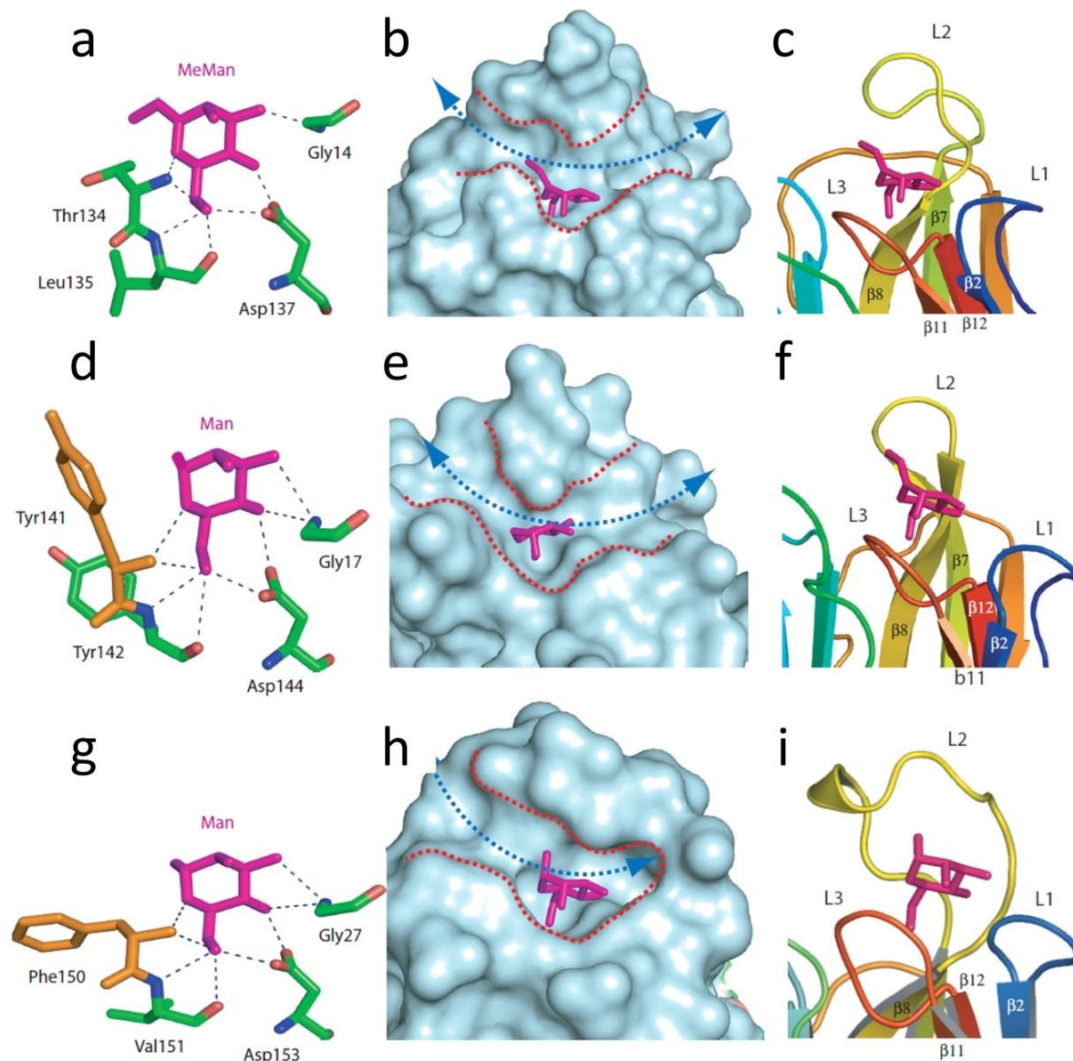
Fig. 2A.6 Molecular modeling of the carbohydrate-binding sites of Orysata, Calsepa and Morniga M. A, d, g: Network of hydrogen bonds anchoring Man to the saccharide binding sites of Orysata (a), Calsepa (d) and Morniga M (g). Hydrogen bonds are represented in blue dotted lines. Aromatic residues that create a stacking interaction with the sugar are colored orange. B, e, h: Topography of the saccharide binding cavity at the surface of the Orysata (b), Calsepa (e) and Morniga M (h) protomers. Cavities are delineated by red dotted lines and the curved blue arrows indicate the overall orientation of the cavities. C, f, i: Ribbon diagrams at the top of the Man-binding lectins showing the overall topography of the carbohydrate-binding sites of Orysata (c), Calsepa (f) and Morniga M (i). L1, L2 and L3 correspond to the loops forming the carbohydrate-binding cavity of the lectins. Strands of β-sheet participating to the binding cavities are numbered. Molecular modeling was performed by Prof Pierre Rougé, Toulouse, France

2A.5 Discussion

This paper describes the characterization of Orysata, a mannose-binding JRL from rice (*Oryza sativa*) expressed in *P. pastoris*. Recombinant Orysata was successfully expressed in *Pichia* strain X-33 with the addition of a signal sequence for secretion of the recombinant protein into the medium. Approximately 12 mg of the recombinant lectin was purified from the medium of a one-liter culture (BMMY, pH 6) induced with methanol for 72 h. Compared to the yield reported for other recombinant lectins that were expressed extracellularly in *Pichia*, the amount of lectin obtained for Orysata is considered to be rather low. However, it should be mentioned that the yield obtained for the nucleocytoplasmic lectin from tobacco was even lower, being only 6 mg/l (Lannoo et al., 2007a). To our knowledge only one JRL has been previously expressed in *Pichia*. The galactose-binding lectin frutalin from breadfruit seeds was successfully expressed at 18-20 mg/l (Oliveira et al., 2008). Much higher yields of recombinant protein can be obtained when *Pichia* cultures are grown in a bioreactor under controlled conditions, as reported for the recombinant lectins from *Aleuria aurantia* (67 mg/l) (Amano et al., 2003), snowdrop (80 mg/l) (Baumgartner et al., 2003) and the bean lectin PHA-E (100 mg/l) (Baumgartner et al., 2002).

After purification, two molecular forms of the lectin were detected by SDS-PAGE and Western blot analysis. Edman degradation revealed them to have identical N-terminal sequences, suggesting that the higher molecular weight fraction might be glycosylated. Indeed a careful analysis of the amino acid sequence revealed one putative N-glycosylation site at position 102 of the mature Orysata sequence (NNT). Far Western blot analysis using Nictaba, a lectin with well-defined specificity towards high-mannose and complex N-glycans (Lannoo et al., 2006a), confirmed that the 23 kDa polypeptide for Orysata is glycosylated whereas the 18.5 kDa polypeptide is unglycosylated, indicating that the recombinant Orysata obtained from the *Pichia* culture is partially glycosylated. This result was further confirmed by PNGaseF treatment of the recombinant Orysata which resulted in a shift of the 23 kDa polypeptide to 18.5 kDa. In this respect it should be mentioned that the JRL frutalin was also partially glycosylated after secreted expression in *Pichia* with a very similar

size difference between the glycosylated and the non-glycosylated lectin polypeptides (Oliveira et al., 2008). Furthermore N-terminal sequence analysis of recombinant Orysata showed that the processing of the α-mating sequence was not fully completed. It has been reported before that cleavage of EA repeats by Ste13 protease is an inefficient process, but these repeats are necessary to enhance proper function of the Kex2 protease (Sreekrishna et al., 1997). In the case of Nictaba and frutalin incomplete processing of the signal peptide was also reported (Oliveira et al., 2008; Lannoo et al., 2007a). The uncleaved part of the α-mating sequence at the N-terminus as well as the histidine tag at the C-terminus of the recombinant lectin apparently do not influence the biological activity of Orysata, since the recombinant lectin reacted with carbohydrate structures and agglutinated red blood cells.

Molecular cloning and characterization of the lectin from rhizomes of *Calystegia sepium* (abbreviated as Calsepa) unambiguously showed that some JRLs show specificity towards mannose (Van Damme et al., 1996). Since then the family of JRLs is subdivided in two classes of lectins with preferential specificity towards galactose (as in the case of jacalin) and mannose (as in the case of Calsepa). In the last decade several so-called mannose-binding JRLs have been identified and characterized from different plant species (Van Damme et al., 2008). Structural analyses as well as detailed studies of the carbohydrate-binding properties have shown that both the galactose-binding and the mannose-binding JRLs are polyspecific lectins with a preference for galactose and mannose, respectively (Bourne et al., 2004; Rougé et al., 2003). Analysis of the carbohydrate-binding specificity of three mannose-binding JRLs on the glycan array revealed differences in their specificity. Clearly Orysata and Morniga M interact much better with high-mannose binding glycans compared to Calsepa. These results are in agreement with the analyses of the sugar binding specificity of Morniga M and Calsepa by frontal affinity chromatography where it was shown that although Morniga M and Calsepa both react with high-mannose structures (especially of Man2-6 type), Calsepa showed a much better interaction with complex N-glycans with bisecting GlcNAc (Nakamura-Tsuruta et al., 2008). Although the frontal affinity chromatography indicated that Morniga M and Calsepa did not react with tri- and tetra-antennary glycans, some interactions with these

glycan structures have been observed on the array. Molecular modeling studies suggest subtle differences in the carbohydrate-binding sites of JRLs. The shortening of the carbohydrate-binding cavity in Morniga M could account for the differences in specificity of the different Man-specific JRLs towards extended oligosaccharide chains, e.g. the α1-O-linked, α3-O-linked and α6-O-linked oligosaccharides.

Until now especially mannose-binding lectins belonging to the group of GNA-lectins such as snowdrop (GNA) and amaryllis (HHA) lectin have been shown to exhibit significant activity against HIV as well as some other viruses such as hepatitis C virus (Botos and Wlodawer 2005; Balzarini 2006; Bertaux et al., 2007). Since very little is known with respect to the antiviral activity of JRLs the anti-HIV activity of three mannose-binding JRLs was tested and compared. Detailed analysis showed that Orysata has potent anti-HIV and anti-RSV activity. Only recently the mannose-binding JRL isolated from the fruit of banana *Musa acuminata* BanLec was also reported to exhibit potent anti-HIV activity (Swanson et al., 2010). It was shown that HHA and BanLec interact with gp120 and can inhibit HIV replication. It is intriguing, however, to notice that the α1,3/α1,6-mannose-specific HHA is 10- to 20-fold more inhibitory to HIV but more than 10-fold less inhibitory to RSV than Orysata. This may point to subtle differences in carbohydrate recognition of both lectins, and is in agreement with the modeling and glycan arrays suggesting that Orysata also recognizes complex-type glycans in addition to high-mannose type glycans. Although the nature of the glycans on the envelope of RSV are not unambiguously determined, they most likely predominantly consist of complex-type glycans since mannose-specific lectins such as GNA and HHA have never found to be endowed with significant anti-RSV activity in cell culture. Taken all data together, the lectin may qualify as a microbicide candidate agent since it not only blocks T-cell infection by cell-free HIV but it also prevents virus transmission (syncytia formation) between HIV-infected cells and uninfected cells. However, additional studies are required to further explore the microbicide potential of Orysata.

Expression of the low abundant rice lectin Orysata in *Pichia* allowed comparing its biological activity to that of other JRLs such as Calsepa and MornigaM which are

expressed in high amounts in plants. Glycan array analyses confirmed earlier reports on the polyspecificity of Calsepa and MornigaM (Bourne et al., 2004; Rougé et al., 2003). Data from molecular modelling suggest that subtle differences in the carbohydrate-binding site of the different JRLs could explain the different specificities and antiviral activities of the JRLs under study.

Chapter 2

Characterization of the lectins under study expressed in *Pichia pastoris*

Chapter 2B

Expression analysis of a type S2 EUL-related lectin from rice in *Pichia pastoris*

Redrafted from:

Al Atalah B, Rougé P, Smith DF, Proost P, Lasanajak Y, Van Damme EJM (2012) Expression analysis of a type S2 EUL-related lectin from rice in *Pichia pastoris.* Glycoconjugate J 29: 467–479

2B.1 Abstract

Rice (*Oryza sativa*) expresses different putative carbohydrate-binding proteins belonging to the class of lectins containing an *Euonymus* lectin (EUL)-related domain, one of them being OrysaEULS2. The OrysaEULS2 sequence consists of a 56 AA N-terminal domain followed by the EUL sequence. In this paper the original EULS2 sequence and some mutant forms have been expressed in *Pichia pastoris*. Subsequently the recombinant proteins were purified and their carbohydrate binding properties determined. Analysis of the original protein on the glycan array revealed interaction with mannose containing structures and to a lesser extent with glycans containing lactosamine related structures. It was shown that mutation of tryptophan residue 134 into leucine resulted in an almost complete loss of carbohydrate binding activity of OrysaEULS2. Our results show that the EUL domain in OrysaEULS2 interacts with glycan structures, and hence can be considered as a lectin. However, the binding of the protein with the array is much weaker than that of other EUL-related lectins. Furthermore, our results indicate that gene divergence within the family of EUL-related lectins lead to changes in carbohydrate binding specificity.

2B.2 Introduction

In 2008 sequencing of the cDNA clone encoding the EEA from spindle tree led to the discovery of a new lectin motif in plants (Fouquaert et al., 2008). Since then, the *Euonymus europaeus* lectin (EUL) domain is considered as the prototype for a new lectin family (Fouquaert et al., 2009a). Screening of the available genome and transcriptome data revealed the broad distribution of the EUL lectin domain in the plant kingdom. A classification system for proteins containing one or more EUL-related lectin domains was proposed based on the architecture of the EUL domain in the different protein sequences (Fouquaert et al., 2009a). According to this classification, some EUL sequences comprise a single EUL domain preceded by variable and unrelated N-terminal domain, whereas other EUL proteins are composed of two EUL domains arrayed in tandem and separated by a linker sequence. In rice, four different types of EUL proteins have been identified: 1) a single-domain protein with a medium long unrelated N-terminal sequence (type S2, referred to as OrysaEULS2), 2) a single-domain EUL protein with a long unrelated N-terminal sequence (type S3, OrysaEULS3), 3) two two-domain proteins with a short linker (type D1, OrysaEULD1A, and OrysaEULD1B), and 4) one two-domain protein with a long linker sequence (type D2, OrysaEULD2) (Fouquaert et al., 2009a).

The OrysaEULS2 protein, originally named as OSR40g3, was first reported in the shoots of rice seedlings subjected to salt stress and ABA treatment (Moons et al., 1997a). The protein is encoded by a gene annotated under two accessions; Os07g0684000 (National Center for Biotechnology Information [NCBI] annotation) and Os07g48500 (The Institute for Genomic Research annotation). OrysaEULS2 is annotated as a 'ricin B-related lectin domain containing protein' because it has two QXW repeats (which is typical for ricin-B domain), but according to the BLASTp search there is no significant sequence similarity with proteins comprising a ricin-B domain and thus it is questionable whether it can be considered as a ricin-B family member (Fouquaert et al., 2008). According to the classification of plant lectins proposed by Fouquaert et al. (2009a) this protein belongs to the group of chimeric EUL proteins and consists of an EUL domain preceded by a 56 unrelated AA sequence, and can be grouped as a type S2 EUL-related lectin. The purification of

OrysaEULS2 is cumbersome because the protein is expressed at very low levels even after the plant was subjected to stress. In an attempt to gain more information on the carbohydrate binding properties of OrysaEULS2 the presumed carbohydrate-binding domain of this protein was recombinantly expressed in the heterologous expression system *Pichia pastoris*. In the last decade this expression system has become a pioneer biological tool to produce proteins of interest (Cereghino and Cregg 2000; Macauley-Patrick et al., 2005; Daly and Hearn 2005). The expression of the recombinant proteins is controlled by the AOX1 promoter and tightly induced to high levels by methanol (Elias et al., 1985; Hartner and Glieder 2006). Several nucleocytoplasmic lectins from tobacco (Nictaba), maize (GNAmaize), rice (Orysata) and *Arabidopsis thaliana* (ArathEULS3) have been expressed successfully in *P. pastoris* (Lannoo et al., 2007a; Fouquaert et al., 2009b; Al Atalah et al., 2011 see chapter 2A; Van Hove et al., 2011).

In this paper we describe the expression and purification of the EUL domain from OrysaEULS2, further referred to as EULS2, in *P. pastoris*. In an attempt to unravel which amino acids are required for the carbohydrate binding activity of EULS2 a three-dimensional model was made for EULS2, and compared to the models for EEA and the EUL domain within ArathEULS3. The molecular structure, biological activity, putative carbohydrate binding site and specificity of the EULS2 domain and its mutants have been investigated. These data will allow us to get better insights into the physiological role and importance of the OrysaEULS2 protein in carbohydrate-mediated stress responses.

2B.3 Materials and methods

2B.3.1 Hydrophobic cluster analysis and molecular modeling

Hydrophobic Cluster Analysis (HCA) (Treiber et al., 2008) was performed to assess the conserved secondary structural features (essentially strands of β-sheet) along the amino acid sequences of the EEA and the EUL domains of *Arabidopsis* (ArathEULS3) and rice (OrysaEULS2). HCA plots were generated using the HCA server

(http://mobyle.rpbs.univ-paris-diderot.fr). Molecular modeling was performed with the YASARA Structure program (Krieger et al., 2002) running on a 2.53 GHz Intel core duo Macintosh computer. All protein models were built from the X-ray coordinates of the holotoxin from Bacillus sphaericus (RCSB PDB code 2VSE) (Gaboriaud et al., 1987), which was used as a template. The toxin shares acceptable percentages of identity (~ 20%) and homology (~ 50%) with all EUL proteins and allowed to build an accurate three-dimensional model for the three EUL lectins. PROCHECK (Laskowski et al., 1993) was used to assess the geometric quality of the three-dimensional models. As an example, about 80% of the residues of the modeled EEA protein were correctly assigned on the best allowed regions of the Ramachandran plot, the remaining residues being located in the generously allowed regions of the plot except for Phe32 and Lys75 (in EEA) which occur in the non-allowed region (result not shown). Cartoons were drawn with PyMOL (http://www.pymol.org) and YASARA. Electrostatic potentials were calculated (Nicholls et al., 1991) and displayed with GRASP using the parse3 parameters (Gilson and Honig 1987). The solvent probe radius used for molecular surfaces was 1.4 Å and a standard 2.0 Å-Stern layer was used to exclude ions from the molecular surface. The inner and outer dielectric constants applied to the protein and the solvent were respectively fixed at 4.0 and 80.0, and the calculations were performed keeping a salt concentration of 0.145 M.

2B.3.2 Expression of EULS2 and its mutant forms in *P. pastoris*

The EasySelect *Pichia* Expression Kit from Invitrogen was used (Invitrogen, Carlsbad, CA USA) to clone and express the EUL domain from OrysaEULS2 and its mutant forms. Therefore the coding sequence for the EUL domain (amino acids 57-204 of the full OrysaEULS2, Genbank accession number AK072989) was amplified from the pFLCI vector (obtained from the Rice Genome Resource Center, Ibaraki, Japan) by PCR using primers EVD 574 and EVD 514 (Table 2B.1). The mutant forms were created by using reverse primers containing several point mutations (Table 2B.1). Cloning and transformation procedure is described in sections 2A.3.2 and 2A.3.3.

Table 2B.1 PCR scheme and primer sequences. Nucleotides that have been mutated are shown in red

Construct	PCR/ Template	Direction	Primer name	Primer sequence 5' to 3'
Original S2d	PCR1/pFL CI vector	Forward	EVD 574	GGCGGAGAATTCACCATGGTGTACTGCCGG GCG AACCCGAA
		Reverse	EVD 514	CCCGCTTTCTAGAATGTAGTAGGGCTGGATC TTCCAGCGC
Mutant 1 W changed to L	PCR2/pFL CI vector	Forward	EVD 574	Above
		Reverse1	EVD 639	GTAGTAGGGCTGGATCTTCCAGCGCTGGTT GTCGCCCTCGCACAACTTCCA
	PCR3/PCR 2 product	Forward	EVD 574	Above
		Reverse 2	EVD 514	Above
Mutant 2 W changed into L and N into S	PCR4/PCR 2 product	Forward	EVD 574	Above
		Reverse1	EVD 640	GTAGTAGGGCTGGATCTTCCAGCGCTGGCT GTCGCCCTC
	PCR5/PCR 4 product	Forward	EVD 574	Above
		Reverse 2	EVD 514	Above
Mutant 3 W changed into L, N into S and Q into L	PCR6/PCR 4 product	Forward	EVD 574	Above
		Reverse	EVD 641	GTAGTAGGGCTGGATCTTCCAGCGCAGGCT GTCGCCCTC
	PCR7/PCR 6 product	Forward	EVD 574	Above
		Reverse 2	EVD 514	Above
AOX1 primers		Forward	EVD 21	GACTGGTTCCAATTGACAAGC
		Reverse	EVD 22	GCAAATGGCATTCTGACATCC

2B.3.3 Expression analysis of the recombinant proteins

Performed like explained in sections 2A.3.3 and 2A.3.4. Through the purification procedure, the mannose-Sepharose 4B column was not used. Instead thyroglobulin column was used to further purify the recombinant proteins.

2B3.4 Purification of the recombinant proteins

Purification of the native EULS2 and its mutant forms was achieved as described in the section 2A.3.4. Starting from a 1 l culture 4 to 6 mg of recombinant EULS2 was obtained.

2B.3.5 N-terminal sequence analysis

N-terminal sequencing was conducted as mentioned in the section 2A.3.5.

2B.3.6 Agglutination assay

The biological activity for the recombinant proteins was examined by performing an agglutination assay using trypsin-treated rabbit red blood cells (BioMérieux, Marcy l'Etoile, France). The procedure is explained in the section 2A.3.6.

2B.3.7 Carbohydrate inhibition test

To test the carbohydrate specificity of the recombinant proteins, a procedure explained in section 2A.3.7 was followed. The concentration required to inhibit the agglutination of trypsin-treated rabbit erythrocytes was estimated at a lectin concentration of 100 μg/ml.

2B.3.8 Glycan array screening

The microarrays are printed as described previously (Blixt et al., 2004) and version 5.0 with 611 glycan targets was used for the analyses reported here (https://www.functionalglycomics.org/static/consortium/resources/resourcecoreh8.shtml). The printed glycan array contains a library of natural and synthetic glycan sequences representing major glycan structures of glycoproteins and glycolipids. Recombinant EULS2 and its mutant forms containing a His tag were purified from *P. pastoris* and detected using a fluorescent-labeled anti-His monoclonal antibody

(Qiagen, Valencia, CA). Since the 488-labeled anti-His did not show binding (data not shown), detection of the immune complex with alexa488-labeled mouse anti-His was amplified using an Alexa633-labeled goat anti-mouse IgG. In addition to any amplification coming from the additional antibody binding the Alexa633 label is approximately 10 times more sensitive than the Alexa488 label. The lectin was diluted to desired concentrations in binding buffer (Tris-bufferred saline containing 10 mM $CaCl_2$, 10 mM $MgCl_2$, 1% BSA, 0.05% Tween 20, pH 7.5) and 70 µl of the lectin solution was applied to separate microarray slides. After 60 min incubation under a cover slip in a humidified chamber at room temperature, the cover slip was gently removed in a solution of Tris-buffered saline containing 0.05% Tween 20 and washed by gently dipping the slides 4 times in successive washes of Tris-buffered saline containing 0.05% Tween 20, and Tris-buffered saline. To detect bound lectin, the labeled anti-His antibody (70 µl at one µg/ml in binding buffer) was applied to the slide under a cover slip. After removal of the coverslip and gentle washing of the slide as described above, this process was repeated with Alexa633 labeled goat anti-mouse IgG (Invitrogen, Eugene, OR) and the slide was finally washed in deionized water and spun in a slide centrifuge for approximately 15 sec to dry. The slide was immediately scanned in a PerkinElmer ProScanArray MicroArray Scanner using an excitation wavelength of 633 nm and ImaGene software (BioDiscovery, Inc., El Segundo, CA) to quantify fluorescence. The data are reported as average Relative Fluorescence Units (RFU) of six replicates for each glycan presented on the array after removing the highest and lowest values.

2B.3.9 Analytical methods

All analytical methods used in this study are described in section 2A.3.11.

2B.4 Results

2B.4.1 Three-dimensional models for EEA and EUL domains in OrysaEULS2 and ArathEULS3

Alignment of the protein sequences encoding EEA, ArathEULS3 and OrysaEULS2 revealed 66% sequence similarity in their C-terminal EUL domain. Both ArathEULS3 and OrysaEULS2 contain an unrelated N-terminal domain which is absent from EEA (Fig. 2B.2b). Molecular modeling was performed for the EUL domains in all three proteins using the three-dimensional structure of the holotoxin from B. sphaericus (2VSE). In spite of the rather moderate percentages of identity (~ 20%) and similarity (~ 50%) between the three EUL domains and the holotoxin, the HCA plots revealed a closely related overall fold for all these proteins. Accordingly, the three-dimensional models built for all EUL proteins exhibited a very similar β-trefoil fold consisting of three bundles of β-sheet organized around a pseudo three-fold symmetry axis (Fig. 2B.1a) and are nicely superposable (Fig. 2B.1b). This β-trefoil structure is reminiscent to that found in ricin-B and other lectins from the ricin-B family (Van Damme et al., 2001). EUL domains in both EEA (Cys16) and the ArathEULS3 (Cys57) contain a single Cys residue, whereas EULS2 contains three Cys residues (Cys3, Cys100, and Cys135), respectively, that are too far from each other to create intra-chain disulphide bonds.
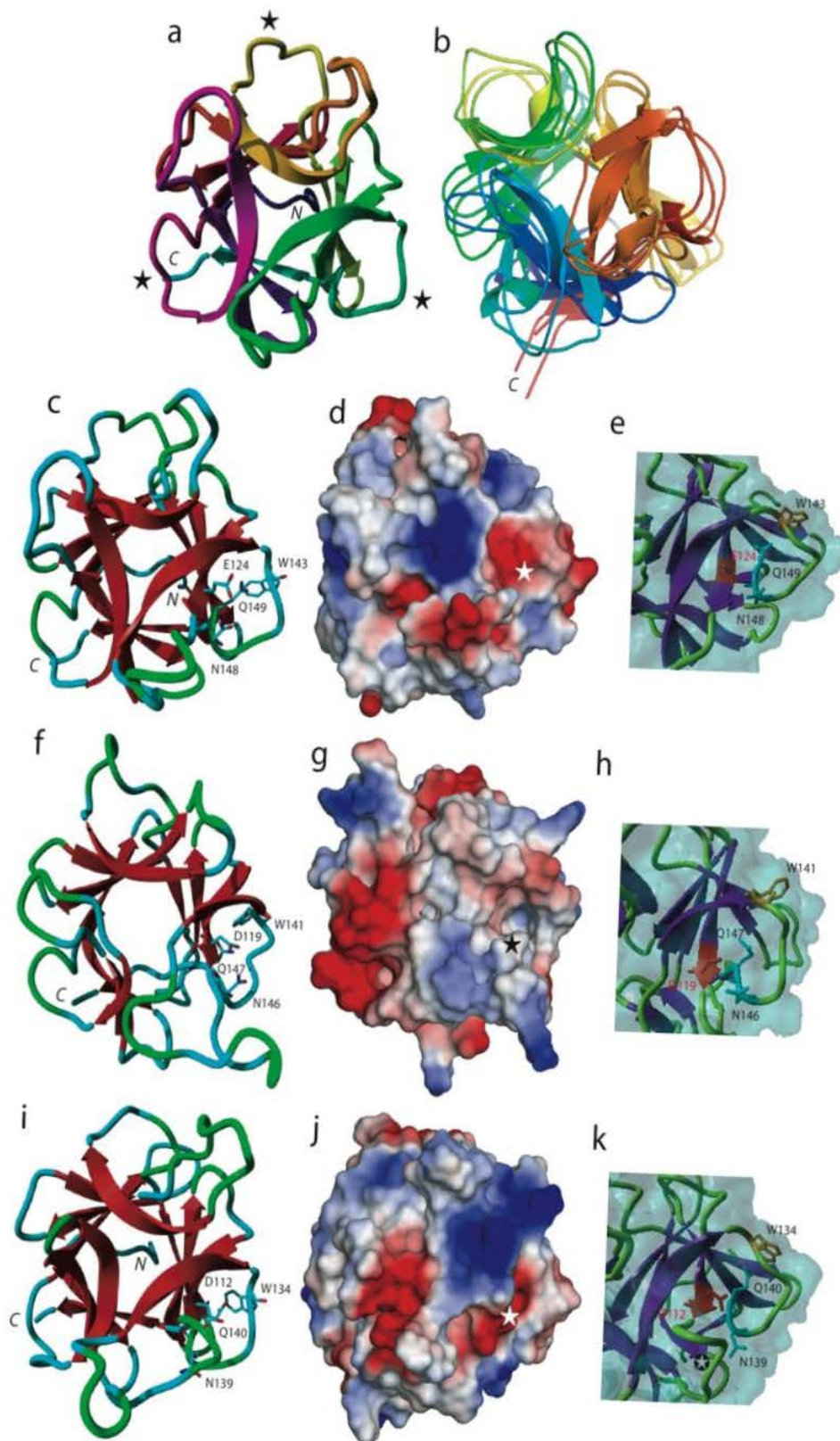
Fig. 2B.1 Legend on next page

Fig. 2B.1 a. Ribbon diagram of EEA showing the three bundles of β-sheet in different colors (orange, green, magenta) associated to the extended loops (indicated by stars) forming the three lobes of the β-trefoil. b. Superposition of the three models for EEA and the EUL domains of ArathEUL3 and OrysaEULS2 showing a similar three-dimensional fold. Panels c, f and i show ribbon diagrams of EEA (c), and EUL domains of ArathEUL3 (f) and OrysaEULS2 (i). The strands of β-sheet and the loops and coil regions are colored red, green and cyan, respectively. N and C indicate the N-terminal and C-terminal ends of the polypeptide chain. The amino acid residues predicted to form a carbohydrate-binding site in a loop located at the C-terminal end of the polypeptide chain are represented in cyan sticks and labelled (according to the sequence). Panels d, g and j show the mapping of the electrostatic potentials on the molecular surface of EEA (d), EUL domain of ArathEULS3 (g) and OrysaEULS2 (j). Electropositively and electronegatively charged regions are colored blue and red, respectively; neutral regions are colored white. The stars indicate the localization of the putative carbohydrate-binding sites. Panels e, h and k show an enlarged view of the carbohydrate-binding sites of EEA (e), EUL domains of ArathEUL3 (h) and OrysaEULS2 (k). The four amino acid residues forming the carbohydrate-binding sites are represented in sticks and labelled (according to the built three-dimensional models). The loop masking Asn139 residue in the EUL domain from OrysaEULS2 is indicated by a star. Molecular modeling was performed by Prof Pierre Rougé, Toulouse, France

2B.4.2 Prediction of putative carbohydrate binding site

Based on the comparison of the three EUL lectins and the C-terminal domain of some bacterial lectins of the ricin-B family and, especially the HA33/A protein from *C. botulinum* (PDB code 1YBI) (Arndt et al., 2005), a putative carbohydrate-binding site consisting of four well conserved residues (Glu124, Trp143, Asn148, Gln149 for EEA; Asp119, Trp141, Asn146, Gln147 for the EUL domain in ArathEULS3 from *Arabidopsis*, and Asp112, Trp134, Asn139, Gln140 for the OrysaEULS2 from rice) was predicted to occur at the C-terminal end of each of the EUL domains (Fig. 2B.1c,f,i). These extremely conserved carbohydrate-binding sites appear as a charged groove as shown from the mapping of the electrostatic potentials on the molecular surface of the lectins (Fig. 2B.1d,g,j). Another aromatic residue located in the vicinity of the putative carbohydrate-binding site, e.g. Tyr147 in EEA and Trp132, Trp139 in the EUL domains of ArathEULS2 and OrysaEULS3 respectively (Fig. 2B.2), could also participate in stacking interactions that often reinforce the binding of a simple sugar to the carbohydrate-binding site of plant lectins (Van Damme et al., 2007a). However, a careful examination of the exposure of the residues forming the carbohydrate-binding sites revealed some discrepancies among the three models, depending on the conformation of a loop located in the close vicinity of the site. In the EUL domain of OrysaEULS2, the loop extends on the Asn139 residue in such a

way that this residue, which plays a crucial role in the binding of saccharides by lectins of the ricin-B family, becomes almost completely buried (Fig. 2B.1k). In this respect, OrysaEULS2 readily differs from EEA and ArathEULS3, since in the latter lectins the corresponding Asn residue remains fully exposed (Fig. 2B.1e,h).

a

```
 1
M RFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPFSNSTNNGL
                       86 87  89    93
LFINTTIASIAAKEEGVSLEKRE_A_EA_EFTM_VYCRANPNYAMTARNGAVVLAPANPKDEY
QHWIKDMRWSTSIKDEEGYPAFALVNKATGQAIKHSLGQSHPVRLVPYNPEVMDESVLWTESRD
VGNGFRCIRMVNNIYLNFDAFHGDKYHGGVRDGTDIVLWKWCEGDNQRWKIQPYYILEQKLISE
            259      264
EDLNSAVDH   HHHHH   *
```

b

```
OrysaEULS2   ------------------------------------------------------------
ArathEULS3   MEHHHQHHRHHQRDDGEDDRQSFGVPPPHVDAPPQPHGLYQSQPHFDPYAPTPQAPAPYR
EEA          ------------------------------------------------------------

OrysaEULS2   ------------------------------------------------------MDFYGR
ArathEULS3   SETQFEPHAPPPYRSEPYFETPAPPPSFGHVSHVGHQSPNESYPPEHHRYGGYQQPSNSL
EEA          ------------------------------------------------------------
OrysaEULS2   REQYGGYGGYGGGGALATPGYAPAAPYGMSQVSIEGNGCGRTLPPQPTVKVYCRANPNYA  10
ArathEULS3   LESHGDHSGVTHVAHHSSNQPQSSSGVYHKPDENRLPDNLAGLAGRATVKVYSKAEPNYN  17
EEA          ----------------------------------MASTIIATGPTYRVYCRAAPNYN  23
                                                    .*  :**.:*  ***

OrysaEULS2   MTARNGAVVLAPANPKDEYQHWIKDMRWSTSIKDEEGYPAFALVNKATGQAIKHSLGQSH  70
ArathEULS3   LTIRDGKVILAPADPSDEAQHWYKDEKYSTKVKDADGHPCFALVNKATGEAMKHSVGATH  77
EEA          MTVGKGVAFLAPIDETNELQYWYKDDTYS-YIKDEAGLPAFSLVNKATGLTLKHSNHHPV  82
                :*  .*  ..***  :  ..:*  *:*  **   :*   :**   *  *.*:*******  ::***    .

OrysaEULS2   PVRLVPYNPEVMDESVLWTESRDVGNGFRCIRMVNNIYLNFDAFHGDKYHGGVRDGTDIV  130
ArathEULS3   PVHLIRYVPDKLDESVLWTESKDFGDGYRTIRMVNNTRLNVDAYHGDSKSGGVRDGTTIV  137
EEA          PVKLVTYNPNVVDESVLWSQADDRGDGYSAIRSLTNPASHLEAAPLNDWS---YNGAIIM  139
               **:*:  *  *:  :******:::  *  *:*:  **  :.*    ::.*    :.        :*:  *:
                              ▼        ▼▼
OrysaEULS2   LWKWCEGDNQRWKIQPYY-  148
ArathEULS3   LWDWNKGDNQLWKIFPF--  154
EEA          GGVWIDAYNQQWKIEPHTG  158
                 *  .. ** *** *.
```

Fig. 2B.2 (a) Sequence of recombinant EULS2 expressed in *Pichia*, preceded by an N-terminal signal peptide (residues 1 to 89) necessary for secretion and a C-terminal tag containing a c-myc epitope and a (His)6 tag (residues 259 to 264). The cleavage sites for the signal peptide are indicated (Kex2 protease site at position 86 and Ste 13 protease sites at positions 87 and 89). The N-terminal sequence of recombinant EULS2 determined by Edman degradation is underlined. (b) Sequence alignment for OrysaEULS2, ArathEULS3 and EEA. Amino acid residues suggested to be important for the formation of the carbohydrate binding site are indicated in green and bold. Mutated residues are indicated by arrows. Trp 132 for OrysaEULS2 and Trp 139 for OrysaEULS3 are shaded in grey and yellow, respectively. The unrelated N-terminal sequences in OrysaEULS2 and ArathEULS3 are shown in red and italics. Residues indicating the start of the EUL domain in the three proteins are underlined. *: identical residues, colon (:) similar residues and dot (.) nearly similar residues

2B.4.3 Purification and characterization of the recombinant proteins from *P. pastoris*

Mutational analysis was conducted to validate the putative carbohydrate binding site in the EULS2 domain of OrysaEULS2. Therefore some amino acids predicted to be part of the carbohydrate binding site were mutated (Fig. 2B.2b). In mutant form 1, a point mutation was introduced into the coding sequence of EULS2 to replace Trp134 (Trp190 in full protein sequence) by leucine. A second point mutation involving the change of Asn139 (Asn195 in full protein sequence) into serine was introduced in the coding sequence of mutant form 1, resulting in mutant form 2. To construct mutant form 3, Glu140 (Glu196 in full protein sequence) was mutated into leucine starting from the coding sequence of mutant form 2. As a result mutant forms 2 and 3 contain two and three amino acid changes, respectively.

The coding sequences for the native EUL domain from OrysaEULS2 and all mutant forms were cloned into the *P. pastoris* expression vector pPICZαA downstream from the α-mating sequence from Saccharomyces cerevisiae (for secretion) and upstream from the c-myc epitope and a C-terminal polyhistidine tag (His tag). The resulting cassettes were electroporated into *Pichia* strain X-33. Transformed colonies were grown in 1 liter cultures and subsequently recombinant proteins were purified using several chromatographic steps. Approximately 4 to 6 mg was purified for EULS2 and its mutant forms from the medium of 1 L *Pichia* culture induced with methanol for three days. The partially purified EULS2 was also retained on a column with immobilized thyroglobulin.

As shown by SDS-PAGE analysis (Fig. 2B.3a), the molecular mass of the recombinant proteins is approximately 20.5 kDa. This result is in a good agreement with the calculated molecular mass from the primary sequence together with the C-terminal c-myc and His tags. In addition, western blot analysis using a monoclonal antibody directed against the His- tag confirmed this result (Fig 2B.3b). Edman degradation of the recombinant EULS2 domain yielded the sequence EAEAEFTMVYXRANPNYAMT and revealed that part of the signal peptide was not completely cleaved (Fig. 2B.2a), resulting in a polypeptide with a calculated molecular mass of 20.7 kDa.
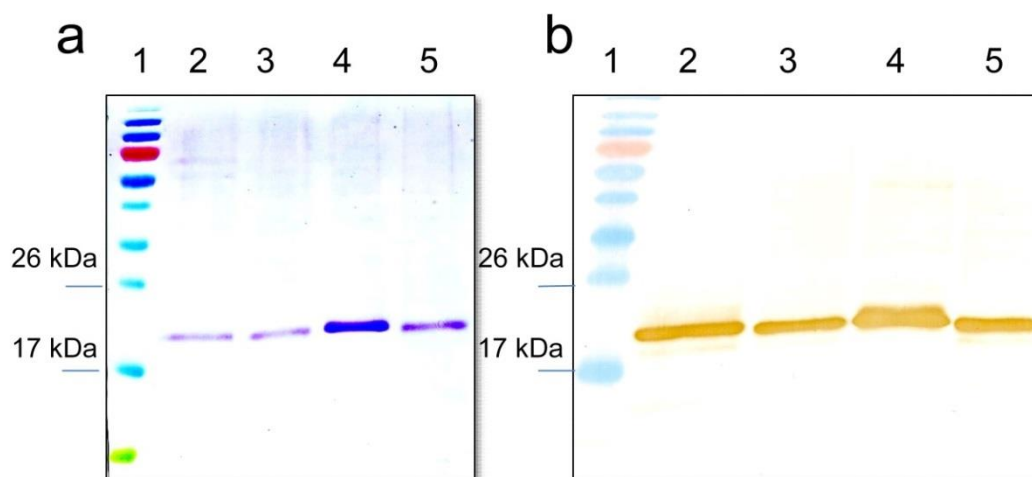
Fig. 2B.3 Purified EUL domain from OrysaEULS2 and its mutant forms were analyzed by SDS-PAGE (a) and Western blot analysis with a monoclonal anti-His antibody (b). Samples are loaded as follows: lane 1: protein ladder (increasing molecular mass: 10, 17, 26, 34, 43, 55, 72, 95, 130, 170 kDa) (Fermentas, St. Leon-Rot, Germany), lane 2: EUL domain from OrysaEULS2, lanes 3-5: mutant forms 1-3. Approximately, 1.5 µg of each protein was loaded in (a) and 0.5 µg in (b)

2B.4.4 Biological activity and carbohydrate binding specificity of the recombinant proteins

The biological activity for the recombinant proteins was investigated using agglutination tests with rabbit erythrocytes. Agglutination of red blood cells was observed for the EULS2 after 30 min, the minimal protein concentration for agglutination being 54 µg/ml. No agglutination activity was seen for any of the three mutant forms. Inhibition assays showed that agglutination of erythrocytes by the recombinant EUL domain from OrysaEULS2 was inhibited by some carbohydrates [mannose (50 mM) and methyl α-mannopyranoside (100 mM)] and glycoproteins [thyroglobulin (100 µg/ml), ovomucoid (200 µg/ml) and asialomucin (100 µg/ml)]. More detailed data were obtained by screening the labeled proteins on the glycan array.

The interactions of EULS2 and its mutant forms at 200 µg/ml with glycans on the array are shown in Fig. 2B.4. The interaction of the protein with the array was relatively weak compared to other related lectins such as EEA (Fouquaert et al., 2008) and ArathEULS3 (Van Hove et al., 2011). The data in Fig. 2B.4 were obtained

using a secondary antibody labeled with Alexa633, which is 10 times more sensitive than the normally used Alexa488 label. Nevertheless, there was sufficient binding of the wild type lectin (Fig. 2B.4a) to discern some specificity since it has been demonstrated that lectin binding to the array measured as RFU is directly related to the relative binding strength of the interaction (Smith et al., 2010). The strongest binding glycans are summarized in Table 2B.2 and indicate that EULS2 exhibits selectivity toward the high-mannose N-linked glycans, especially Man3GlcNAc2 (glycans #51 and 52) with somewhat lower binding to Man5GlcNAc2 (#217), Man6GlcNAc2 (#216), Man7GlcNAc2 (#211), Man8GlcNAc2 (#212), and Man9GlcNAc2 (#213). Interestingly, no binding was observed by the lectin to the high-mannose structures without the chitobiose core including Man3 (#214), Man5 (#215 and 315), Man8, (#316), and Man8, (#317) (data not shown in Table 2B.2) indicating the binding required the chitobiose core. All mutant forms of EULS2 showed a strongly reduced binding (at least 10-fold lower) to the glycan array (Fig. 2B.4b-d).
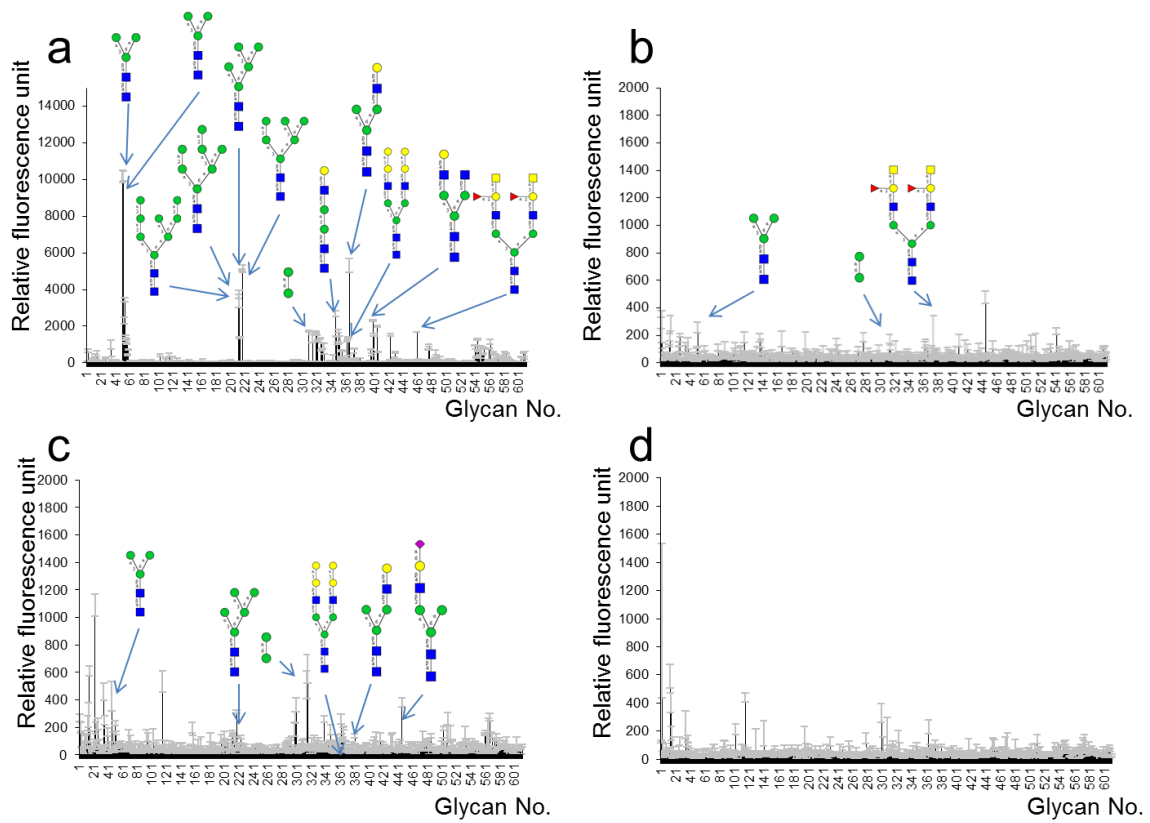
Fig. 2B.4 Comparative analysis of the glycan array binding of the recombinant EUL domain from OrysaEULS2 (a) and mutant forms 1-3 (b-d) tested at 200 µg/ml. The complete primary data set for each protein is available on the website of the Consortium for Functional Glycomics (www.functionalglycomics.org)

Table 2B.2 Overview of the top 30 glycans interacting with the EULdomain from OrysaEULS2. Lactosamine structures are underlined. Terminal GlcNAc is shown in italic. The glycan with the highestrelative fluorescence unit (RFU) is assigned a value of 100. The rank is the percentile ranking. Similar glycan structures are grouped together and ordered in a descending way according to the RFU% in each group

Glycan no	Structure	RFU%
	N-linked high-mannose glycans (in bold)	
51	Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp13	100.0
50	Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	99.8
217	Mana1-6(Mana1-3)Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	50.5
216	Mana1-6(Mana1-3)Mana1-6(Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	49.9
211	Mana1-6(Mana1-2Mana1-3)Mana1-6(Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	35.6
212	Mana1-2Mana1-6(Mana1-3)Mana1-6(Mana1-2Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	34.1
213	Mana1-2Mana1-6(Mana1-2Mana1-3)Mana1-6(Mana1-2Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	13.6
314	Mana1-6Manb-Sp10	13.1
	Lactosamine related structures (underlined)	
404	Gala1-4Galb1-3GlcNAcb1-2Mana1-6(Gala1-4Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	19.7
319	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	16.6
459	Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-2)Mana1-6(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	16.5
320	Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	15.1
54	Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	14.2
364	Gala1-3Galb1-4GlcNAcb1-2Mana1-6(Gala1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	13.6

325	Galb1-3GlcNAcb1-2Mana1-6(Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	13.1
543	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp24	12.7
362	Fuca1-2Galb1-4GlcNAcb1-2Mana1-6(Fuca1-2Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	11.7
55	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	11.5
	Mannose /Lactosamine hybrid structures	
346	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp12	26.4
398	GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp12	22.8
352	Mana1-6(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	12.3
309	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	17.1
350	Galb1-4GlcNAcb1-2Mana1-3Manb1-4GlcNAcb1-4GlcNAc-Sp12	17.0
399	Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp12	14.9
347	Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp12	14.2
321	GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	11.7
	GlcNAc related structures	
53	GlcNAcb1-2Mana1-6(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp13	33.7
52	GlcNAcb1-2Mana1-6(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	22.6
422	GlcNAcb1-2(GlcNAcb1-6)Mana1-6(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	15.0

2B.5 Discussion

In an attempt to gain more information on the carbohydrate-binding properties of the extended family of EUL proteins in rice we selected the OrysaEULS2 protein as a model. This protein is of particular interest because expression of the protein is specifically induced after stress treatment of rice plants; in particular NaCl and ABA treatment enhance the expression of the lectin (Moons et al., 1997a). Since inducible lectins are expressed in response to (a) biotic stress factor (s), it is believed that these lectins are involved in stress signaling (Van Damme et al., 2004b). In this paper we describe the expression of the EUL domain from OrysaEULS2 in *P. pastoris*. Recombinant EULS2 was secreted in the medium and purified successfully using different chromatographic steps, including an affinity chromatography on immobilized thyroglobulin, suggesting its binding to glycan structures.

After purification one polypeptide of approximately 20 kDa was detected by SDS-PAGE and western blot analysis. Edman degradation for the native EULS2 yielded an amino acid sequence corresponding to the N-terminus of the EULS2 domain containing five residues of the *Saccharomyces* signal peptide used to secrete the recombinant protein into the medium. Incomplete processing of the α-mating sequence has been reported before (Lannoo et al., 2007a; Al Atalah et al., 2011 see chapter 2A; Oliveira et al., 2008). The EA repeats are necessary for the function of Kex2 protease but removing them with the Ste 13 protease has been shown to be an inefficient process (Sreekrishna et al., 1997). The recombinant EULS2 protein agglutinated rabbit red blood cells, albeit at a rather high protein concentration of 54 μg/ml. Since the specific agglutination activity of EEA required only 1.7 μg/ml, it can be concluded that the agglutination activity of EULS2 is low. Nevertheless the protein was bound to a thyroglobulin column and its agglutination was inhibited by mannose, methyl α-mannopyranoside and some glycoproteins. Previous data have clearly shown that most plant lectins show a much better interaction with glycans or more complex sugars, rather than monosaccharides (Van Damme et al., 2007a). Carbohydrate-binding activity of EUL domain from OrysaEULS2 was further confirmed by glycan array analyses showing interaction of the protein especially with high-mannose N-glycans and to a lesser extent with N-glycans containing lactosamine structures (Galβ1-3GlcNAc and Galβ1-4GlcNAc). Therefore it can be concluded that the OrysaEULS2 protein is a functional carbohydrate binding protein.

A three dimensional model was built for the EUL domain in OrysaEULS2 and compared to the EUL domain in ArathEULS3 and EEA. The EUL domains in all three proteins showed three bundles of β-sheet forming a β-trefoil fold. Despite the strong resemblance in overall fold between the three proteins, there are also clear differences in the overall structure and the charge distribution on the protein surface. In particular the positioning of Asn139 in the EUL domain from OrysaEULS2 could influence the carbohydrate binding properties of this protein. Previously it was shown that the hydrogen bond network anchoring galactose to both carbohydrate binding sites of the ricin-B chain predominantly involves O3 and, to a lesser extent, O4 of the sugar. An Asn residue creates a hydrogen bond with O3 of galactose which reinforces the binding of this oxygen to the lectin (PDB code 2AAI) (Rutenber et al., 1991). An Asn residue also occurs in the carbohydrate binding sites of the ricin-B domain of Streptomyces olivaceoviridis xylanase (PDB code 1ISZ) (Fujimoto et al., 2002) and the C-terminal domain of the HA33/A protein from *Clostridium botulinum* (PDB code 1YBI) (Arndt et al., 2005), which similarly participate in the binding of the sugar. Residues Asn148 of EEA and Asn146 of the EUL domain in ArathEULS3, which occupy a similar position in the respective lectins, could also contribute to the binding of the sugars to the active carbohydrate binding sites. However, the corresponding Asn139 in the EUL domain from OrysaEULS2 is apparently masked by an extended loop that protrudes in the vicinity of the carbohydrate binding site, and thus this residue should no longer be available for a hydrogen bond interaction with the sugar. According to this structural discrepancy, the EUL domain in OrysaEULS2 is predicted to display a weaker carbohydrate-binding capacity, compared to the EEA and ArathEUL3 lectins, which is in agreement with the low specific agglutination activity observed for EULS2. Possibly this loop can also be (partly) responsible for some of the changes in specificity between the different EUL lectins.

To validate our proposed model for the EUL domain in OrysaEULS2 and especially the position of the carbohydrate binding site, three mutant forms were created, expressed in *Pichia* and the mutant proteins were purified. All three mutants clearly showed a reduced binding to the glycan array since the glycan interaction measured as relative fluorescence units was approximately 10 times lower compared to the native protein. This is also in agreement with the absence of agglutination activity for the mutant proteins.

Due to the low binding of recombinant EUL domain from OrysaEULS2 with the array it is difficult to interpret the results for the different mutants quantitatively. In any case reduced binding was observed for all mutants indicating that the mutation of a single residue (Trp134 as in mutant 1) already affected the carbohydrate-binding properties of the EUL domain. The importance of Trp for the configuration of a functional carbohydrate binding site was reported for many lectins belonging to different lectin families (Schouppe et al., 2010). The results obtained with the other mutants do not allow us to conclude if the additional mutation of residues Asn139 and Gln140 yielded a stronger reduction of the carbohydrate binding activity.

Several nucleocytoplasmic lectins have been identified in rice, among which multiple putative lectins with EUL domains and one jacalin-related lectin, Orysata. The latter protein has been characterized in detail and its carbohydrate binding properties were investigated using the glycan array technology. Orysata shows affinity towards high-mannose and complex N-glycans (Al Atalah et al., 2011; see chapter 2A). The proteins reacts equally well with both types of glycans. A closer investigation of the carbohydrate specificity of Orysata also revealed that among the complex N-glycans interacting with Orysata there are several glycans containing lactosamine structures. In this respect the specificity of Orysata resembles that of the EUL domain within OrysaEULS2. However, it should be mentioned that OrysaEULS2 reacts much better with the high-mannose structures compared to the more complex structures.

OrysaEULS2 belongs to the family of EUL-related lectin since its EUL domain shows 46% sequence identity to the sequence of EEA. It was previously reported that the specificity of EEA is directed towards two major groups of glycans, being the blood type B (Galα1–3(Fucα1–2) Galβ1–4GlcNAc) oligosaccharides and N-linked high-mannose glycans (Fouquaert et al., 2008). Since the fluorescence units for the high-mannose N-glycans were roughly 10-fold lower than for the blood group B oligosaccharides it was concluded that EEA has a much higher affinity for the blood group B substances. A comparative analysis of the glycan array data for EEA and the recombinant EUL domain from OrysaEULS2 suggests the opposite binding pattern for OrysaEULS2 in that the interaction with high-mannose N-glycans is much better than with complex structures. However, it should be mentioned that the binding of OrysaEULS2 with the array was much weaker than for EEA.

Recently another lectin from the family of EUL-related lectins was analyzed on the glycan array. ArathEULS3 from *Arabidopsis thaliana* belongs to the group of S3 type EUL proteins which is ubiquitous in the plant kingdom, and therefore more widespread than the EULS2 type. Both the full protein ArathEULS3 and its EUL domain were expressed in *Pichia* and the recombinant proteins were analyzed on the glycan array. Both proteins yielded similar results and reacted with glycans containing one or more Lewis X (Galβ1-4(Fucα1-3) GlcNAc), Lewis Y (Fucα1–2Galβ1–4(Fucα1–3)GlcNAc) or lactosamine motifs (Van Hove et al., 2011). Both for ArathEULS3 and EEA there is evidence that lactosamine structures in both N- and O-glycans will react with these lectins, whereas in EULS2 and Orysata only lactosamine structures in N-glycans were reactive.

From the comparative analysis between EEA, OrysaEULS2 and ArathEULS3, belonging to the type S0, type S2 and type S3 class of EUL-related lectins, respectively, it is clear that even though these three lectins belong to the same family, their specificity has evolved in a different way. Our results indicate that gene divergence within the family of EUL-related lectins lead to changes in carbohydrate binding specificity. Previously it was also shown that gene divergence within the legume lectin family (Loris et al., 1998), the jacalin-related lectins (Rougé et al., 2003) and the GNA-related lectins (Fouquaert et al., 2009b) has resulted in changes in carbohydrate-binding specificity. Furthermore evidence also shows that ricin-B domains display considerable plasticity in their carbohydrate-binding specificity. For instance, the type 2 ribosome-inactivating protein from Dutch Iris exhibits specificity towards Gal/GalNAc but also towards mannose, indicating that ricin-B domain can also accommodate mannose (Hao et al., 2001). Similarly the carbohydrate-binding module of xylanase 10A from *Streptomyces lividans* which shows structural similarity to the ricin-B domain cannot only bind lactose and galactose but also interacts with the polysaccharide xylan (Notenboom et al., 2002). Structural evidence suggests that subtle changes in the amino acids building the carbohydrate-binding site or surrounding the site can provoke changes in its specificity (Van Damme et al., 2007a). Hence one should be very careful when trying to predict the carbohydrate-binding properties of lectins.

Blood group B, Lewis X, Lewis Y and lactosamine structures are well-studied in higher animals (including human beings) (Stanley et al., 2009), bacteria and viruses (Preston et al., 1996; Wang et al., 2000; Monzavi-Karbassi et al., 2004). However, in plants only the Lewis A

(Galβ1-3(Fucα1-4) GlcNAc) motif has been identified (Leonard 2002; Melo et al., 1997b; Dam et al., 2011). At present very little is known about the distribution and physiological importance of lactosamine related structures in plants. Therefore, more research is needed in order to elucidate the importance of the carbohydrate-binding activity for the physiological role of OrysaEULS2.

Chapter 2

Characterization of the lectins under study expressed in *Pichia pastoris*

Chapter 2C

Characterization of a type D1A EUL-related lectin from rice expressed in *Pichia pastoris*

Redrafted from:

Al Atalah B, Vanderschaeghe D, Bloch Y, Proost P, Plas K, Callewaert N, Savvides SN, Van Damme EJM (2013) Characterization of a type D1A EUL-related lectin from rice expressed in *Pichia pastoris*. Biol Chem, in press

2C.1 Abstract

OrysaEULD1A is one of the five EUL genes determined in rice (*Oryza sativa*) which encode putative carbohydrate-binding proteins belonging to the *Euonymu*s lectin (EUL)-related family. The OrysaEULD1A sequence comprises two almost identical EUL domains (91% sequence similarity and 72% sequence identity) separated by a 23 AA linker domain and preceded by an 19 AA N-terminal sequence. In the present study, OrysaEULD1A domain 1, OrysaEULD1A domain 2 and the full OrysaEULD1A were expressed in *Pichia pastoris*. After purification of the recombinant proteins, their carbohydrate binding specificity was analyzed and compared. All recombinant proteins agglutinate red blood cells indicating that the full OrysaEULD1A and its domains are true lectins. Multi-angle light scattering analysis showed that the protein corresponding to EULD1A domain 1 is mainly monodiperse. On the contrary, recombinant OrysaEULS2 is polydisperse. Interestingly, all recombinant lectins showed clear specificity towards galactosylated structures. Our results indicate that although the amino acids, responsible for the formation of the carbohydrate binding site, are identical for all EUL proteins, they show different carbohydrate specificities. This promiscuity of the carbohydrate-binding site is most probably due to the gene divergence which took place within the EUL family.

2C.2 Introduction

Since the discovery of these stress inducible lectins, at least six carbohydrate binding domains have been identified within the group of nucleocytoplasmic plant lectins (Lannoo and Van Damme 2010). Of particular interest are the *Euonymus* lectin sequences which are present in all sequenced genomes. The EEA was cloned a few years ago (Fouquaert et al., 2008). Since then EEA has become the prototype of a new family of proteins referred to as the *Euonymus* lectin-related (EUL) family, grouping all proteins with one or more EUL-related sequences. *In silico* screening for available genomes and transcriptomes revealed that the EUL family is omnipresent in the plant kingdom (Fouquaert et al., 2009a). Based on the architecture of the EUL domain identified in different protein sequences, a classification system was elaborated (Fouquaert et al., 2009a). Most EUL sequences consist of a single EUL domain preceded by a variable and unrelated N-terminal domain. However some EUL sequences contain two in tandem arrayed EUL domains separated by a linker sequence, preceded by an unrelated N-terminal domain which can vary in length. These proteins with a single EUL domain and with two (double) EUL domains are referred to as the type S and the type D, respectively. Based on the length of the N-terminal domains, the presence of a C-terminal unrelated domain and the length of the linker sequence separating the EUL domains more subtypes can be distinguished (for more details see Fouquaert et al., 2009a).

In rice, four types of EUL proteins have been identified (Fouquaert et al., 2009a, Fouquaert and Van Damme 2012; Al Atalah et al., 2013 see chapter 3A). Two proteins with a single EUL domain preceded by an unrelated N-terminal domain are classified as type S2 (OrysaEULS2) and type S3 (OrysaEULS3). Both sequences differ from each other in the length and the sequence of the N-terminal domain. Furthermore the rice genome harbours three sequences with tandem arrayed EUL domains, referred to as type D1 and D2, differing from each other in the sequence and the length of the linker sequences between the two EUL domains. Within the type D1 two almost identical sequences can be distinguished. The corresponding proteins are called OrysaEULD1A and OrysaEULD1B.

Until recently, only few reports provided evidence that EUL sequences encode true lectins, particularly OrysaEULS2 from the monocot *Oryza sativa* (Al Atalah et al., 2012; see chapter 2B), ArathEULS3 from the dicot *Arabidopsis thaliana* (Van Hove et al., 2011) and PhypaEULS3

from the lower plant *Physcomitrella patens* (Fouquaert and Van Damme 2012) have been characterized in some detail. Although the EUL domain sequence is quite conserved, the carbohydrate-binding sites have evolved to recognize different glycan structures. Whereas EEA exhibits specificity towards blood-group B related structures, OrysaEULS2 shows affinity against high-mannose N-glycans (Al Atalah et al., 2012; see chapter 2B). ArathEULS3 and PhypaEULS3 display a clear preference for galactosylated epitopes (Fouquaert and Van Damme 2012).

OrysaEULD1A is composed of a 22 AA unrelated N-terminal domain and two almost identical EUL domains separated by a 19 AA linker sequence. To gain more knowledge about the physiological role of OrysaEULD1A, it is a prerequisite to prove its biological activity. Since OrysaEULD1A is expressed at very low concentrations it is impossible to purify reasonable amounts of the protein from rice plants. Furthermore it would be a challenge to separate the different EUL types. To overcome this problem, the D1A sequence was expressed in *Pichia pastoris* under the control of the AOX1 promoter, known to induce high levels of protein after treatment with methanol (Hartner and Glieder 2006). In recent years, multiple nucleocytoplasmic lectins have been successfully expressed and purified from *P. pastoris* (Lannoo et al., 2007a; Fouquaert et al., 2009b; Al Atalah et al., 2011 see chapter 2A; Van Hove et al., 2011; Al Atalah et al., 2012 see chapter 2B; Stefanowicz et al., 2012).

In this study, we characterized for the first time an EUL protein of the type D consisting of two EUL domains. The full sequence of OrysaEULD1A as well as its individual EUL domains, OrysaEULD1A domain 1 and OrysaEULD1A domain 2 (further referred to the full EULD1A, EULD1A domain 1 and EULD1A domain 2, respectively) were expressed in *P. pastoris*. Afterwards, the molecular structure and biological activity of the recombinant proteins was characterized, and compared to other EULs which have been reported previously.

2C.3 Methods and materials

2C.3.1 Expression of EULD1A and its domains in *P. pastoris*

Expression for the full EULD1A, EULD1A domain 1 and EULD1A domain 2 sequences was conducted following the EasySelect *Pichia* Expression Kit from Invitrogen (Invitrogen,

Carlsbad, CA USA). Therefore, the coding sequence for full EULD1A (Genbank accession number AK103324) as well as the sequence encoding EULD1A domain 1 and EULD1A domain 2 were amplified from the pFLCI vector (obtained from the Rice Genome Resource Center, Ibaraki, Japan) by PCR using different sets of primers: EVD 515 / 5'GGCGGAGAATTCACCATG TTTGGGTTCGGGCACCACCAC3' and EVD 716 /5'CCCGCTTTCTAGAATCAATGCAATGCAGCTCCC CCTCCCCGGAATCGA3' to amplify the full EULD1A, EVD 515 and EVD 644 /5'CCCGCTTTCTAGA ATGGTGGCGGCGCCGGGGGAGGAGGAGT3' to amplify EULD1A domain 1, EVD 645 /5'GGC GGAGAATTCGGCGCCGCCACCGCCTGCACCATCGGCG3' and EVD 516 /5'CCCGCTTTCTAGAATC CAGGGGAGGATCTTCCAGCGCTG3' to amplify EULD1A domain 2. The following PCR conditions were used: 2 min denaturation at 94°C, 25 cycles of 15 sec 94°C, 30 sec 60°C, 1 min 72°C, ending with an additional 5 min elongation at 72°C. EcoRI and XbaI cleavage sites were added at the 5' and 3' ends, respectively, of the PCR primers to allow cloning of the amplified fragments into the pPICZαA vector. Sequence analysis of the cloned fragments was performed using the primers EVD 21 5'GACTGGTTCCAATTGACAAGC3' and EVD 22 5'GCAAATGGCATTCTGACATCC3' located on the vector (carried out by LGC Genomics, Berlin, Germany). Subsequently, the pPICZαA vector harboring each of the constructs was transformed into the *Pichia* strains X-33 and KM71H as described previously (Al Atalah et al., 2012; see chapter 2B). This vector contains the α–mating sequence that enables secretion of the recombinant proteins into the medium which in turn allows easy purification. This vector also contains a polyhistidine tag which allows detection of the recombinant proteins as well as purification by metal affinity chromatography. After *Pichia* transformation, several colonies per construct were tested for EUL expression and the colonies that yielded the highest expression were selected to grow larger *Pichia* cultures following the procedure described by Al Atalah et al. (2011; see chapter 2A), except that *Pichia* cells were grown and expression was induced at 22°C instead of 30°C.

2C.3.2 Purification of the recombinant proteins

Purification of the recombinant proteins was performed as explained in section 2A.3.4 without using the mannose-Sepharose 4B column. The Coomassie Protein Assay Kit (Thermo Fischer Scientific, Rockford, IL USA) was used to measure the protein concentration following the Bradford dye-binding procedure (Bradford, 1976). After each purification step, the purity

of the protein samples was checked by SDS-PAGE (Laemmli 1970) and/or western blot analyses as described by Al Atalah et al. (2012; see chapter 2B).

2C.3.3 N-terminal protein sequencing

The recombinant proteins were sequenced as described in the section 2A.3.5.

2C.3.4 Agglutination assay

To investigate the biological activity for the recombinant proteins agglutination assays were performed using trypsin-treated rabbit red blood cells (BioMérieux, Marcy l'Etoile, France). The procedure explained in 2A.3.6 was followed.

2C.3.5 Carbohydrate inhibition assay

The carbohydrate specificity for each of the proteins under study was analyzed by carbohydrate inhibition assays, using several sugars (mannose, methyl α-mannopyranoside, trehalose, glucose, galactose, GlcNAc, GalNAc, methyl α-galactopyranoside and lactose at 0.5 M) and glycoproteins (thryoglubulin, ovomucoid, asialomucin, mucin and fetuin at 10 mg/ml). The concentration required to inhibit the agglutination of trypsin-treated rabbit erythrocytes was estimated at a lectin concentration of 20 µg/ml for the full EULD1A and the EULD1A domain 2 as well as 100 µg/ml for the EULD1A domain 1 and OrysaEULS2.

2C.3.6 Protein deglycosylation

Recombinant proteins for full EULD1A and EULD1A domain 1 were digested with N-glycosidase F (PNGase F). Briefly, 2 µg recombinant proteins for each sample were mixed with 1x denaturation buffer (0.5% SDS and 0.04 M dithiothreitol) in a volume of 10 µl. The samples were boiled at 100°C for 10 min and afterwards cooled down. To start the digestion, the denatured samples were mixed with 2 µl of 10x reaction buffer (0.5 M sodium phosphate pH 7.5), 2 µl 10% NP-40 and 5.5 µl pure water to adjust the volume into 20 µl. Finally, 0.5 µl of PNGase F was added to each sample and all samples were incubated at 37°C for 4 h. RNAseB (2 µg) was used as a positive control. After digestion, protein samples were analyzed by SDS–PAGE.

2C.3.7 Multi-angle light scattering analysis (MALLS)

Molecular masses were determined by MALLS. One hundred µl protein samples (0.5 mg/ml) were separated using a HPLC system (Agilent 1260 Infinity) coupled to a size-exclusion column (Wyatt WTC 030S5) equilibrated with 20 mM DAP (pH 8.5) running buffer and eluted at 0.5 ml/min. The in-line MALLS setup consisted of an UV-detector (Shimadzu SPD-10A), a miniDAWN TREOS MALLS detector (Wyatt) and a T-rEX refractive index detector (Wyatt) (thermostated at 25°C). Data analysis was carried out using the ASTRA 6 software (Wyatt), a refractive index increment (dRI) of 0.185 ml/g was used for the protein concentration and consequent molecular mass determination.

2C.4 Results

2C.4.1 Cloning and expression of the recombinant proteins in *P. pastoris*

Due to the high sequence identity between the two EUL domains PCR amplification of the full EULD1A sequence was complicated in that two PCR fragments corresponding to the full EULD1A sequence and the EULD1A domain 1 sequence were always amplified. To overcome this problem, a reverse primer was designed in the 3' un-translated region of the OrysaD1A sequence and a point mutation was created in the stop codon. As a consequence the original stop codon was mutated into a Ser residue and the coding sequence was extended (11 AA at protein level) with a sequence unrelated to the EUL sequence, allowing designing a specific primer to amplify the full EULD1A sequence (Fig. 2C.1).

Sequences encoding full EULD1A, EULD1A domain 1 and EULD1A domain 2 were cloned under the control of the methanol-inducible AOX1 promoter present in the pPICZαA expression vector. The coding sequence was cloned in-frame with the yeast α-factor mating signal sequence at the N-terminus and with a c-myc epitope followed by a polyhistidine tag at the C-terminus (Fig. 2C.1). Afterwards, the expression cassettes were introduced into the Pichia strains X-33 and KM71H by electroporation. Expression for each of the proteins under study was first investigated by performing small scale induction experiments at 30°C. However, western blot analysis revealed no expression of the protein for the full EULD1A sequence in none of the tested *Pichia* strains. At 30°C the expression for the EULD1A

domain 1 was very faint whereas the expression for the EULD1A domain 2 sequence was moderate. In an attempt to increase the expression levels of the recombinant proteins, the *Pichia* cultures were grown at 22°C and methanol-induced expression was also conducted at 22 C°. At this temperature, the full EULD1A sequence yielded good expression in both *Pichia* strains but expression was higher in the KM71H strain compared to the X-33 strain (Results not shown). Furthermore, the expression of the EU1D1A domain 1 and EULD1A domain 2 sequences was also considerably enhanced, with no difference in growth and protein expression between both strains. Based on these results, the KM71H strain was selected for expression of the full EULD1A sequence and the X-33 strain was used for the expression of the individual EUL domains. Finally, all recombinant proteins were purified from the *Pichia* culture medium by means of ion exchange and metal-affinity chromatography. Starting from 1 l culture, the yield of purified protein was 1 mg for the full EULD1A construct, 3 mg for the EULD1A domain 1 and 0.5 mg for the EULD1A domain 2 constructs.

a

<u>EFTMFGFG</u>HHHNQAPAAPSDPNQIFKIFCRANENYCLTVRDSAVVLAPVNPKDEHQHWFKDM
RFSTKVKDGEGMPAFALVNKATGLAVKHSLGQSHPVKLVPFNPEYEDASVLWTESKDVGKGF
RCIRMV**NNT**RLNLDAFHGDKDHGGVRDGTTVVLWEWCKGDNQSWKILPWGPEAHSSSPGAAT
ACTIGGVSVHTVRVFSAAGEDYCLTVRNGTACLAPKNPRDDYQHWIKDMRHSNKIRDEEGYP
AFALVNKVTGEAIKHSTGQGHPVKLVPYNPEYQDESVLWTESKDVGKGFRCIRMVNNIYLNF
DAFHGDKDHGGIHDGTEIVLWKWCEGDNQRWKILPW<u>**S**</u>IPGRGSCIALILEQKLISEEDLNSA
VD*HHHHHH*

b

<u>EFTMFGFGHHHNQAPAAP</u>SDPNQIFKIFCRANENYCLTVRDSAVVLAPVNPKDEHQHWFKDM

RFSTKVKDGEGMPAFALVNKATGLAVKHSLGQSHPVKLVPFNPEYEDASVLWTESKDVGKGF

RCIRMV**NNT**RLNLDAFHGDKDHGGVRDGTTVVLWEWCKGDNQSWKILPWGPEAHSSSPGAAT

ILEQKLISEEDLNSAVD*HHHHHH*

c

<u>EFGAATACTIGG</u>VPVHTVRVFSAAGEDYCLTVRNGTACLAPKNPRDDYQHWIKDMRHSNKIR

DEEGYPAFALVNKVTGEAIKHSTGQGHPVKLVPYNPEYQDESVLWTESKDVGKGFRCIRMVN

NIYLNFDAFHGDKHGGIHDGTEIVLWKWCEGDNQRWKILPWILEQKLISEEDLNSAVD*HHHH*

HH

Fig. 2C.1 Sequences for the full EULD1A (a), EULD1A domain 1 (b) and EULD1A dpmain 2 (c) expressed in *P. pastoris*. Residues sequenced by Edman degradation are underlined. The putative glycosylation site is indicated in bold. Amino acid mutated (accidently) by Taq DNA polymerase enzyme is shaded in grey (the original amino acid is proline). The Serine residue resulting from the mutation of the stop codon is shown in bold and underlined. Extra amino acids added to the C-terminus of the full D1A sequence are shaded in grey and the font is shown in white. The myc epitope is shaded in black and the font is shown in white. The polyhistidine tag is shown in italics

2C.4.2 Characterization of the recombinant proteins

SDS-PAGE analysis of the recombinant proteins purified by affinity chromatography on a nickel column, followed by concentration of the protein by anion exchange chromatography, yielded three polypeptides of approximately 45, 42 and 15 kDa for the full EULD1A protein (Fig. 2C.2a, lane 4). A very similar pattern was observed after western blot analysis using a monoclonal antibody specifically directed against the polyhistidine tag (Fig. 2C.2b, lane 4).

However, it should be mentioned that the 42 kDa polypeptide was hardly visible on gel and on blot. Taken into account the calculated molecular mass of the full protein including the c-myc and His tags (42.672 kDa), the size of the 42 protein band would be expected. The identity of the protein bands was confirmed by Edman degradation which yielded EFTMFGF. The smallest polypeptide of 15 kDa most probably corresponds to a degradation product, as confirmed by N-terminal protein sequencing.

The recombinant protein corresponding to EULD1A domain 1 yielded two polypeptides of 26 and 23 kDa. Again the lower polypeptide can hardly be detected on a coomassie stained gel or on western blot (Fig. 2C.2a, b, lane 2). Taking into account the additional tags present on the recombinant protein, its mass was calculated as 23.448 kDa. Protein sequencing revealed identical sequences for both polypeptides, being EFTMFGFGHHHNQAP. This sequence is identical to the N-terminus of the OrysaD1A sequence or the D1A domain.

The recombinant protein for EULD1A domain 2 yielded one clear polypeptide of 22 kDa, which is in good agreement with the calculated mass for the protein (21,379 kDa) (Fig. 2C.2a, b, lane 3). Edman degradation yielded the sequence EFGAATA<u>X</u>TIG, which corresponds to the N-terminal sequence of EULD1A domain 2.
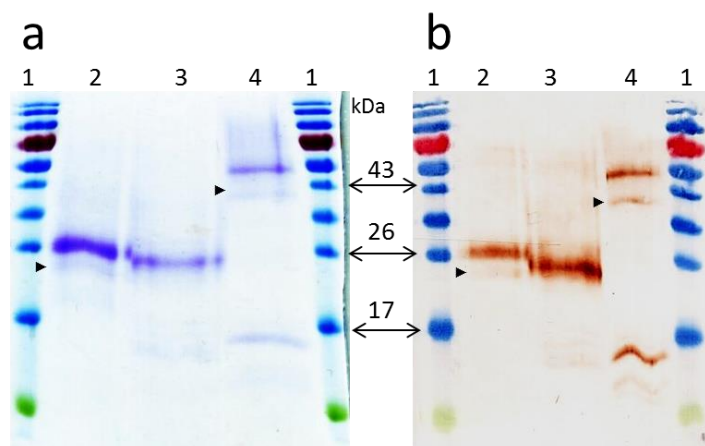
Fig. 2C.2 SDS-PAGE (a) and Western blot analysis (b) of purified full EULD1A, EULD1A domain 1 and EULD1A domain 2. Samples are loaded as follows: lane 1: protein ladder (increasing molecular mass: 10, 17, 26, 34, 43, 55, 72, 95, 130, 170 kDa) (Fermentas, St Leon-Rot, Germany), lane 2: EULD1A domain 1, lane 3: EULD1A domain 2, lane 4: full EULD1A. ▶: indicates the position of the un-glycosylated polypeptides. Approximately, 1 µg of each protein was loaded in (a), 0.5 µg for EULD1A domain 2 and 0,2 µg for each of full D1A and EULD1A domain 1 were loaded in (b)

EULD1A domain 1 was further characterized by SEC-MALLS. Low sample recovery was observed due to interaction with the stationary phase of the SEC-column, this resulted in lower quality measurements. The eluting fraction shows a single broad peak on the SEC profile containing a monodisperse region with a molecular weight (MW) of 35.1 kDa (± 9.0%). The remainder of the peak is polydisperse with maximal MW of around 70 kDa. In addition, OrysaEULS2 was also characterized by SEC-MALLS. This protein elutes as two peaks. The largest molecular species has a MW of 60.5 kDa (± 12.2%) and the second smaller molecular species has a MW about half that [33.9 kDa (± 14.4%)].

2C.4.3 Glycosylation analysis

A close investigation of the OrysaD1A sequence revealed the presence of a putative glycosylation motif NNT located within the EULD1A domain 1 (Fig. 2C.2a, b). As a result expression of the full EULD1A sequence or its corresponding EULD1A domain 1 could result in (partial) glycosylation of the secreted proteins. To determine if these recombinant proteins are N-glycosylated, they were digested with PNGase F. Subsequent SDS-PAGE analysis revealed a shift of the protein polypeptide corresponding to the full protein by appr. 23 kDa. Similarly PNGase F treatment of the recombinant protein corresponding to EULD1A domain 1 also reduced its molecular mass, resulting in one polypeptide of 3 kDa after removal of the N-glycan (Fig. 2C.3).
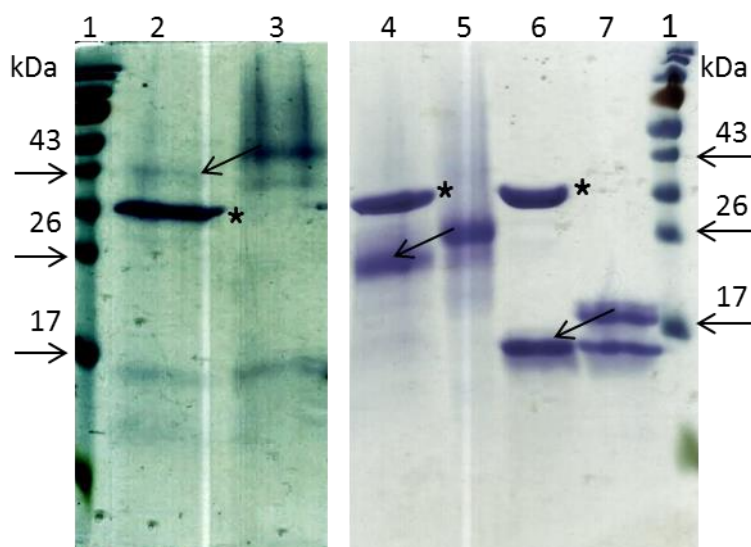

Fig. 2C.3 Legend on next page

Fig. 2C.3 Deglycosylation of the recombinant full EULD1A and EULLD1A domain 1. Samples are loaded as follows: lane 1: protein ladder (increasing molecular mass: 10, 17, 26, 34, 43, 55, 72, 95, 130, 170 kDa) (Fermentas, St Leon-Rot, Germany), lane 2: full EULD1A treated with PNGase F, lane 3: full EULD1A without treatment, lane 4: EULD1A domain 1 treated with PNGase F, lane5: EULD1A domain 1 without treatment, lane 6: RNAseB treated with PNGase F, lane 7: RNAseB without treatment. The protein band corresponding to PNGase F is indicated with an asterisk. The arrow depicts the shift of the band after PNGase F treatment. In each well, 2 µg protein was loaded

2C.4.4 Agglutination activity and specificity of the recombinant proteins

The biological activity for the different recombinant proteins under study was investigated by agglutination assays using trypsin-treated rabbit red blood cells. Using twofold serial dilutions, the lowest concentration of the purified protein needed for agglutination was determined, being 7 µg/ml for the full EULD1A, 52 µg/ml for EULD1A domain 1 and 8 µg/ml for EULD1A domain 2.

Carbohydrate inhibition assays revealed that the agglutination of rabbit erythrocytes mixed with the recombinant proteins was best inhibited by the sugars: galactose (50 mM), GalNAc (50 mM), methyl α-galactopyranoside (50 mM) and lactose (0.5 mM). Among the glycoproteins tested, only asialomucin (500 µg/ml) and mucin (1000 µg/ml) inhibited the agglutination of the different proteins at similar concentrations. No inhibitory effect on agglutination was observed for any of the other sugars tested including mannose, methyl α-mannopyranoside, GlcNAc, glucose and trehalose, nor for the glycoproteins ovomucoid, thyroglobulin and fetuin.

Table 2C.1 Comparison of the carbohydrate specificity of EUL proteins from rice. -: no inhibition of agglutination. IC50 is the inhibitory concentration required to give a 50% inhibition of the agglutination of trypsin-treated rabbit erythrocytes at a lectin concentration of 20 µg/ml for the full EULD1A and the EULD1A domain 2 as well as 100 µg/ml for the EULD1A domain 1 and OrysaEULS2

	IC 50			
	EULD1A	EULD1A domain 1	EULD1A domain 2	EULS2 domain
Sugars (mM)				
Mannose	-	-	-	50
Methyl α-mannopyranoside	-	-	-	100
Galactose	50	50	50	-
GalNAc	50	50	50	50
Methyl α-galactopyranoside	50	50	50	50
Lactose	0.5	0.5	0.5	0.5
GlcNAc	-	-	-	-
Glucose	-	-	-	-
Trehalose	-	-	-	-
Glycoproteins (µg/ml)				
Ovomucoid	-	-	-	200
Thyroglobulin	-	-	-	100
Asialomucin	500	500	500	100
Mucin	1000	1000	1000	500
Fetuin	-	-	-	-

2C.5 Discussion

Screening of the available genome sequences from plants revealed that the EUL domain is widespread in terrestrial plants, ranging from flowering plants to spermatophyte, ferns, mosses and liverworts. The EUL domain sequence is quite conserved and can be retrieved from all completed plant genomes, suggesting some physiological importance for this class of proteins. A more detailed analysis of the EUL sequences revealed that EUL sequences from dicot species usually contain one single EUL domain which most often is preceded by an unrelated N-terminal sequence. However, a search for EUL domains in monocotyledons revealed that these monocot species express a set of single-domain EUL sequences, but in addition also have two-domain EUL sequences. Similar to other grass species the rice (*O. sativa)* genome contains at least 5 EUL sequences with different domain architecture.

Until now only two single EUL domain sequences have been studied in some detail for their carbohydrate binding properties, in particular the S3 type protein from *Arabidopsis* (Van Hove et al., 2011) and the S2 type from rice (At Atalah et al., 2012). Here we describe for the first time the characterization of a EUL-related protein composed of two almost identical (91% and 72% sequence similarity and identity at amino acid level, respectively) EUL domains. The full EULD1A sequence from rice as well as its individual EUL domains EULD1A domain 1 and EULD1A domain 2, have been successfully expressed in *P. pastoris*. On average, the yield obtained for the recombinant proteins can be considered low compared to the yield reported previously for some other plant lectins (Al Atalah et al., 2011 see chapter 2A; Lannoo et al., 2007a). However a yield of 0.5-3 mg recombinant protein per liter is comparable to the yield obtained for OrysaEULS2 (4-6 mg/l, Al Atalah et al., 2012; see chapter 2B). Protein expression of the full EULD1A or its EUL domains was achieved by growing *Pichia* cultures at 22°C, a temperature considerably lower than the temperature of 30°C which is usually recommended. Lowering the temperature considerably increased the expression level of the proteins. Low temperature reduces the rate of protein synthesis and this will give enough time for the nascent peptides to fold properly. Previously, several heterologous proteins were expressed successfully in *P. pastoris* at low temperatures 20-22°C (Murasugi et al., 2001; Jahic et al., 2003; Murasugi and Tohma 2003). All D1A sequences were cloned behind the α-mating sequence of *Saccharomyces*, resulting in secretion of the recombinant proteins into the medium. Although this secretion facilitates

purification of the protein it also caused N-glycosylation of the full EULD1A as well as the EULD1A domain 1. Judging from the SDS-PAGE and the western blot analysis it was clear almost all recombinant proteins were glycosylated since the unglycosylated polypeptide was hardly visible (Fig. 2C.2 lanes 2, 4). PNGase F digestion and glycan analysis confirmed the glycosylation of the recombinant protein.

Judging from the ability of the full EULD1A to agglutinate rabbit erythrocytes and bind carbohydrate structures, it can be concluded that this recombinant protein is a functional lectin. Interestingly, both individual EUL domains of the full EULD1A were also active. The protein concentration needed to agglutinate rabbit red blood cells was very similar (7 and 8 µg/ml) for the full EULD1A and the EULD1A domain 2, but was higher for the EULD1A domain 1 being 52 µg/ml. These agglutination titers are similar to the ones reported before for OrysaEULS2 (54 µg/ml, Al Atalah et al., 2012; see chapter 2B) and PhypaEULS3 (92 µg/ml, Fouquaert and Van Damme 2012). The high concentration required for the EULD1A domain 1 might be explained by the results of the MALLS analysis which showed that most of this protein was monodisperse (35 KDa) and some was polydisperse (70 KDa). On the contrary, most of the OrysaEULS2 was polydisperse (60 KDa) and some was most likely monodisperse (33.9 kDa). These results are in a line with the fact that gene divergence within the EUL family led to differences in specificity and most likely a different function. Agglutination of the recombinant protein for OrysaEULD1A was inhibited by different sugars and glycoproteins. Interestingly, the full EULD1A, EULD1A domain 1 and EULD1A domain 2 showed a very similar carbohydrate-binding specificity directed towards galactose related sugars or galactose containing glycoproteins (Table 2C.1). Mannose or related compounds were unable to inhibit the agglutinations caused by full EULD1A or its domains. On the contrary, mannose or related compounds inhibited the agglutination caused by OrysaEULS2 (Table 2C.1) but not by PhypaEULS3 (Fouquaert and Van Damme 2012). Furthermore, galactose related sugars or galactose containing glycoproteins inhibited the agglutination caused by both lectins. These results clearly indicate that the carbohydrate-binding properties of OrysaEULD1A are very different from these of OrysaEULS2 and similar to PhypaEULS3.

The promiscuity of the EUL domain has been reported before in a previous paper where three different proteins containing a single EUL domain were compared for their

carbohydrate-binding properties. Most probably the differences in specificity related to small changes in the amino acid sequences, especially amino acids located in the vicinity of the carbohydrate-binding site. Previously sequence comparison and molecular modeling of the EUL sequence suggested that, the putative carbohydrate binding site of the EUL domain consists of four residues Aspartic acid, Tryptophan, Asparagine and Glutamine (Al Atalah et al., 2012; see chapter 2B). Sequence alignment of the C-terminal sequence of all EUL domains present in the rice genome indicates that the residues predicted to be part of the carbohydrate-binding site are identical for all EUL domains in rice (Fig. 2C.4 in red). Nevertheless it was shown for some of these sequences that they encode EUL proteins with different carbohydrate binding properties. Some subtle amino acid changes can be observed in the sequence alignment, suggesting that these amino acid substitutions might be important. It has been suggested that Tryptophan 35 (Fig. 2C.4), located in the vicinity of the carbohydrate-binding site for OrysaEULS2, might participate in stacking interactions that often reinforce the binding of a simple sugar to the carbohydrate binding site of plant lectins (Al Atalah et al., 2012 see chapter 2B; Van Damme et al., 2007). Furthermore, observations suggest that some loops created during protein folding might influence the carbohydrate binding specificity of the protein by changing the binding site or weakening its reactivity. For instance in the case of OrysaEULS2, there is a loop that covers the Asparagine residue (Fig. 2C.4, N 42) of the binding site, making it impossible for this residue to contribute to the binding of sugar residues (Al Atalah et al., 2012; see chapter 2B). Genome sequence analyses combined with transcriptome analyses provided evidence that through evolution lectin domains have been used as building blocks to create new chimeric proteins with multiple domains and with multiple activities. The question rises why plants evolved proteins with multiple lectin domains. It has been suggested before that plants used domain duplication followed by divergent evolution as a mechanism to generate multispecific lectins to adapt to various environmental stresses (Jiang et al., 2010a). Evidence shows that through evolution domain duplications can result in alteration of the lectin specificity of the individual lectin domains, as previously also shown for the group of legume lectins, jacalin-related lectins and GNA-related lectins (Loris et al., 1998; Raval et al., 2004; Van Damme et al., 2007). For instance the single-domain GNA-related lectins show exclusive specificity towards mannose but domain duplication events also yielded lectins that exhibit a more complex specificity (Van Damme et al., 2007b).

Furthermore it is clear that in the case of rice this plant species accumulates multiple lectin genes. Over the last 5 years at least 8 (Jiang et al., 2010a) types of lectin domains have been reported in rice, most of which can be considered as stress inducible lectins. The jacalin-related lectin Orysata and the EUL-related lectin OrysaEULS2 have both been shown to exhibit specificity towards mannose (Al Atalah et al., 2011; 2012; see chapters 2A and 2B respectively). This report yields evidence that at least OrysaEULD1A is directed towards galactosylated carbohydrates. It can be envisaged that from a plant point of view it is advantageous to possess a set of carbohydrate binding domains with variable specificities, enabling the plant to bind a multitude of carbohydrate structures which are present either as free molecules or attached to glycoproteins inside the plant cell.

```
              1              15                    35      42
S2 domain     FRCIRMVNNIYLNFDAFHGDKYHGGVRDGTDIVLWKWCEGDNQRWKIQPY-Y
S3 domain     FRCIRMVNNIYLNFDALHGDKDHGGVRDGTTLVLWEWCEGDNQRWKIVP--W
D1A domain1   FRCIRMVNNTRLNLDAFHGDKDHGGVRDGTTVVLWEWCKGDNQSWKILP--W
D1A domain2   FRCIRMVNNIYLNFDAFHGDKDHGGIHDGTEIVLWKWCEGDNQRWKILP--W
D1B domain1   FRCIRMLNNIRLNFDAFHGDKDHGGVHDGTTIVLWEWAKGDNQCWKILP--W
D1B domain2   FRCIRMVNNIYLNFDALHGDKDHGGVRDGTTVALWKWCEGDNQRWKIVP--W
D2 domain1    FRCIRMVNNIYLNLDAFHGDKSHGGVHDGTTVVLWEWCKGDNQCWKILP--W
D2 domain2    FRCVRMVNNIYLNFDAFHGDKDHGGVHDGTTVVLWEWCKGDNQRWKILP--W
              ***:**:**  **:**:**** ***:.*** :.**:*.:**** *** *  :
EEA           YSAIRSLTNPASHLEAAPLN---DWSYNGAIIMGGVWIDAYNQQWKIEPHTG
              : .:* :.*    :::*    :   .    :*: :    * .. ** *** *
```

Fig. 2C.4 Sequence alignment of ~50 amino acids located at the C-terminus of EUL domains in rice. Residues expected to form the carbohydrate-binding site are shown in bold and red. Green asterisks, colons and dots indicate the sequence identity and similarity among the EUL domains. Black asterisks, colons and dots show the sequence identity and similarity between the rice EUL domains and EEA

Chapter 3

Promoter and Q PCR analysis of EUL lectins from rice

Chapter 3A

Promoter analysis for three types of EUL-related rice lectins in transgenic *Arabidopsis*

Redrafted from:

Al Atalah B, Fouquaert E, Van Damme EJM (2013) Promoter analysis for three types of EUL-related rice lectins in transgenic *Arabidopsis*. Plant Mol Biol Rep 31: 1315–1324

3A.1 Abstract

In this study, the promoter activity for three types of *Euonymus*-related lectins (EUL) from rice, further referred to as OrysaEULS2, OrysaEULS3 and OrysaEULD1A was analyzed. *In silico* promoter analyses showed that the EUL promoters from rice contain next to the typical promoter elements some motifs that are considered to be stress responsive elements. Furthermore *Arabidopsis thaliana* plants were transformed with a promoter::GUS construct for each of the proteins under study. Subsequently, one-insertion homozygous lines were selected and analysed for GUS activity. Experiments were performed under normal growth conditions or after application of different stress conditions, in particular treatments with 150 mM NaCl, 100 mM mannitol and 100 µM ABA for 24 h. GUS activity was detected with the OrysaEULS3 and OrysaEULD1A promoters especially in the cotyledons and the young true leaves, respectively, but not with the OrysaEULS2 promoter. The activity of OrysaEULS3 and OrysaEULD1A promoters was increased after ABA and mannitol treatments but decreased after NaCl treatment. We hypothesize that the *Euonymus*-related rice proteins have a role in sensing and responding to external stresses as well as in the growth of the plant.

3A.2 Introduction

Since the molecular cloning of the *Euonymus europaeus* lectin (Fouquet et al., 2008), the *Euonymus* lectin (EUL) domain is considered as the prototype for a new lectin family (Fouquaert et al., 2009a; Fouquaert and Van Damme 2012). The EUL sequence is well conserved and can be retrieved from all completely sequenced plant genomes (Fouquaert et al. 2009a). Hence, the EUL domain is widespread within the plant kingdom. Furthermore, expression analyses revealed that EUL domains are present in many stress response proteins suggesting a role of this lectin domain in stress signalling. In rice, several types of EUL proteins have been identified (Table 3A.1). The EUL type S2 (referred to as OrysaEULS2) consists of a single EUL domain preceded by medium long unrelated N-terminal domain. The EUL type S3 (OrysaEULS3) resembles the type S2 but contains an N-terminal domain with a different and longer sequence. The EUL type D1 from rice includes two almost identical proteins: OrysaEULD1A and OrysaEULD1B (Fig. 3A.1). Both D1 proteins are composed of two EUL domains separated by a short linker and preceded by an unrelated N-terminal sequence. Similarly the EUL type D2 (OrysaEULD2) contains two EUL domains but the linker sequence is longer and the N-terminal domain is different (Fouquaert et al., 2009a) (Fig. 3A.1). OrysaEUL proteins have been described previously as OSR40 (*Oryza sativa* repeats 40 kDa) proteins responsive towards salt and ABA treatments in shoot and root tissues of rice seedlings (Moons et al. 1997a, Kawasaki et al., 2001) as well as in the panicles of the adult plant (Dooki et al., 2006).

Abiotic stresses such as salinity and drought are considered as a serious threat to agricultural production (Boyer 1982; Mittler 2006), decreasing the average yields for most major crops by more than 50% (Bray et al., 2000). More than 20% of mass land and almost 50% of the irrigated lands are affected by salinity (Bray et al., 2000). Rice is one of the most important crops providing food for more than half of the world population, but unfortunately its yield is decreased 50% under moderate salt conditions (Zeng et al., 2002). Furthermore, drought is the major environmental constraint reducing the productivity of rice in rainfed areas (Farooq et al., 2009; Serraj et al., 2009). Because the world population will reach 9 billion in 2050 the demand for food production is increasing (Godfray et al., 2010; Tester and Langridge 2010) but at the same time the arable land is decreasing due to severe environmental changes. It is therefore a big challenge to satisfy this huge demand for food

and feed, and thus researchers are attempting to understand the physiological response of rice (and other crops) towards stress conditions in order to enhance the tolerance of these crops towards biotic and abiotic stresses, and as a consequence increase their yield.

Recently the rice EUL proteins OrysaEULS2, OrysaEULS3 and OrysaEULD1A have been expressed in *Pichia pastoris* and purified (Al Atalah et al., 2012 see chapter 2B, unpublished data). It was shown that these EUL proteins are true lectins and can interact with different carbohydrate structures. The *Arabidopsis* system has been extensively used as a model to investigate the promoter activity mainly because of the ease of transformation, growth and analysis. Several promoters from rice have been successfully investigated in *Arabidopsis thaliana* (Shiyou et al., 2007; Ya-Qin et al., 2007; Yang et al., 2012). In this work, the promoter activity for OrysaEULS2, OrysaEULS3, and OrysaEULD1A was investigated under normal plant growth conditions and after stress treatment of *Arabidopsis thaliana* harboring a promoter::β-glucuronidase (GUS) construct. Our data show differential expression of the three promoter sequences and contribute to a better understanding of the role of nucleocytoplasmic plant lectins in stress tolerance.

Table 3A.1 Overview of EUL-related lectins from rice

Schematic representation	Lectin term[a]	Old term[b]	Accession numbers of target genes		
			Locus ID[c]	MSU ID[d]	NCBI ID[e]
	OrysaEULS2	OSR40g3	Os07g0684000	LOC_Os07g48500	AK072989
	OrysaEULS3	r40c1	Os01g0104400	LOC_Os01g01450	CB681539
	OrysaEULD1A	OSR40g2	Os07g0683900	LOC_Os07g48490	AK103324
	OrysaEULD1B	OSR40c1	Os03g0327600	LOC_Os03g21040	AK069815
	OrysaEULD2	OSR40c2	Os07g0683600	LOC_Os07g48460	OSJNBa0060 O17.11

: EUL domain. , , , : Unrelated N-terminal domains. : Inter-domain linker. [a]: (Fouquaert et al., 2009a). [b]: (Moons et al., 1997a). [c]: The Rice Annotation Project Database (Tanaka et al., 2008). [d]: Rice Genome Annotation Project (Ouyang et al., 2007). [e]: National Center for Biotechnology Information.

```
OrysaEULS2   MDFYGRREQYGGYGGYGGGGALATPGYAPAAPYGMSQVSIEGNGCGRTLPPQP------- 53
OrysaEULS3   MEFPHRHHHHGHRGDDDDDDRRRHPAPAPAYGHDSAPPPGPYGQAPPPADPYARHPPSHD 60
OrysaEULD1A  MFGFGHHHNQAPAAPSDPN---------QIFKIFCRANENYCLTVRDSAVVLAPVNPKDE 51
OrysaEULD1B  MFGFGHHGHHGQDQPPQHHGGGGGGAHQPTFKIFCRADEGYCVAVREGNVVLAPTNPRDE 60
OrysaEULD2   MFSHHGHGHGQYQPPATGPQHE------PTFKIFCRADEGYCLTVRHDAVVLAPTNPRDD 54
                                         *******:*.**::**.. *****.**:*:

OrysaEULS2   ------------------------------------------------------------
OrysaEULS3   YAHPPPAYGGGGYGNVVHVSHEVSDHQRPTPHYGGSEYISPVQETRPYHGGGGAPPV--- 117
OrysaEULD1A  HQHWFKDMRFSTKVKDGEGMPAFALVNKATGLAVKHSLGQSHPVKLVPFNPEYEDASVLW 111
OrysaEULD1B  HQHWYKDMRFSAKIKDEEGNPAFALVNKATGLAIKHSLGQGHPVKLAPFNPEYPDESVLW 120
OrysaEULD2   CQHWYKDMRHSTRVKDEEGHPAFALVNRATGLAVKHSLGQSHPVKLVPYNPEYQDESVLW 114
             ***:****.*::.** ** *******:****:******.*****.*:**** * ****

OrysaEULS2   ------------------------------------------------------------
OrysaEULS3   ------------------------------------------------------------
OrysaEULD1A  TESKDVGKGFRCIRMVNNTRLNLDAFHGDKDHGGVRDGTTVVLWEWCKGDNQSWKILPWG 171
OrysaEULD1B  TESGDVGKSFRCIRMLNNIRLNFDAFHGDKDHGGVHDGTTIVLWEWAKGDNQCWKILPWG 180
OrysaEULD2   TESKDVGHGFRCIRMVNNIYLNLDAFHGDKSHGGVHDGTTVVLWEWCKGDNQCWKILPWG 174
             *** ***:.******:** **:****** .****:****:***** ***** ******

OrysaEULS2   ------------------------------------------------------------
OrysaEULS3   ------------------------------------------------------------
OrysaEULD1A  PEAHSSSPGAATACTIGGVPVH-------------------------------------- 193
OrysaEULD1B  DEAYAGGSANAPRGGNEP------------------------------------------ 198
OrysaEULD2   PEAYAPPPPPAYGHQAYPPPPPNREPGHGYHPAPAFYPPQPPPSHDEPGYGYRPPPVGPP 234

OrysaEULS2   ----------------TVKVYCRANPN-YAMTARNGAVVLAPANPKDEYQHWIKDMRWST 96
OrysaEULS3   ----------------THRIYCKAGEDNYSLAVRDGKVCLVRSDRDDHTQHWVKDMKYST 161
OrysaEULD1A  ----------------TVRVFSAAGED-YCLTVRNGTACLAPKNPRDDYQHWIKDMRHSN 236
OrysaEULD1B  ----------------TVRIFCKADEG-FSVTVRGGSVCLAPTNPRDEYQHWIKDMRHSN 241
OrysaEULD2   GAGYGNRLPRALASEPTVRILCRADEA-YSLTVRNGAVCLAPTNPRDDFQHWVKDMRHST 293
                              * :: . *.   :.::.*.* . *.  :  *. ***:***: *.

OrysaEULS2   SIKDEEGYPAFALVNKATGQAIKHSLGQSHPVRLVPYNPEVMDESVLWTESRDVGNGFRC 157
OrysaEULS3   RVKDEEGYPAMALVNKATGDALKHSIGQSHPVRLVRYNPEYMDESVLWTESRDVGSGFRC 221
OrysaEULD1A  KIRDEEGYPAFALVNKVTGEAIKHSTGQHPVKLVPYNPEYQDESVLWTESKDVGKGFRC 296
OrysaEULD1B  SIKDEEGYPAFALVNRVTGEAIKHSQGEGHPVKLVPYNPGYQDESVLWTESRDVGHGFRC 301
OrysaEULD2   SIKDEEGYPAFALVNKATGEAIKHSLGQSHPVRLVPYNPEYLDESVLWTESKDVGHGFRC 353
              ::*******:****:.**:*:*** *:.***:** ***    *********:*** ****

OrysaEULS2   IRMVNNIYLNFDAFHGDKYHGGVRDGTDIVLWKWCEGDNQRWKIQPYY 204
OrysaEULS3   IRMVNNIYLNFDALHGDKDHGGVRDGTTLVLWEWCEGDNQRWKIVPW- 268
OrysaEULD1A  IRMVNNIYLNFDAFHGDKDHGGIHDGTEIVLWKWCEGDNQRWKILPW- 343
OrysaEULD1B  IRMVNNIYLNFDALHGDKDHGGVRDGTTVALWKWCEGDNQRWKIVPW- 348
OrysaEULD2   VRMVNNIYLNFDAFHGDKDHGGVHDGTTVVLWEWCKGDNQRWKILPW- 400
             :*************:**** ***::*** :.**:**:******** *:
```

Fig. 3A.1 Amino acid sequence alignment of rice EUL proteins. This alignment was performed only for the EUL domains. Identical amino acids within the EUL domains are indicated by asterisks and similar residues by dots or colons. EUL domains are shaded in yellow. N-terminal sequences are shaded in grey. Inter-domain linkers are shaded in green

3A.3 Materials and methods

3A.3.1 Plant material and growth conditions

Arabidopsis thaliana ecotype Columbia (Col-O) (Lehle Seeds, Round Rock, Texas, USA) was used in this investigation. To set up the *in vitro* culture, dry seeds were surface sterilized in 70% (v/v) ethanol for 2 min and afterwards for 10 min in 5% (v/v) NaOCl. After rinsing the seeds thoroughly with sterile distilled water they were germinated on a filter paper put on top of solid Murashige and Skoog (1962) medium (MS) (4.3 g/l MS micro and macro

nutrients containing vitamins (Duchefa, Haarlem, The Netherlands), 30 g/l sucrose, pH 5.7 (adjusted with 0.5 M NaOH) and 8 g/l plant agar (Duchefa)). To synchronize the germination, *Arabidopsis* seeds were kept in the dark at 4°C for 3 days. Afterwards, all plates were moved to a growth chamber at 22°C with a 16/8 h light/dark photoperiod. Three weeks later, *Arabidopsis* plantlets were transferred to artificial soil (Jiffy-7, columnar diameter 44 mm, AS Jiffy Products, Drobak, Norway) and allowed to grow into adult plants.

3A.3.2 Preparation of Promoter::GUS constructs

Genomic DNA was prepared from the shoots of 5-day old rice seedlings using the protocol described by Lannoo et al. (2006b). Promoter sequences for OrysaEULS2, OrysaEULS3 and OrysaEULD1A were obtained from the Osiris promoter database (Morris et al., 2008). To amplify these promoter sequences from genomic DNA, two nested PCR reactions were performed using different sets of primers (Table 3A.2). The non-complete *attB* sequences were completed by an additional PCR reaction using the primers EVD 2 and EVD 4 (Table 3A.2). In order to investigate these promoter sequences, they were cloned into the destination vector: pKGWFS7.0 (Karimi et al., 2002) following the Gateway Cloning technology (Invitrogen, USA) by conducting two consecutive reactions referred to as the BP and LR reaction, as described before (Al Atalah et al., 2011; see chapter 2A). As such, the GUS gene was cloned under the control of each of the promoters under study. Promoter::GUS constructs were introduced in *Escherichia coli* top10F cells using heat shock transformation. All constructs were sequenced after the BP reaction using the primers Donr-F and Seql-E (Table 3A.2) (achieved by LGC Genomics, Berlin, Germany).

Table 3A.2 Primer sequences used in this study

Promoter construct/PCR[a]	Forward primer/ name and sequence (5' to 3')	Reverse primer/ name and sequence (5' to 3')
OrysaEULS3/1	EVD570/CTTTGTCTCTCTTATCTTCTCCC	EVD572/ATGAGGGAACTCCATGATGGA
OrysaEULS3/2	EVD571/AAAAAGCAGGCTTCTCTTCTCCCTTTTGCTATCTCTA	EVD573/AGAAAGCTGGGTGCATGATGGATGGATTTGGGGG
OrysaEULS2/1	EVD659/CACATGCCCACAAACGAGCCATAG	EVD661/TACTGCTCGCGCCGCCCGTAAAA
OrysaEULS2/2	EVD660/AAAAAGCAGGCTTCAAACGAGCCATAGTGCCCGTCCG	EVD/662AGAAAGCTGGGTGTCGCGCCGCCCGTAAAAGTCCAT
OrysaEULD1A/1	EVD663/CTGCTCCAACGTCCACAAACTCCA	EVD665/TGCCTGGTTGTGGTGGTGCCCGA
OrysaEULD1A/2	EVD664/AAAAAGCAGGCTTCACAAACTCCACAGGACAACCTA	EVD666/AGAAAGCTGGGTGTGGTGGTGCCCGAACCCAAACAT
Completing the *attB* sequence	EVD2/GGGGACAAGTTTGTACAAAAAAGCAGGCT	EVD4/ACCACTTTGCTCAAGAAAGCTGGGT
Sequencing primers	Donr-F/TCGCGTTAACGCTAGCATG	Seql-E/GTTGAATATGGCTCATAACAC

a: 1: first nested PCR, 2: second nested PCR

3A.3.3 *Arabidopsis* transformation with promoter::GUS constructs

The binary expression vector was extracted from *E. coli* top10F cells after LR reaction and transformed into the *Agrobacterium* strain GV3101 by electroporation using the following pulse conditions: 2 kV, 25 μF and 200 Ω using the Gene Pulser[R] (BioRad, Hercules, CA, USA). Several colonies were tested after selection on YEB (consisting of 5 g/l beef extract, 5 g/l peptone, 5 g/l sucrose, 1 g/l yeast extract and 15 g/l bacteriological agar) medium containing spectinomycin (50 μg/ml). One positive colony was chosen to transform 6-week old *Arabidopsis* plants using the floral dip transformation method (Clough and Bent 1998). T0 seeds were selected on MS medium containing kanamycin (50 μg/ml) following the procedure described by Harrison et al. (2006) and afterwards several independent lines were tested at DNA level. To select for one-insertion lines, a protocol adapted from Jain et al. (2006) and Giri et al. (2011) was used. Briefly, 100 T1 seeds were sown on MS medium containing kanamycin and then lines which met the ratio 3:1 (survival:dead) were considered as one-insertion lines. T2 seeds were harvested individually for 10 sub-lines per construct and sown on MS containing kanamycin. T3 seeds were harvested from the sub-lines which showed 100% survival and were considered as homozygous. Afterwards, seeds

from homozygous sub-lines were pooled for each construct. In this investigation, T3 seeds (homozygous one-insertion lines) were used to investigate the activity of OrysaEULS2, OrysaEULS3 and OrysaEULD1A promoters.

3A.3.4 Stress treatments

Five, ten and fifteen-day old *Arabidopsis* seedlings expressing the GUS gene under the control of one of the EUL promoters were exposed to salt (150 mM), mannitol (100 mM) and ABA (100 µM) by transferring the filter paper on which the plantlets were grown to new plates containing several filter papers immersed in MS with the indicated stress factors. Seedlings were kept on this medium for 24 h. GUS staining was analysed and compared for seedlings grown on MS plates and seedlings subjected to stress treatments. Three independent lines for each construct were used in this analysis. Seedlings expressing the GUS gene under the 35S cauliflower mosaic virus promoter were kept on the MS medium and served as a positive control.

3A.3.5 Qualitative GUS assay

To perform the histochemical staining of β-glucuronidase, a protocol adapted from Jefferson (1987) as described by Delporte et al. (2011) was followed. GUS assays were performed during different stages of plant development under different growth conditions. Three seedlings were analysed for each treatment. The staining procedure was performed on the whole seedling in the first three stages of development and on detached parts (roots, leaves, stems, flowers and siliques) for developmental stages 4 and 5.

3A.3.6 Bioinformatics analysis

DNA fragments containing approximately 1900 bp upstream from the start codon were considered to have the near full length promoter sequence for each gene under study. To investigate the presence of putative stress regulatory cis-elements, all promoter sequences were analysed using the publicly accessible database PLACE (Higo et al., 1999). The plant promoter database (Shahmuradov et al., 2003) was used to check for the typical promoter elements including CAAT, TATA boxes and transcription start site (TSS). In this database, different nucleotide matrices were built from several species including rice and by comparing these matrices with our promoter sequences, the typical promoter elements

were determined. Using the Osiris promoter database (Morris et al., 2008), the 5'
untranslated region was determined for each promoter sequence under study. The
nucleotide A in the start codon (ATG) is designated as +1.

3A.4 Results

3A.4.1 *In silico* promoter sequence analysis

All promoter sequences under study were searched for the typical promoter elements and
the TSS (Suppl. File 3A.1). The TSS in the promoter sequences of OrysaEULS2, OrysaEULS3
and OrysaEULD1A is located at positions -107, -1075 and -76, respectively. Several TATA and
CAAT boxes were identified in each promoter sequence but the boxes closest to the TSS and
not located in the 5' untranslated region were considered as active elements. The TATA box
is located at positions -127/-132, -1198/-1203 and -104/-109 in the OrysaEULS2, OrysaEULS3
and OrysaEULD1A promoter sequences, respectively. The CAAT box is detected at positions -
144/-147, -1179/-1182 and -379/-382 in the promoter sequences of OrysaEULS2,
OrysaEULS3 and OrysaEULD1A, respectively. Since all promoter sequences contain the
typical promoter elements, they are expected to represent active promoters.

The presence of stress regulatory elements in the promoter sequences of OrysaEULS2,
OrysaEULS3 and OrysaEULD1A was analyzed using the PLACE database. Eleven groups of
putative stress elements were identified (Suppl. File 3A.2). Nine stress elements were
detected in all three promoter sequences though with different frequencies. Based on the
frequency of occurrence they were divided into three classes. The first class with high
frequency elements contains only light responsive elements, the second class groups
medium frequency elements (13> freq >7) including ABA and gibberellin responsive
elements, and the third class contains low frequency elements (freq 1-5) including salt,
drought, cold, freezing, pathogen and anaerobic responsive elements (Fig. 3A.2). Auxin and
brassinosteroid responsive elements were found only in the OrysaEULS3 promoter sequence
at a low frequency (freq 2) (Fig. 3A.2). The frequency of occurrence for each stress element
was modified by subtracting the value reflecting that this motif can occur only by chance.
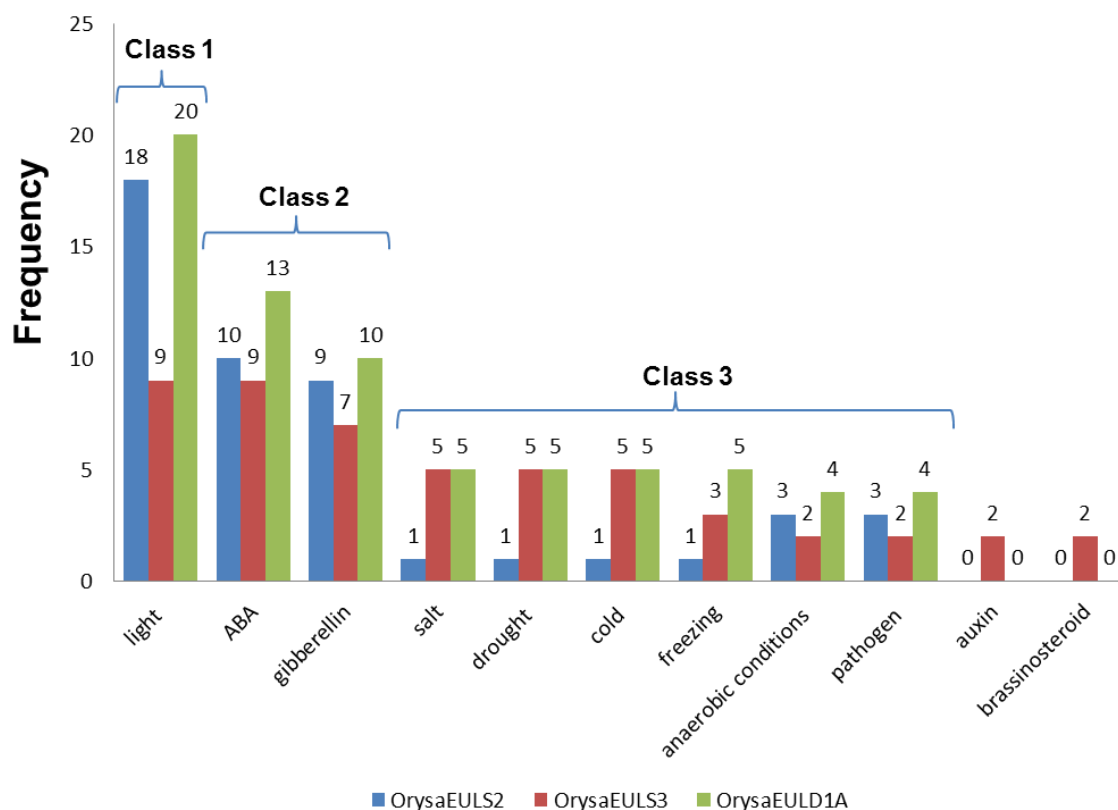
Fig. 3A.2 Chart showing the frequencies of stress elements in the promoter region of the lectins under study

3A.4.2 GUS Expression patterns for EUL promoters in *Arabidopsis*

Transgenic *Arabidopsis* lines expressing promoter::GUS fusion constructs for the promoter sequences from OrysaEULS2, OrysaEULS3 and OrysaEULD1A were generated and a comparative analysis of the promoter activity for each of the EUL genes was made using GUS activity staining of plant tissues. Histochemical staining of the different *Arabidopsis* transgenic lines yielded different GUS staining patterns. Five developmental stages were analyzed starting with 1) 5 day old (excluding the vernalization period) seedlings having the fully open cotyledons and the emergence of the first 2 true leaves, 2) 10 day old seedlings with 6 true leaves, 3) 15 day old plants with 8 true leaves, 4) 32 day old adult flowering plants and 5) 43 day old adult plants harboring pre-mature and some dry siliques. The promoter GUS activity was examined under normal growth conditions (for all five developmental stages) as well as under different stress conditions (during the first three plant stages), including 150 mM salt, 100 mM mannitol, and 100 µM ABA treatments. These stress conditions were selected taken into account previous reports showing that the rice

EUL proteins are responsive to salt and ABA (Moons et al., 1997a; Kawasaki et al., 2001), as well as the occurrence of putative stress response elements for ABA, salt and drought in the promoter sequences. Seedlings harboring a GUS construct under the 35S promoter were used as a positive control during all experiments at several developmental stages of the plant (Suppl. Fig. 3A.1).

3A.4.2.1 *OrysaEULS2 promoter activity*

In spite of the presence of the typical promoter elements in the OrysaEULS2 promoter sequence (Suppl. File 3A.1), no GUS expression was observed for the OrysaEULS2::GUS promoter construct neither under normal growth conditions nor under any of the stress treatments (Suppl. Fig. 3A.2).

3A.4.2.2 *OrysaEULS3 promoter activity*

Different GUS staining patterns for the OrysaEULS3 promoter were observed during development of *A. thaliana* transgenic lines. In plant **stage 1** a strong GUS staining was present in the hypocotyl and the petioles of the cotyledons of *Arabidopsis* seedlings grown under normal conditions. GUS activity was also seen in the cotyledons as well as in the trichomes of the true leaves but not in the leaf tissue (Fig. 3A.3a_1). OrysaEULS3 promoter activity was similar in seedlings subjected to stress treatments, but the GUS staining was reduced especially in response to NaCl treatment (Fig. 3A.3d_1). The GUS staining of trichomes was not apparent after ABA and NaCl treatment (Fig. 3A.3c_1 and 3d_1 respectively) but was still present after mannitol treatment of the seedlings (Fig. 3A.3b_1).

In plant **stage 2**, GUS activity for the OrysaEULS3 promoter was clearly reduced in the hypocotyl, the petioles of the cotyledons of the control plants. GUS staining was observed in the trichomes as well as on the attachment site of the trichomes on the leaves. Low GUS staining was present in the laminae of the true leaves (Fig. 3A.3a_2). OrysaEULS3 promoter activity was clearly observed in seedlings subjected to mannitol and ABA treatments, but was almost absent in NaCl treated seedlings (Fig. 3A.3b_2, c_2, d_2).

Plantlets of **stage 3** grown under normal conditions showed OrysaEULS3 promoter activity in the hypocotyl, the petioles of oldest true leaves, at the base of the youngest leaves but not

in the cotyledons and the trichomes (Fig. 3A.3a$_3$). All stress treated plants showed lower GUS activity of OrysaEULS3 promoter compared to the control plants (Fig. 3A.3b$_3$, c$_3$, d$_3$).

Plant **stages 4 and 5** showed GUS activity for the OrysaEULS3 promoter in the flowers and the pre-mature siliques, but not in the mature siliques or the seeds (Fig. 3A.3a$_{4,5}$). At this stage no staining was observed in the leaves, stems or roots (Suppl. Fig. 3A.2b).

OrysaEULS3 promoter activity was never observed in root tissues for none of the plant stages and treatments tested.
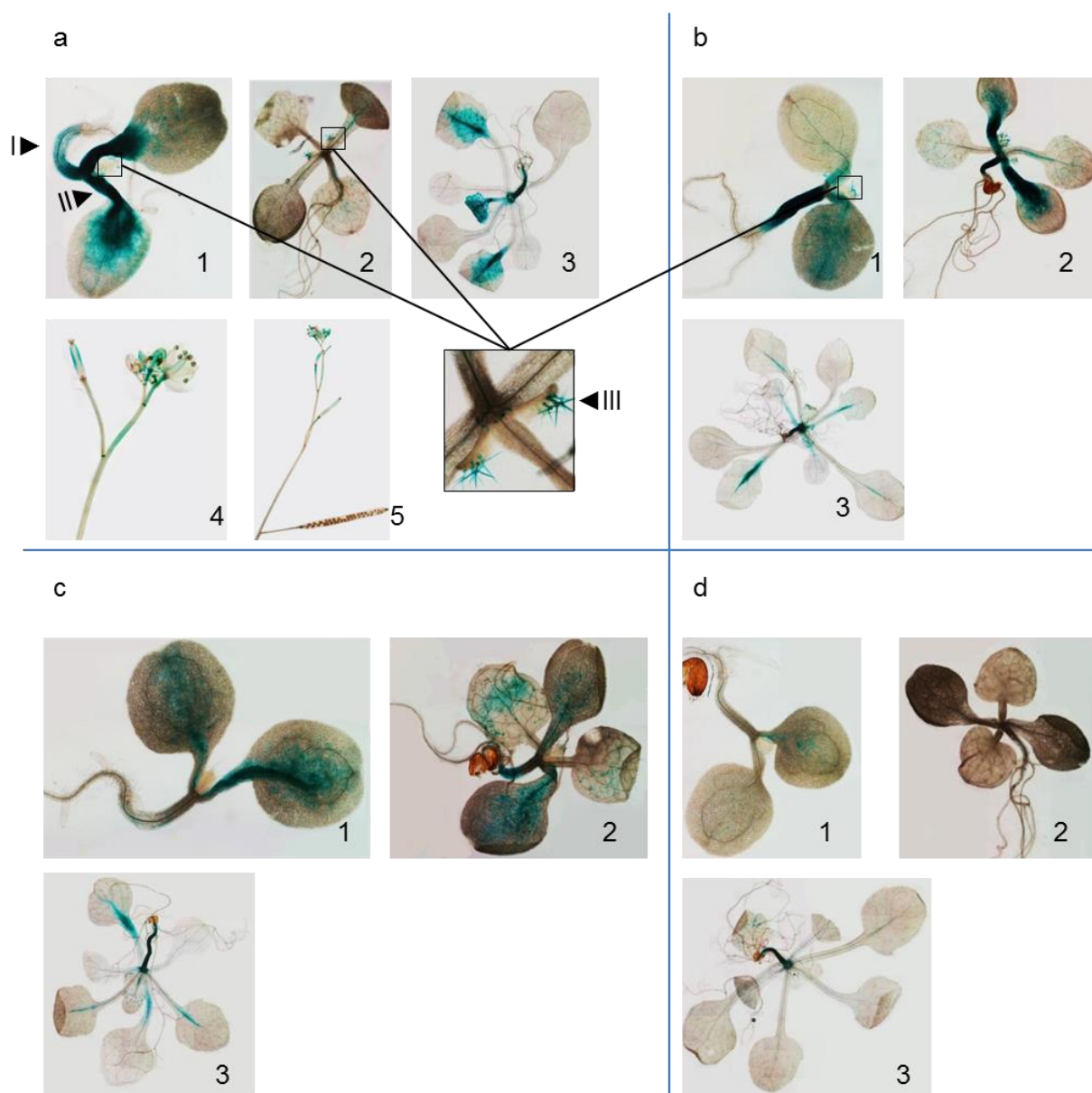

Fig. 3A.3 Legend on next page

Fig. 3A.3 Histochemical analysis of the GUS reporter gene under the control of OrysaEULS3 promoter in *Arabidopsis* plants. GUS expression was analyzed under normal growth conditions (MS) (a) and after stress (100 mM mannitol (b), 100 µM ABA (c) and 150 mM NaCl (d)) treatment for 24 h. Pictures represent one transgenic line during different developmental stages, but are representative for all transgenic lines tested. Developmental stages: 1, stage 1, 5-day old seedling; 2, stage 2, 10-day old seedling; 3, stage 3, 15-day old seedling; 4, stage 4, flowering plant and 5, stage 5, flowering plant with immature and mature siliques. I►: hypocotyl, II►: petiole, III►trichomes. Images were taken using a Nikon eclipse TE2000-e epi-fluorescence Microscope (Nikon Benelux, Brussels, Belgium)

3A.4.2.3 *OrysaEULD1A promoter activity*

OrysaEULD1A::GUS promoter activity showed different GUS staining patterns compared to those observed for the OrysaEULS3::GUS construct. In plant **stage 1**, GUS staining was detected in the first 2 emerging true leaves under normal growth conditions as well as under all stress treatments tested (Fig. 3A.4a_1, b_1, c_1, d_1). Unlike for the OrysaEULS3::GUS construct, no OrysaEULD1A::GUS expression was observed in the cotyledons, nor in the hypocotyl or in the trichomes. The intensity of the GUS staining was increased after mannitol and ABA treatments and decreased after NaCl treatment (Fig. 3A.4b_1, c_1, d_1).

Similar to plant stage 1, GUS activity was confined to the true leaves in plant **stage 2**, but no staining was present neither in the cotyledons, nor in the hypocotyl or the trichomes (Fig. 3A.4a_2, b_2, c_2, d_2). There was no clear difference between the GUS staining patterns observed for the plants grown under normal conditions and the plants subjected to ABA (Fig. 3A.4a_2 and c_2 respectively). Although the GUS staining was similar for plantlets grown on MS medium and plants subjected to mannitol treatment (Fig. 3A.4a_2 and b_2), the GUS staining was slightly reduced in response to NaCl treatment (Fig. 3A.4d_2). GUS activity for the OrysaEULD1A promoter in root tissue (in the tip of the lateral roots) was only observed after plants were subjected to mannitol treatment (Fig. 3A.4b_2).

In plant **stage 3**, the GUS expression was very low and confined to the youngest true leaves. Similar results were observed for all growth conditions (Fig. 3A.4a_3, b_3, c_3, d_3).

During plant **stages 4 and 5** GUS activity was observed in the flowers and the pre-mature siliques, but not in the mature siliques or the seeds (Fig. 3A.4$a_{4,5}$), nor in leaf, stem or root tissues (Suppl. Fig. 3A.2b).
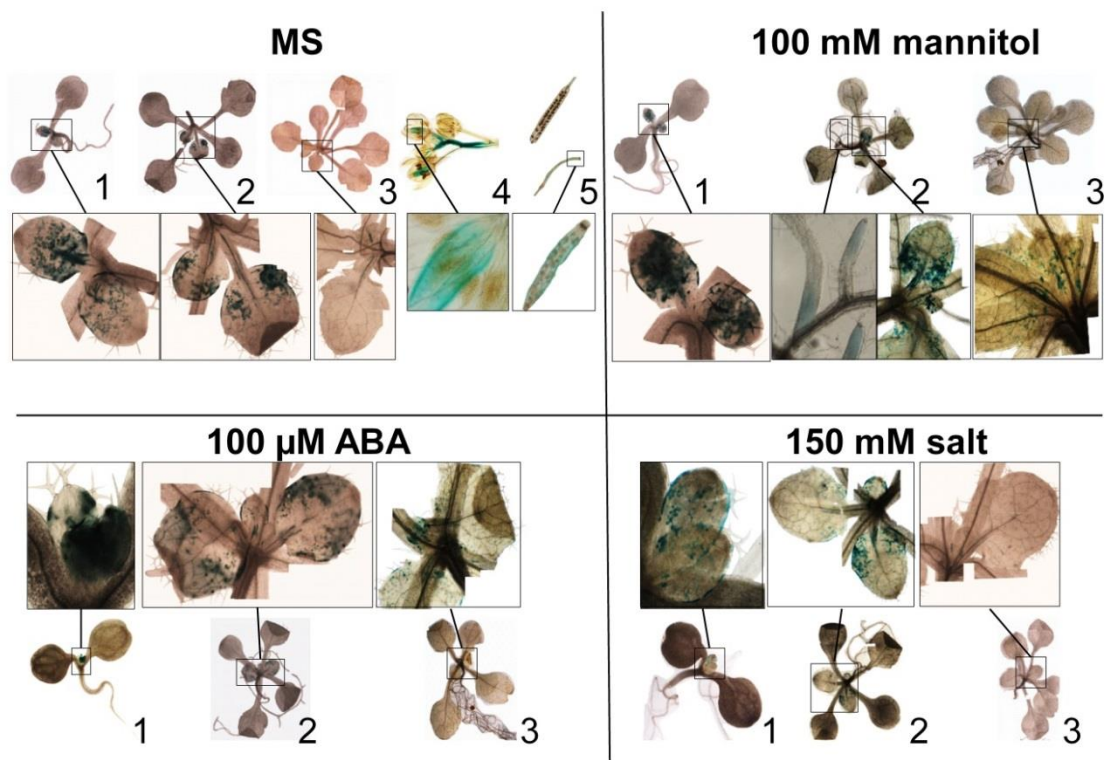
Fig. 3A.4 Histochemical analysis of GUS activity under the control of OrysaEULD1A promoter in *Arabidopsis* plants. GUS expression was analyzed under normal growth conditions (MS) and after 24 h stress (100 mM mannitol, 100 µM ABA and 150 mM NaCl) treatment. Pictures represent one transgenic line during different developmental stages, as in figure 3A.3. Images were taken by a Leica microscope (Leica, Nussloch, Germany)

3A.5 Discussion

During the last decade several inducible plant lectins have been reported (Zhang et al., 2000; Chen et al., 2002; Lannoo et al., 2006a; Imanishi et al., 1997; Nakagawa et al., 2000; Yong et al., 2003). Most of these lectins were not detected in plant tissues grown under normal conditions, but lectin expression was increased considerably after the application of some stress factor. Nevertheless the inducible lectins are still expressed at low concentrations in response to abiotic and biotic stresses. Therefore, the hypothesis has been put forward that these lectins might play an important role in plant stress physiology (Van Damme et al., 2008). In an attempt to gain more knowledge and understanding how EUL lectin expression levels change in response to external stresses, the promoter activity for some EUL lectins from rice was analyzed in this study.

In silico analysis revealed that the 1900 bp sequences preceding the OrysaEULS2, OrysaEULS3 and OrysaEULD1A coding sequences contain multiple typical promoter elements

(Suppl. File 3A.1) and thus these promoter sequences were considered to be functional. Analysis of these promoter sequences for putative stress regulatory elements, allowed the identification of 11 stress elements, nine of which were present in all promoter sequences under study (Suppl. File 3A.2). These nine stress elements were further subdivided into 3 classes based on the frequency of their occurrence in the promoter sequence (Fig. 3A.2). Based on these observations, several stress treatments were selected for this study: ABA related stress elements (class 2) were present at medium frequency, whereas putative elements previously attributed to NaCl and mannitol responses (class 3) were found with low frequency. ABA, salt and mannitol treatments have been studied intensively in plants since they are known to be important for salinity and drought tolerance of plants. Increased ABA levels have been shown to regulate plant adaptation to environmental stresses (like drought and salinity) (Zhu 2002; Nambara and Marion-Poll 2005; Cutler et al., 2010; Hirayama and Shinozaki 2010). In addition, both salt and drought stresses are considered as serious problems for crop production, and therefore are of considerable interest for plant research aiming at developing plants that can cope with these stress factors.

To study the effect of the selected stress treatments on the promoter activity of the EUL lectins, transgenic *Arabidopsis* lines harboring the promoter::GUS constructs were generated for the different EUL lectins. Subsequently, histochemical GUS staining allowed analyzing promoter activity during development of plants grown under normal growth conditions as well as in plants treated with ABA, salt or mannitol. Although the OrysaEULS2 promoter sequence contained all the promoter elements necessary for a functional promoter, no expression was observed for the OrysaEULS2 promoter::GUS fusion construct under any of the conditions tested. This result is in a good agreement with the very low frequency (freq 1) of salt and drought elements in the OrysaEULS2 promoter (Fig. 3A.2). In contrast the OrysaEULS3 and OrysaEULD1A promoter GUS fusion constructs showed clear but different GUS staining patterns. When plants were grown under non-stressed conditions in MS medium, GUS staining directed by the OrysaEULS3 promoter was clear in the seedlings, in particular the hypocotyl, the petioles, the cotyledons, the true leaves and the trichomes showed a clear blue staining (Fig. 3A.3). In contrast, GUS activity controlled by the OrysaEULD1A promoter was confined to the young true leaves (Fig. 3A.4). Since the GUS staining analysis was performed qualitatively and the activity of the promoters under study

was weak, it is hard to draw conclusions about their activity. However, some differences were observed between plants grown under stress treatments and the control. After NaCl treatment a fainter blue staining was seen in the plants harboring the OrysaEULS3 promoter GUS as well as the OrysaEULD1A GUS constructs. A reduced GUS activity was observed especially in plant stages 1 and 2. After mannitol and ABA treatments, the promoter activity was increased in plant stage 1 for OrysaEULD1A. OrysaEULS3 promoter activity was increased in stage 2 after mannitol and ABA treatments especially in the petioles and the lamina of the cotyledons as well as in the hypocotyl. The activity of OrysaEULS3 and OrysaEULD1A promoter in young tissues in particular the cotyledons and the true leaves, respectively, suggested a possible role in expansion and cell division during the early stages of tissue development. Both the OrysaEULS3 and OrysaEULD1A promoter sequences showed similar GUS staining of the flower tissues and immature siliques but not in the mature siliques nor in the seeds (Fig. 3A.3a$_{4,5}$, 4a$_{4,5}$) pointing towards a possible role for OrysaEULS3 and OrysaEULD1A in flower and silique development.

Although the dicot species *Arabidopsis thaliana* is often used as a model plant for promoter::GUS studies, mainly because of the ease of transformation, growth and analysis, it is known that promoter elements from monocot species might not be recognized. For instance, it was shown that the rice YY2 promoter shows different GUS staining patterns in *Arabidopsis* and rice plants, in that the complete anthers showed GUS activity in *Arabidopsis* plants whereas the GUS staining was confined to the tapetum tissue in rice plants (Kuriakose et al., 2009; Khurana et al., 2013). Therefore we cannot exclude that the promoter activity of these GUS constructs transformed in rice might yield additional data.

We hypothesize that EUL proteins fulfill a physiological role in sensing and responding to different stresses as well as in the developmental stages of the plant. Further work is required to investigate this hypothesis, e.g. by overexpressing the EUL proteins and analyzing the effect of these proteins under normal growth as well as under different stress conditions. Similarly, knockout lines in one or more EUL genes can be investigated to study the impact of EUL expression on growth and development of the rice plant.

Chapter 3

Promoter and Q PCR analysis of EUL lectins from rice

Chapter 3B

Responsiveness of EUL-related rice lectins towards important abiotic and biotic stresses

Manuscript submitted

Al Atalah B, De Vleesschauwer D, Xu J, Fouquaert E, Höfte M, Van Damme EJM (2013) Responsiveness of EUL-related rice lectins towards important abiotic and biotic stresses.

3B.1 Abstract

In the rice genome, there are 5 genes encoding putative carbohydrate-binding domains belonging to the *Euonymus* lectin (EUL)-related domain family. In this study, the expression level of all these EUL domains was investigated under different stresses and compared to the expression level of Orysata, a well known stress related lectin. Q-PCR was conducted on rice seedlings exposed to abiotic stresses, particularly 150 mM NaCl, 100 mM mannitol, and 100 µM ABA as well as biotic stresses being *Xanthomonas oryzae* (*Xoo*) and *Magnaporthe oryzae* infections. All EUL proteins were up-regulated after ABA and NaCl treatments in the roots whereas the expression level was less and more variable in the shoots. All abiotic stresses induced Orysata in both tissues except mannitol treatment which did not show an effect in the roots. Proteins from the D type were up-regulated after *Xoo* infection except for OrysaEULD1A which was down-regulated. In addition, some of the EUL proteins were down-regulated by *Magnaporthe*, particularly OrysaEULS3, OrysaEULD1A and OrysaEULD2 indicating that this pathogen is targeting them. Interestingly, Orysata was up-regulated after *Xoo* and down-regulated after *Magnaporthe* infections. Our results show that rice expresses multiple carbohydrate-binding proteins in response to several abiotic and biotic stress conditions. We hypothesize that the *Euonymus*-related proteins fulfill a role in sensing and responding to these stresses.

3B.2 Introduction

One family of inducible lectins that is ubiquitous in plants comprises all proteins related to the *Euonymus europaeus* lectin (EUL) (Fouquaert et al., 2009a). Though the EUL domain is quite well conserved, the carbohydrate-binding domain is promiscuous in that it shows interaction with different carbohydrate structures depending on the species and the type of protein under study (Fouquaert and Van Damme, 2012).

In rice, four different types of EUL proteins have been identified (Fouquaert et al. 2009a; Fouquaert and Van Damme 2012). Both the S2 and the S3 types are composed of a single EUL domain preceded by a variable N-terminal sequence. Only one protein was identified belonging to each of the S types, referred to as OrysaEULS2 for type S2 and OrysaEULS3 for type S3, respectively. In addition two D type proteins consisting of two EUL domains separated by a linker sequence and preceded by an N-terminal sequence were found. The size of the linker and N-terminal domains differs in case of the D1 and D2 types. Two almost identical proteins referred to as OrysaEULD1A and OrysaEULD1B belong to the D1 type, whereas one protein was classified as a D2 type, namely OrysaEULD2 (Fouquaert et al., 2009a; Al Atalah et al., 2013 see chapter 3A).

Rice (*Oryza sativa L*.) is a cereal crop that feeds half of the population on earth and provides (together with wheat and maize) 50% of the total consumption of calories (Maclean et al., 2002). Changes in the global climate increase the temperature and the atmospheric CO_2 levels, and alter the rainfall patterns. Due to these changes, abiotic stress (drought, salt, high temperature…) is considered the most harmful factor for growth and productivity of crops worldwide. Therefore, abiotic stress is conceived the primary cause of crop loss, reducing average yields by more than 50% for most major crops (Bray et al., 2000; Boyer and Westgate 2004; Rodríguez et al., 2005). In addition to abiotic stress,   biotic stresses provoked by fungi, bacteria, viruses, insects,… are also decreasing the yield of crops. For instance, 10 to 15% reduction of rice yield was estimated to be due to diseases caused by bacterial blight (*Xanthomonas oryzae*), blast (*Magnaporthe oryzae*), sheath blight (*Rhizoctonia solani*), and rice tungro bacilliform virus (Dai et al., 2007). Based on these facts, Identifying stress tolerance genes and understanding their functions have become the most urgent tasks in order to increase the yield of crops under unfavorable conditions.

In the present study, the expression profile for the whole set of EUL proteins from rice was analyzed under different abiotic and biotic stress conditions. To do this, qRT-PCR was performed to measure the relative expression level of the EUL genes after different stress treatments. Quantitative RT-PCR has become a pioneer technology to measure the gene relative expression, because of its rapidity, accuracy and sensitivity (Gingeras et al., 2005; Nolan et al., 2006; Van Guilder et al., 2008). In addition, qRT-PCR is considered the most appropriate method to confirm or confute data generated by large scale microarrays (Wang et al., 2006b). The data for the different EUL sequences under study have been compared to the expression pattern of a previously identified stress related rice lectin belonging to the family of jacalin-related lectins (Claes et al., 1990, Garcia et al., 1998, Zhang et al., 2000). Our findings can contribute in deciphering the possible role of the inducible EUL lectins under stress conditions.

3B.3 Materials and methods

3B.3.1 Plant material

Oryza sativa cv *Nipponbare* were obtained from the National Rice Research Centre (Genetic stock Oryza collection, Stuttgart, USA). To set up the *in vitro* culture, the seeds were de-husked and surface sterilized in 70% (v/v) ethanol for 5 min and subsequently in 5% (v/v) NaOCl for 30 min. Afterwards the seeds were extensively washed with sterile water.

3B.3.2 Abiotic stress assays

Ten seeds were sown in a jar (9 cm diameter) containing 100 ml solid Murashige and Skoog (1962) medium (MS) (4.3 g/l MS micro- and macronutrients containing vitamins (Duchefa, Haarlem, The Netherlands), 30 g/l sucrose, pH 5.7 (adjusted with 0.5 M NaOH), and 8 g/l plant agar (Duchefa)). The jars were kept at 28°C, 16 h light / 8 h dark cycle.

Five ten day old rice seedlings were transferred to a jar with 100 ml liquid MS containing the desired stress factor (150 mM NaCl, 100 mM mannitol or 100 µM ABA). Five seedlings were transferred to MS and used as a negative control. Two jars were used for each treatment. The stress treatment was applied for 6 h under the conditions mentioned above. Afterwards,

the shoots and the roots for 10 plants per treatment were separated and immediately frozen at − 80°C.

3B.3.3 Biotic stress assays

Rice seeds (japonica cultivar Nipponbare) were surface sterilized with 1% NaOCl for two min, rinsed three times with sterile, demineralized water and incubated on wet sterile filter paper for five days at 28° C. In case of *Xanthomonas oryzae* bioassays, germinated seedlings were transferred to sterilized vermiculite supplemented with half strength Hoagland solution (Xu et al., 2013). Two weeks later, the plants were transferred to plastic containers containing modified Hoagland solution (Xu et al., 2013) and grown for another three weeks under growth chamber conditions (28°C, relative humidity: 60%, 12/12 light regime). Seedlings used for *Magnaporthe oryzae* inoculations were grown in perforated plastic trays (23 by 16 by 6 cm) filled with commercial potting soil (Structural; Snebbout, Kaprijke, Belgium) that had been autoclaved twice on alternate days for 21 min.

3B.3.3.1 *Xanthomonas* assay

Xanthomonas oryzae pv. *oryzae* (*Xoo*) strain PXO99 (Philippine race 6) (Song et al., 1995) was cultured on Sucrose Peptone Agar (SPA) medium at 28°C. To perform the assay, a single colony was transferred to liquid SP medium and grown for 48 h at 28°C. Six week old plants were inoculated by clipping the fifth and sixth stage leaves with scissors and dipping them in a solution of *Xoo* cells in water (1 x 10^9 CFU.ml^{-1}). Inoculated plants were kept in a dew chamber (≥ 92% relative humidity; 28°C) for 24 h and thereafter transferred to greenhouse conditions for disease development. Leaf samples were collected 2, 4 and 8 days post-inoculation (dpi) and frozen at -80°C until use.

3B.3.3.2 *Magnaporthe* assay

Inoculations with Mo were performed exactly as described in De Vleesschauwer et al. (2009). Briefly, leaves of 4-week-old plants (5-leaf stage) were evenly sprayed with a *M. oryzae* spore suspension containing 5 x 10^4 conidia.ml^{-1}. Inoculated plants were incubated at high relative humidity (≥ 92% relative humidity; 28°C) for 24 h and, thereafter, returned to growth chamber conditions for disease development. Six days post inoculation, inoculated leaves were found to display many sporulating blast lesions, confirming the effectiveness of

the pathogen inoculation. Leaf samples of mock and pathogen-inoculated plants were collected at 1 and 2 dpi and frozen at -80°C until further use.

3B.3.4 RNA extraction and cDNA synthesis

Total RNA was extracted from rice roots and shoots after each abiotic treatment and from the infected leaves after each biotic treatment using the RNeasy Plant Mini Kit (QIAGEN, Valencia CA, USA) following the manufacturer's instructions. After DNase I treatment (ThermoScientific, Erembodegem, Belgium), the cDNA was synthesized using the M-MLV Reverse Transcriptase Kit (Invitrogen, Carlsbad, CA, USA). Briefly, in a total volume of 26 µl, 2 µg of DNase I treated RNA were mixed with 2 µl of 2 µM oligo-dT and 2 µl of 10 mM dNTP. After incubation for 5 min at 65°C all samples were cooled down immediately. Subsequently the samples were mixed with 8 µl of the 5x first strand buffer and 4 µl of 0.1 M dithiothreitol. After 2 min incubation at 37°C, 2 µl M-MLV reverse transcriptase (200 U/µl) was added. After incubation for 50 min at 37°C all samples were transferred to 75°C for 15 min to stop the reaction. Finally, water was added to each sample to reach the concentration of 20 ng/µl. To check the quality of the cDNA, a standard RT-PCR using a pair of reference gene primers (Suppl. Table 3B.1) was performed and the amplicons were visualized on a 1.5% agarose gel (Invitrogen).

3B.3.5 Quantitative RT-PCR

The SensiMixTMSYBR No-ROX Kit (BIOLINE, London, UK) was used to perform the qRT-PCR reactions. In each reaction, 10 µl of 2 x SensiMix, 1 µl cDNA (20 ng/µl), 1 µl of 10 µM from each primer and 7 µl water were mixed in a total volume of 20 µl. All reactions were performed in the Rotor-Gene 3000 (Corbett Life Science, Qiagen, Venlo, The Netherlands) using Rotor Discs (Qiagen), and the results were generated by the Rotor-Gene 6 software. The thermal profile consisted of 10 min at 95°C as a pre-denaturation step, 45 cycles of 25 sec at 96°C, 25 sec at 60°C, and 20 sec at 72°C. To test the amplicon specificity, a melting curve was generated by increasing the temperature from 72°C to 95°C. The relative expression level of the target genes under stress conditions were compared with the control by using the REST 384 software (Corbett Research, (Pfaffl et al. 2002)). This software allows determining the statistical significance of the results and compared the relative expression between a sample and a control group. Primers for qRT-PCR were designed by primer 3.0

software: http://frodo.wi.mit.edu/ (Suppl. Table 3B.1) to amplify <150 bp amplicons. Three independent biological replicates were used for the abiotic assays as well as the *Xanthomonas* assay whereas two biological replicates were performed for the *Magnaporthe* assay. Three and two technical replicates were used for the abiotic and biotic assays, respectively.

3B.4 Results

The qRT-PCR technology was performed to compare and quantify the relative expression levels for the whole set of EUL proteins from rice under normal growth condition as well as after several abiotic and biotic stresses. The fold change for all treatments is shown in Suppl. Table 3B.2.

3B.4.1 EUL expression patterns in response to abiotic stresses

Ten day old seedlings were exposed for 6h to several abiotic stress conditions including NaCl (150 mM) to mimic salinity stress, mannitol (100 mM) to mimic drought stress and ABA (100 µM). The effect of the different stresses on the expression level of the different rice EULs was analyzed. The data were normalized by including two stably expressed reference genes, namely expressed protein (EXP, LOC_Os03g27010) (Kyndt et al., 2012) and the eukaryotic initiation factor 5C (EIF5C, LOC_Os11g21990.1) (Narsai et al., 2010) (Suppl. Table 3B.1). Orysata, which belongs to the jacalin-related lectin family, was first described as a salt inducible protein (Claes et al., 1990). Since Orysata has been reported to be expressed in response to salt or drought stresses as well as jasmonic acid and ABA treatment (Zhang et al., 2000; De Souza et al., 2003), the expression of Orysata was also included in the abiotic stress experiments as a positive control. Furthermore the expression levels for the EUL lectins and Orysata have been compared.

After ABA treatment, all EUL proteins under study were up-regulated significantly in the root tissues (Fig. 3B.1a), but only the relative expression levels for OrysaEULS2, OrysaEULD1A and OrysaEULD2 were up-regulated significantly in the shoots (Fig. 3B.1b). Furthermore, the expression levels of OrysaEULS3 and OrysaEULD1B in the shoots were not affected by ABA treatment (Fig. 3B.1b, Suppl. Table 3B.2 respectively). Similarly the expression of Orysata

was significantly up-regulated in roots and shoots after ABA stress. Interestingly, all tested rice lectins showed higher relative expression levels in the root compared to the shoot tissues in response to ABA treatment (Fig. 3B.1).

Quantitative analyses of lectin expression in plants subjected to NaCl stress revealed that the relative expression levels for all EUL proteins as well as Orysata were up-regulated significantly in the root tissues (Fig. 3B.1a). In the shoots, similar to the expression pattern for Orysata the expression of all EUL D types (OrysaEULD1A, OrysaEULD1B and OrysaEULD2) was up-regulated significantly (Fig. 3B.1b) whereas the expression of the S type EULs (OrysaEULS2 and OrysaEULS3) showed no response to NaCl (Fig. 3B.1b, Suppl. Table 3B.2).

In contrast to ABA and NaCl, the mannitol treatment provoked less effect on the expression of the rice lectins. The expression pattern of the EUL proteins was not significantly changed neither in the root nor in the shoot tissues (Fig. 3B.1, Suppl. Table 3B.2). Only the expression level of Orysata was up-regulated significantly in the shoot tissues (Fig. 3B.1b) whereas it was not regulated in the root tissues.
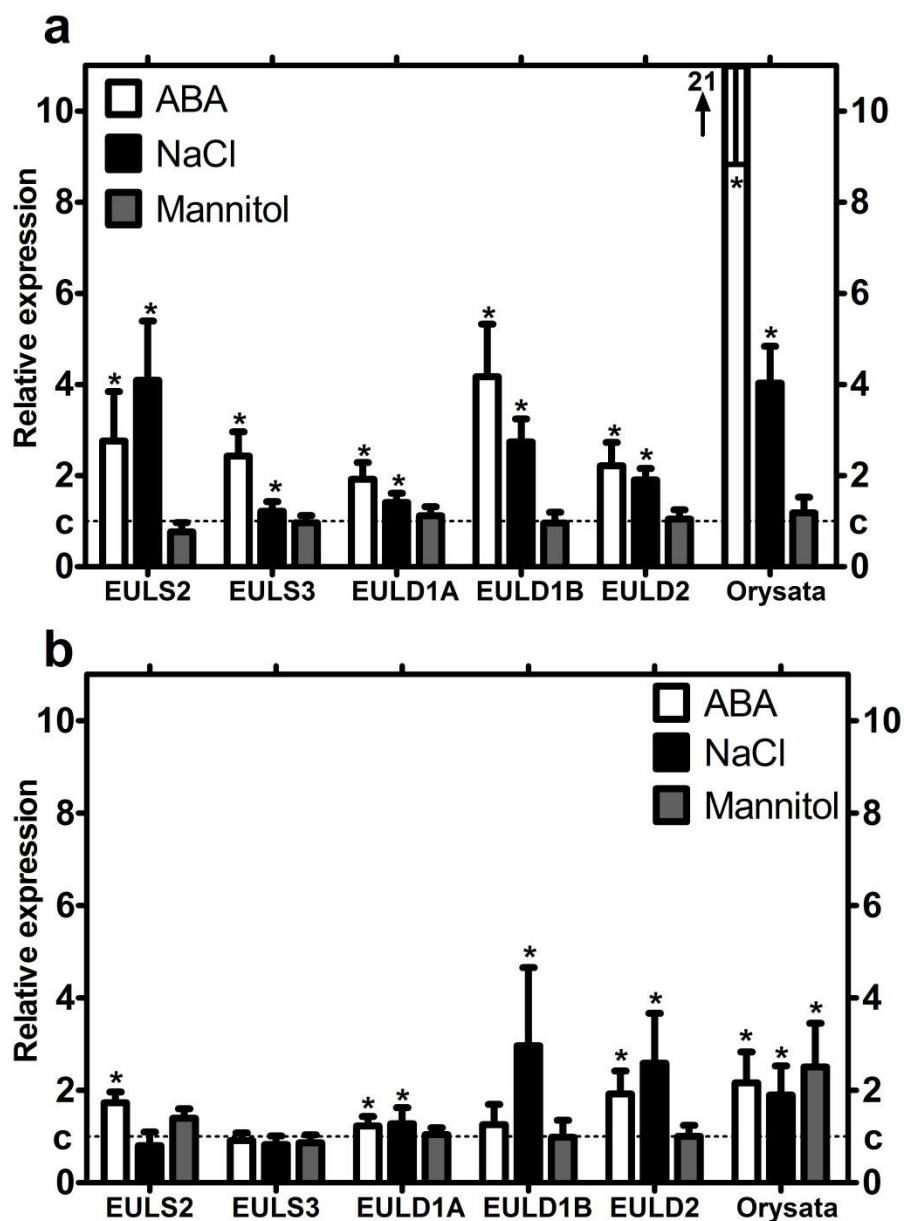
Fig. 3B.1 Expression analyses of rice lectins under different abiotic stress conditions. Relative expression level for all proteins under study in root (a) and shoot (b) tissues are shown. Quantitative RT-PCR was performed on RNA extracted from shoot and root tissues of 10-day old rice seedlings grown in MS medium (control) or treated for 6 h with 150 mM NaCl, 100 mM mannitol or 100 μM ABA. Gene expression levels were normalized using two internal reference genes: EXP (LOC_Os03g27010) and EIF5C (LOC_ Os11g21990,1). Bars represent means and standard error of three biological replicates, each containing a pool of 10 plants per treatment. Asterisks indicate statistically significant differences of expression in comparison with the control. EULS2: OrysaEULS2; EULS3: OrysaEULS3; EULD1A: OrysaEULD1A; EULD1B: OrysaEULD1B; EULD2: OrysaEULD2

3B.4.2 EUL expression patterns in response to infection assays

For the biotic stress assays, the fifth and sixth stage leaves of susceptible rice plants were infected with the pathogens *Magnaporthe oryzae* and *Xoo*, respectively. Similar to the abiotic stress experiments, the data were normalized by using two reference genes. For the *Xoo* experiment we selected the reference genes EXP and EIF5C. However, since these references genes were not stable after *Magnaporthe* infection, other reference genes were used to normalize the data obtained from the *Magnaporthe* experiment namely, Eukaryotic elongation factor 1-alpha (EEFa1, LOC_Os03g08020.1) and ubiquitin 5 (UBQ5, LOC_Os01g22490.1) (Jain et al., 2006) (Suppl. Table 3B.1). In both sets of experiments, clear disease symptoms were observed starting 4 dpi, indicating that pathogen inoculation was successful. Since the pathogen *Mo* can infect the rice plant faster compared to the pathogen *Xoo*, samples were collected 1 and 2 dpi after Mo infection whereas the sampling time was 2, 4 and 8 dpi for Xoo infection.

After *Xoo* infection, the expression levels for OrysaEULD1B, OrysaEULD2 and Orysata were up-regulated significantly 4 and 8 dpi whereas they were not affected 2 dpi (Fig. 3B.2a, Suppl. Table 3B.2). Among the EUL type D, OrysaEULD1A was the only protein that showed down-regulation of its expression, being significant 8 dpi whereas the expression was not regulated 2 and 4 dpi (Fig. 3B.2a). No significant changes in the expression for the S type EULs were observed (Fig. 3B.2a, Suppl. Table 3B.2).

The relative expression for the rice lectins was totally different after *Magnaporthe* infection. At 1 dpi a significant down-regulation of the expression was observed for most of the proteins under study. Particularly, the expression level of OrysaEULS3 was down-regulated significantly 1 dpi but not affected 2 dpi whereas OrysaEULS2 was not regulated after all tested time points (Fig. 3B.2b). After 1 dpi, the relative expression was down-regulated significantly for OrysaEULD1A and not affected for OrysaEULD1B 1 dpi whereas there was no significant regulation for both proteins 2 dpi (Fig. 3B.2b, Suppl. Table 3B.2). In spite of the fact that the expression level for OrysaEULD2 was down-regulated significantly 1 dpi, it was up-regulated significantly 2 dpi (Fig. 3B.2b). In addition, the expression pattern for Orysata showed significant down-regulation at all tested time points (Fig. 3B.2b).
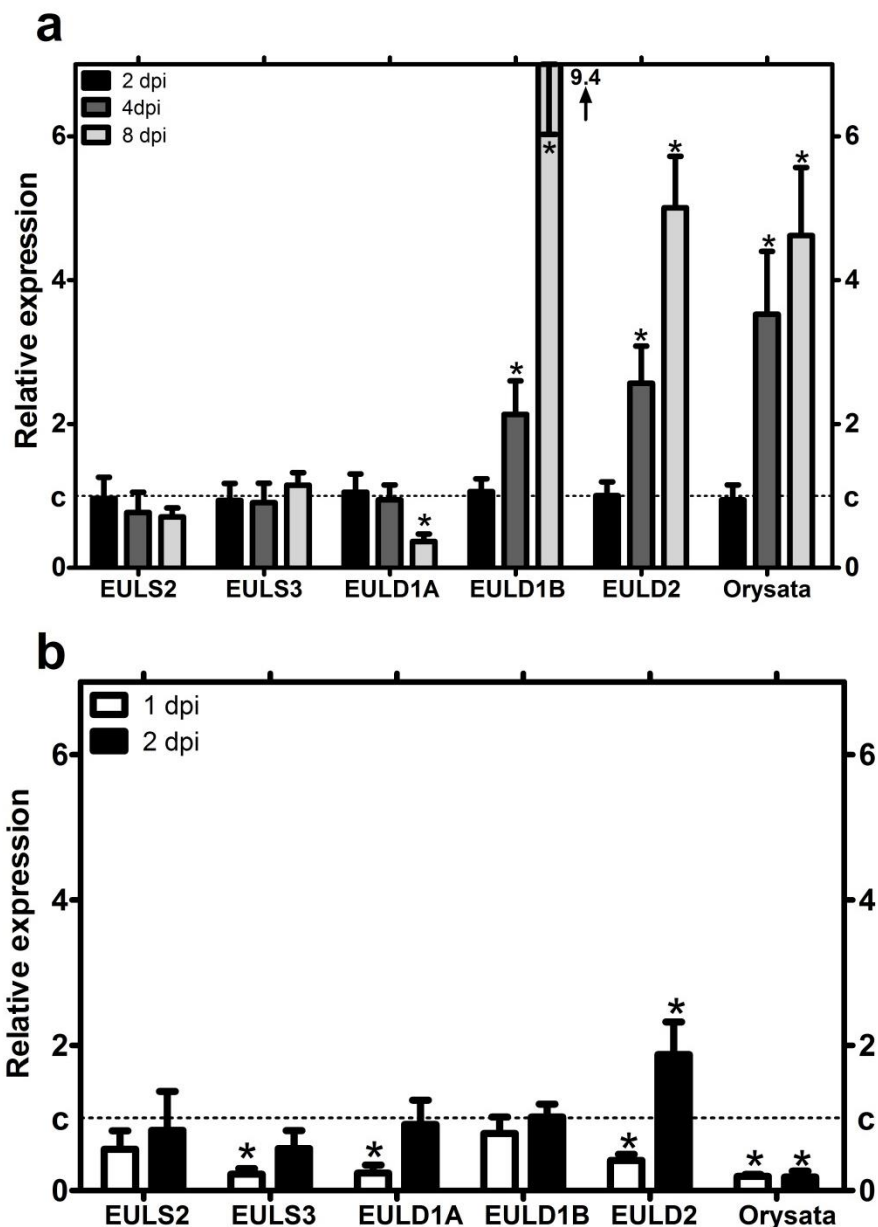
Fig.3B.2 Expression profile for all EUL lectins as well as Orysata after biotic infections. The relative expression for all tested proteins is shown in panel (a) after *Xoo* infection and in panel (b) after *Magnaporthe orysae* infection. Bars represent means and standard error of three biological replicates for *Xoo* and two biological replicates for *Magnaporthe*. To normalize the data obtained after each treatment, genes EXP (LOC_Os03g27010) and EIF5C (LOC_Os11g21990,1) were used to for *Xoo* treatment and genes Eukaryotic elongation factor 1-alpha (EEFa1, LOC_Os03g08020.1) and ubiquitin 5 (UBQ5, LOC_Os01g22490.1) were used for *Magnaporthe*. Asterisks indicate statistically significant differences of expression in comparison with the control. EULS2: OrysaEULS2; EULS3: OrysaEULS3; EULD1A: OrysaEULD1A; EULD1B: OrysaEULD1B; EULD2: OrysaEULD2

3B.5 Discussion

Inducible EUL lectins are of much interest because these carbohydrate-binding proteins are present in all terrestrial plants and their expression is enhanced in response to different stress situations (Imanishi et al., 1997; Zhang et al., 2000; Chen et al., 2002; Lannoo et al., 2006a). Recently, the hypothesis has been put forward that inducible lectins might play a role in the stress physiology of plant cells (Van Damme et al., 2008). As a result of exposure to different abiotic and biotic stresses, potential yields of economically important crops (including rice) are lost yearly (Mittler and Blumwald 2010, Seo et al., 2011). Since rice is both a model and a crop plant, it provides an excellent system to study the effect of stress on growth and yield under field as well as under laboratory conditions.

On a global level, a major limitation of crop production is imposed by different abiotic and biotic stresses resulting in 30%–60% yield losses each year (Dhlamini et al., 2005). Among the abiotic stresses, salinity and drought have received a lot of attention. Salinity adversely affects both the quantity and the quality of the crop yield (Gepstein et al., 2006; Blumwald and Grover 2006; Hauser et al., 2011). Drought is the most important limiting factor for crop production in many regions of the world (Bouman et al., 2005; Passioura 2007; Witcombe et al., 2008). For the biotic stresses, bacterial leaf blight caused by *Xoo* as well as rice blast caused by *Magnaporthe oryzae*, give rise to devastating crop losses in rice (Baker et al., 1997; Talbot et al., 2003; Sana et al., 2010). Additionally, ABA is a plant stress hormone and one of the most important signaling molecules in plants. It is well known that ABA plays a key role in regulating many developmental and adaptive processes after abiotic stresses (Meng et al., 2009; Santner et al., 2009; Cutler et al., 2010) as well as biotic stresses (Asselbergh et al., 2008; Cao et al., 2011). Taking into account these facts several stresses were selected for investigation in this study. Salinity and drought were chosen to represent abiotic stresses, whereas *Xoo* and *Magnaporthe* were selected for the biotic assays. In addition ABA was included due to its importance in responsiveness to both abiotic and biotic stresses.

To improve crop yield, it is necessary to understand the response mechanism of plants to abiotic and biotic stresses with the ultimate goal of improving crop performance under these unfavorable conditions. The present study investigates the expression profile for the whole

set of EUL proteins under several abiotic and biotic conditions in an attempt to gain better knowledge about their physiological importance. Recently, it was shown that the EUL domain from OrysaEULS2 (Al Atalah et al., 2012; see chapter 2B), the full protein of OrysaEULD1A and each of its domains (see chapter 2C) can be considered as true lectins since they can specifically interact with carbohydrate structures. The question rises why rice expresses multiple EUL type lectins.

3B.5.1 Responsiveness towards abiotic stresses

Since the Ct values for **OrysaEULS2** are ~25 in the shoot tissues and ~31 in the root tissues, it can be concluded that the expression of this protein was low in the shoots and even lower in the roots compared to the expression of the reference genes (Ct value ~20). This observation may explain why no GUS activity was detected when the GUS gene was expressed under the control of the OrysaEULS2 promoter sequence in *Arabidopsis* (Al Atalah et al., 2013; see chapter 3A). Quantitative PCR analysis after stress treatment with ABA for 6h revealed that OrysaEULS2 expression was significantly up-regulated 1.7 fold in the shoots (Fig. 3B.1b). Furthermore, the up-regulation of OrysaEULS2 was observed in the roots after ABA (2.8 fold) and salt (4 fold) treatments (Fig. 3B.1a). Our findings match earlier results which showed that OrysaEULS2 was up-regulated in the shoots after ABA and salt treatments (Moons et al., 1997a). Furthermore, OrysaEULS2 was not detected in the roots under normal growth conditions by RNA gel blot hybridization or under mild salt (100 mM for 3 days) by two dimensional blot analyses (Moons et al., 1997a) which is in good agreement with the very low expression levels observed in our study (Ct values ~31).

OrysaEULS3 expression was not changed in the shoots but was up-regulated significantly in the roots after application of ABA (2.4 fold) and salt (1.2 fold) stresses (Fig. 3B.1a, b respectively). The GUS staining patterns for the promoter of OrysaEULS3 showed down-regulation of the promoter activity in the shoots after salt treatment when tested in *Arabidopsis* (Al Atalah et al., 2013; see chapter 3A). The up-regulation of OrysaEULS3 expression in the roots after salt treatment was also inferred from the microarray data by Cheng et al. (2009). In addition, this protein was reported to be up-regulated in rice panicles in response to salt stress (Dooki et al., 2006).

In our qPCR analysis, **OrysaEULD1A** was up-regulated significantly in the shoots and the roots after treatment of the plantlets with ABA and salt (Fig. 3B.1). Relative expression levels for OrysaEULD1A were up-regulated by a factor of 1.2 and 1.9 in the shoots and roots, respectively, after ABA treatment whereas salt treatment resulted in 1.3 and 1.4 fold up-regulation in the shoots and the roots, respectively (Suppl. Table 3B.2). Mannitol treatment caused no significant effect on the expression level of OrysaEULD1A neither in the root nor in the shoot tissues (Fig. 3B.1, Suppl. Table 3B.2). The expression level of OrysaEULD1A is somewhat higher in the roots compared with the shoots in response to ABA and salt treatments. Our results are consistent with some data reported previously. OrysaEULD1A was detected by two-dimensional protein blot analysis from root extracts of rice seedlings exposed to mild salt stress (100 mM NaCl) for 3 days, (Moons et al., 1997a). Microarray data revealed that OrysaEULD1A was up-regulated almost 1.6 fold in the root of Pokkali, a rice salt tolerant variety, seedlings after salt treatment (150 mM) for 24 h (Kawasaki et al., 2001). Furthermore, OrysaEULD1A was reported as an ABA inducible protein in scutellar tissues of rice seedlings (Asakura et al., 2007).

As for the expression of **OrysaEULD1B**, it was up-regulated significantly 4.2 and 2.7 fold after ABA and salt treatments respectively in the root tissues (Fig. 3B.1a) whereas it was up-regulated significantly only in response to salt application in the shoots. Surprisingly and unlike OrysaEULD1A, OrysaEULD1B showed no expression change in the shoots after applying ABA exogenously. Our results are consistent with several reports showing that OrysaEULD1B can be considered as a stress responsive protein. OrysaEULD1B expression was shown to be up-regulated at high levels in the roots of seedlings after salt (Moons et al., 1997a, Kawasaki et al., 2001, Cheng et al., 2009) and ABA treatments (Moons et al., 1995, 1997). In addition, EST sequences encoding OrysaEULD1B have been obtained from libraries generated from the leaves (but not in the roots) of one month old drought-stressed seedlings from indica rice (Babu et al., 2002). Indeed, none of the EULs showed any responsiveness towards mannitol which was used to mimic drought. OrysaEULD1B was also up-regulated in the shoots of 12 day seedlings exposed to 140 mM of NaCl for 24h (Chao et al., 2005). Furthermore, OrysaEULD1B was reported as an ABA inducible protein in scutellar tissues of rice seedlings (Asakura et al., 2007).

Similar to OrysaEULD1A, **OrysaEULD2** was up up-regulated significantly 1.9 and 2.2 fold after ABA treatment; 2.6 and 1.9 after salt treatment in the roots and the shoots, respectively (Fig. 3B.1).

Our qPCR results showed that **Orysata** was up-regulated in the shoots significantly by 2.2, 1.9 and 2.5 fold after ABA, salt and mannitol treatments, respectively. A marked increase of the relative expression level of Orysata was observed in response to ABA treatment since lectin expression was up-regulated significantly by 21 fold in the roots. In addition Orysata was up-regulated in the roots 4 fold in response to salt treatment (Fig. 3B.1a). These observations are in good agreement with different reports in literature. Claes et al. (1990) have shown that Orysata expression was up-regulated in the sheath and the lamina of 8 day seedlings exposed to 1% NaCl for 7 days as well as in the sheath and the root (but not in the lamina) after 20 µM ABA treatment for 3 days. Furthermore, Orysata was also expressed in the sheaths of plants treated with NaCl and ABA (De Souza et al., 2003). Orysata was also up-regulated in leaf segments from rice seedlings treated with ABA for 8 h under light at 30°C (Jiang et al., 2010b).

3B.5.2 Responsiveness towards biotic stresses

Our expression data revealed that **OrysaEULS2** expression was not significantly changed upon *Xoo* and *Mo* treatments (Fig 3B.2 a, b respectively), suggesting that this protein might not be involved in the plant response to both pathogens. Conversely, the expression of **OrysaEULS3** was 4.5-fold down-regulated at 1 dpi after *Mo* infection (Fig. 3B.2b) whereas it was not significantly altered after *Xoo* infection (Fig. 3B.2a, Suppl. Table 3B.2). Down-regulation of OrysaEULS3 triggered by *Mo* but not *Xoo* may suggest a specific role of this protein in resisting fungal pathogen attack. Supporting this hypothesis, OrysaEULS3 was previously also shown to be specifically up-regulated during incompatible, but not compatible, interactions of rice with the fungal brown spot pathogen *Cochliobolus miyabeanus* (Van Bockhaven et al., 2011).

Compared to the S type EULs, proteins belonging to the D type class of EULs tended to respond more strongly to pathogen attack. Both *Mo* and *Xoo* infection caused a 4.2 fold and 2.7 fold repression of **OrysaEULD1A** at 1 dpi and 8 dpi, respectively. These results are in good agreement with other reports describing the down-regulation of this protein as a result

of microbial attack or herbivore infestation. Feeding experiments with the brown planthopper (a piercing-sucking insect) revealed that OrysaEULD1A expression was down-regulated in both the brown planthopper-resistant rice variety B5 and the brown planthopper-susceptible variety MH63 (Zhang et al., 2004). In addition, OrysaEULD1A expression was down-regulated 1.8 fold in the galls of 12-day-old rice roots after 7 days inoculation of root knot nematode (*Meloidogyne graminicola*) (Kyndt et al., 2012). Opposite to OrysaEULS3, OrysaEULD1A may therefore be involved in regulating basal plant defense responses that are effective against a wide variety of attackers.

Contrary to OrysaEULD1A, **OrysaEULD1B** expression remained stable after *Mo* infection, whereas it was strongly induced in response to *Xoo* infection, showing an approximate 2- and 9-fold up-regulation at 4 and 8 dpi, respectively. Such activation in response to *Xoo* infection was also observed for OrysaEULD2, which corroborates earlier findings reporting a 5-fold induction in the galls of 12-day-old rice roots at 7 days after inoculation with the root knot nematode (*Meloidogyne graminicola*) (Kyndt et al., 2012). In addition, the levels of **OrysaEULD2** were increased 1.9 fold in the roots of 12-day-old rice seedlings inoculated with the root rot nematode (*Hirschmanniella oryzae*) (Kyndt et al., 2012). However, following *Mo* inoculation, we found OrysaEULD2 expression to be significantly down-regulated at 1 dpi and up-regulated at 2 dpi. Such opposite expression pattern in response to *Xoo* and *Mo* infection was also seen for Orysata, the transcription of which was consistently down-regulated after *Mo* infection but strongly up-regulated at 4 and 8 days post *Xoo* inoculation.

One possible explanation to reconcile these apparently conflicting observations is that OrysaEULD2 and Orysata may fulfill antagonistic roles in resistance to *Mo* and *Xoo*, acting as either a positive or negative regulator of plant immunity depending on the type of attacker encountered. However, given the strong overlap in rice defenses against *Mo* and *Xoo* and the large number of plant resistance elicitors capable of inducing resistance against both pathogens (De Vleesschauwer et al., 2013; Van Bockhaven et al., 2013), this scenario appears rather unlikely. Therefore, rather than pointing to a dichotomous role in *Mo* and *Xoo* resistance, down-regulation of OrysaEULD2 and Orysata in *Mo*-inoculated plants may suggest that *Mo* is actively interfering with the expression of these proteins in order to counteract the buildup of plant resistance. In other words, it is not unlikely that pathogens such as *Mo* are targeting plant lectins as a virulence strategy. Supporting this view, *Mo* has

been shown repeatedly to induce rice disease susceptibility by targeting specific plant proteins through means of hitherto specialized effector proteins (Kawano and Shimamoto, 2013; Liu et al., 2013). Moreover, it was already shown that over-expression of Orysata suppressed hyphal growth of *Mo* and subsequently enhanced the resistance of rice plants, thus linking Orysata to plant immunity (Shinjo et al., 2011). Testing the disease resistance of rice plants with increased or reduced expression of EUL and/or Orysata and assessing the impact of purified proteins on fungal biology will be of particular help in further elucidating the immune-regulatory role of these proteins.

3B.5.3 Speculation on the physiological role of EULs

The physiological role of plant lectins most likely relies on their specific carbohydrate-binding activity and specificity. Inducible lectins (including the EUL family) are up-regulated upon exposure to different stresses. Based on these facts, the concept was developed that lectin-mediated protein-carbohydrate interactions play an important role in the stress physiology of the plant cell. Our data clearly indicate that rice expresses multiple carbohydrate-binding proteins in response to different stresses. Whereas Orysata is known to specifically recognize and bind mannose containing glycan structures (Al Atalah et al., 2011; see chapter 2A), the carbohydrate-binding properties of the EUL proteins are more diverse, in that different EUL sequences will recognize different carbohydrate structures (Fouquaert and Van Damme 2012). It is known that e.g. OrysaEULS2 will react with oligomannosidic structures, but will also recognize more complex carbohydrate structures containing lactosamine and GlcNAc structures (Al Atalah et al., 2012; see chapter 2B). Preliminary data for OrysaEULD1A indicate that this protein reacts with galactosylated structures.

It is generally accepted that plants can change their gene expression and protein accumulation in response to biotic and abiotic stresses. The root is usually the first plant organ that senses salt stress. Some salt stress-responsive genes were found to be mainly, or more strongly, induced in roots than in other organs (Yan et al., 2005). The EUL proteins are expressed at higher levels in the roots compared to the shoots in response to different stress treatments. At present we can only speculate about the physiological relevance of the carbohydrate-binding activity of the different EUL lectins in planta, but similar to what is known for Orysata (Shinjo et al., 2011), it is assumed that these lectins could play an

important role in signaling processes as a result of biotic or abiotic stress responses. To our knowledge this is the first paper dealing specifically with the regulation of EUL lectins after abiotic and biotic stresses. The quantitative data obtained from our qRT-PCR analyses clearly establish EUL lectins as a family of stress-related proteins and suggest that they might play a widespread role in the signal transduction circuitry determining plant stress tolerance or susceptibility. Future work will be focused on unraveling the mode of action and physiological importance of EUL proteins under various stress conditions, using transgenic rice lines with increased or reduced expression of EULs.

Chapter 4

Performance of transgenic lines over-expressing OrysaEULS2, OrysaEULD1A and Orysata after biotic and abiotic stress application

Chapter 4A

Can EUL proteins from rice increase the stress tolerance of *Arabidopsis* plants?

4A.1 Abstract

Plants synthesize tiny amounts of inducible lectins in response to environmental stresses. One family of these proteins groups is the *Euonymus*-related lectins (EUL) which is widely distributed within the plant kingdom. In the present research, two of the EUL proteins, OrysaEULS2 and OrysaEULD1A as well as Orysata, a protein belonging to the jacalin-related lectin family were investigated for their contribution to the stress tolerance of the plant. All genes encoding the proteins under study were cloned behind the 35S promoter and transformed into *Arabidopsis* plants. Subsequently, the performance of several one-insertion homozygous transgenic lines was analyzed in response to different stresses including NaCl, mannitol (drought), ABA and *Pseudomonas* infection. These experiments were performed on different developmental stages of transgenic *Arabidopsis* particularly, the germination stage, seedlings as well as the adult plants. Transgenic *Arabidopsis* over-expressing OrysaEULS2 and OrysaEULD1A showed hypersensitivity of ABA on germination level. Plants showed short stem height in the early stages of development and bolted earlier compared to the wild type. Only after moderate mannitol treatment, the transgenic lines showed a better responsiveness compared to the control plants. On adult level, transgenic plants harboring the proteins under study enhanced the tolerance against salinity and drought. After *Pseudomonas* infection, only OrysaEULD1A conferred tolerance to the transgenic *Arabidopsis* plants. Localization studies revealed that OrysaEULS2 and OrysaEULD1A are localized to the nucleus and the cytoplasm of the plant cell. Our findings indicate that EUL proteins can confer tolerance to *Arabidopsis* plants against environmental stresses and hence can be used to increase the tolerance of crops to cope with external stress factors.

4A.2 Introduction

Since the world population will reach more than 9 billion in 2050 (Godfray et al., 2010; Tester and Langridge 2010), it is a major challenge to satisfy the increasing demand for food production. Abiotic stresses, especially salinity, drought, temperature and oxidative stress, together with the biotic stresses have a crucial impact on the productivity and yields of crops. As sessile organisms, plants are continuously threatened by different biotic and abiotic stress factors and therefore have evolved sophisticated mechanisms to face these threats. Hundreds of genes respond to stress at transcriptional and translational level (for reviews see Cushman and Bohnert 2000; Sreenivasulu et al., 2004; Yamaguchi-Shinozaki and Shinozaki 2005; Umezawa et al., 2006). Understanding the functions of these stress-inducible genes will give the chance to reveal the mechanisms that plants trigger to tolerate stresses.

Transgenic plants open the door to improve crop tolerance to abiotic and biotic stresses. These transgenic strategies most often rely on the transfer of one or several genes that are involved in signaling and regulatory pathways or that encode proteins which confer stress tolerance (Wang et al., 2003; Vinocur and Altman 2005; Valliyodan and Nguyen 2006; Sreenivasulu et al., 2007; Kathuria et al., 2007). Among these stress related proteins is a group of proteins called lectins. Plant lectins are a complex and heterogeneous group of carbohydrate-binding proteins that specifically recognize and bind to carbohydrate structures (Van Damme et al., 2008). In the last decade, accumulating data have demonstrated that plants synthesize low amounts of lectins in response to abiotic and biotic stresses (Van Damme et al., 2004a; Lannoo and Van Damme 2010). These inducible lectins are synthesized on free ribosomes and locate to the nucleus and/or the cytoplasm of the plant cell (Van Hove et al., 2011, Al Atalah et al., 2011 see chapter 2A). The physiological importance of these inducible plant lectins remains ambiguous. However, they are suggested to be involved in defense and have a role in signal transduction in response to stress factors (Van Damme et al., 2004a; Van Damme et al., 2008). Until recently, six classes of inducible lectins have been reported (Lannoo and Van Damme 2010).

One class of inducible lectins is the group of Jacalin-related lectins (JRL) which comprises all proteins that have one or more domains similar to jacalin, a protein isolated from *Artocarpus*

integrifolia (Bunn-Moreno and Campos-Neto 1981). The majority of the JRLs are mannose-specific lectins, but some of them show specificity towards galactose related structures (Peumans et al., 2000b). The responsiveness of mannose-binding JRLs towards stress has been reported in *Arabidopsis* (Chisholm et al., 2001), wheat (Subramanyam et al., 2008; Ma et al., 2010), barley and rice (Qin et al., 2003). In 2008, a new group of inducible lectins was discovered after cloning of the *Euonymus europaeus* agglutinin. Since then this lectin is considered as the prototype for the class of so-called *Euonymus*-related lectins (EUL) (Fouquaert et al., 2008). Based on the extensive screening of all available genomes and transcriptomes, proteins with one or more EUL domains were identified throughout the plant kingdom (Fouquaert et al., 2009a). Similar to JRLs, EUL proteins were also shown to be stress related in *Arabidopsis* (Fouquaert et al., 2009a), banana (Carpentier et al., 2007), maize (Riccardi et al., 2004) and rice (Moons et al., 1997a; Kawasaki et al., 2001; Dooki et al., 2006; Cheng et al., 2009; Kyndt et al., 2012).

In the present study, the physiological importance of several inducible lectins from rice, including Orysata, belonging to the jacalin-related family as well as OrysaEULS2 and OrysaEULD1A, representing the single EUL domain and the double EUL domain proteins, respectively, was investigated in response to different stresses. Coding sequences for each of the proteins under study were cloned after the 35S promoter and transformed into *Arabidopsis* plants. Subsequently, one-insertion homozygous lines were selected and exposed to different abiotic and biotic stresses. A comparison was made between the transgenic plants grown under normal conditions and stress conditions, as well as with wild type plants. Our investigation suggests that OrysaEULS2, OrysaEULD1A as well as Orysata are stress-related proteins that can be used to improve the stress tolerance of crops.

4A.3 Methods and materials

4A.3.1 Plant material and growth conditions

Arabidopsis thaliana ecotype Columbia (Col-0) (Lehle Seeds, Round Rock, Texas, USA) was used for stable transformation. Sterilization of the seeds and growth conditions were as described in section 3A.3.1. For the homozygous lines, the seeds were sown on MS plates

whereas MS plates with kanamycin (50 µg/ml) were used for the heterozygous lines. To grow plants *in vitro*, round (9 cm diameter) and square plates (12 x 12 cm) were used. For each round plate, 25 ml MS (for the MS composition see section 2A.3.1) was poured and 40 ml MS was used for each square plate (the square plates were only used for root growth assays). For the *in vivo* culture, seeds were sown first on round plates and after two weeks *Arabidopsis* plantlets were transferred to artificial soil (Jiffy-7, columnar diameter 44 mm, AS Jiffy Products, Drobak, Norway). Unless it is mentioned otherwise, *Arabidopsis* plants were watered on Mondays and Thursdays by adding 2 L for each tray and the water was drained after 30 min. To synchronize the germination, *Arabidopsis* seeds were kept in the dark at 4°C until day 3. When counting the days, the day of sowing was designated as day 0.

For transient transformation, seeds from *Nicotiana benthamiana* were sown directly on soil and kept at 28°C, 16/8 light/dark cycle. Tobacco plants were watered once a week.

4A.3.2 Construction of vectors expressing the lectins under the 35S promoter, and N- and C- EGFP fusion constructs

Expression constructs were prepared as explained in sections 2A.3.3 and 2A.3.4. To prepare the 35S::constructs, the following primers were used to add the non-complete *attB* sequences to the flanking regions of the coding sequences encoding the proteins under study: EVD 545 5'AAAAAGCAGGCTTCACCATGACGCTGGTGAAGATTGGCCT3' and EVD 546 5'AGAAAGCTGGGTGTCAAGGGTGGACGTAGATGCCAATTGC3' to clone Orysata; EVD 541 5'AA AAAGCAGGCTTCACCATGGACTTTTACGGGCGGCGCGA3' and EVD 542 5'AGAAAGCTGGGTGT CAGTAGTAGGGCTGGATCTTCCAGCG3' to clone OrysaEULS2; EVD 543 5'AAAAAGCAGGCTTCA CCATGTTTGGGTTCGGGCACCACCA3' and EVD 563 5'AGAAAGCTGGGTGTGCAATGCAGCTCCCC CTCCCCGGAAT3' to clone OrysaEULD1A. All lectin sequences were cloned after the 35S promoter in the expression vector pK7WG2.0 (Karimi et al., 2002) allowing the sequences to be translated as proteins without any tags. To prepare the N- and C-terminal EGFP fusion constructs, the following primers were used to add the non-complete *attB* sequences: EVD 541 (above) and EVD 702 5'AGAAAGCTGGGTGGTAGTAGGGCTGGATCTTCCAGCG3' to prepare the OrysaEULS2::EGFP construct; EVD 543 (above) and EVD 839 5'AGAAAGCTGGGT GCAATGCAATGCAGCTCCCCCTCCCCGGAATCGA3' to prepare the OrysaEULD1A::EGFP. A point mutation was created in the reverse primers EVD 702 and 839 to inactivate the stop codon

allowing the translation of the chimeric proteins. For the construct::EGFP, the expression vector pK7FWG2 was used whereas the expression vector pK7WGF2 was used for the EGFP::construct.

4A.3.3 *Agrobacterium* transformation

For stable transformation, the binary vectors were mobilized into the *Agrobacterium* strain GV301 by electroporation as described in section 3A.3.3, whereas the C58C1-Pmp90 strain was transformed by tri-parental mating as described in section 4B.3.3.

4A.3.4 Transient transformation of *N. benthamiana* and microscopic analysis

Four week old healthy plants were used for the transient transformation. The *Agrobacterium* strain C58C1-Pmp90 was grown in 5 ml YEB (consisting of 5 g/l beef extract, 5 g/l peptone, 5 g/l sucrose, 1 g/l yeast extract and 15 g/l bacteriological agar) medium containing spectinomycin (50 µg/ml) and gentamycin (20 µg/ml) at 28°C for two days in the dark. After centrifugation of 2 ml of the *Agrobacterium* culture for 10 min at 7000 g, the cells were washed twice using the infiltration buffer [containing 50 mM MES, 2 mM Na_2HPO_4 and 0.5% (w/v) glucose, adjusted to pH 5.6 using 1 M NaOH]. Then, the cells were re-suspended in the infiltration buffer containing 100 µM acetosyringone and the OD_{600} was adapted to 0.1. Using 1 ml syringes, the *Agrobacterium* solution was injected in the lower epidermis cells of the tobacco leaves (4 injection points per leaf, 2 leaves per plant). Afterwards, the plants were watered and kept under normal growth conditions as mentioned above. Two days post injection; the infected areas for each leaf were mounted on a microscopic slide and analyzed under the confocal microscope (Nikon Instruments, Badhoevedorp, The Netherlands) using a 60x objective lens. Images were processed with ImageJ (http://rsbweb.nih.gov/ij/).

4A.3.5 Stable transformation of *A. thaliana*

Floral dip transformation of *A. thaliana* was performed as mentioned in section 3A.3.3 and was used to introduce the constructs of interest into the *Arabidopsis* genome.

4A.3.6 Selection for one-insertion homozygous lines expressing the native constructs

One-insertion homozygous lines expressing the native constructs were generated as explained in section 3A.3.3.

4A.3.7 Localization studies with transgenic *Arabidopsis* expressing the EGFP fusion construct

After floral dip transformation, the transformants were selected on MS medium plates containing kanamycin (50 µg/ml) following the fast selection procedure as described by Harrison et al. (2006). Positive transformants were allowed to grow into adult plants. After one month, one rosette leaf per line was mounted on a microscopic slide and analyzed using the confocal microscope (see section 4B.3.4).

4A.3.8 Preparation of crude extracts and western blot analysis

Crude protein extracts from the transgenic heterozygous (generation 1 = T1) and homozygous (T3) lines expressing the native constructs were performed by crushing 3 rosette leaves in an eppendorf tube using a micro-pestle after adding 100 µl of 20 mM un-buffered 1,3-diaminopropane. Afterwards, extracts were centrifuged at 13,000 g for 5 min. Subsequently, the lectin concentration was measured as mentioned in section 4B.3.4 and western blot analysis was performed as described in section 4B.3.6 except that the anti-OrysaEULS2 antiserum [produced by Thermo scientific by injecting two rabbits with recombinant EUL domain from OrysaEULS2 (Al Atalah et al., 2012; see chapter 2B)] was used as the primary antibody to detect OrysaEULS2 and OrysaEULD1A expression in the transgenic lines. Similarly the anti-Orysata antiserum [produced by Thermo scientific by injecting two rabbits with recombinant Orysata (Al Atalah et al., 2011; see chapter 2A)] was used to detect the expression of Orysata.

4A.3.9 Stress treatments

Three homozygous transgenic lines for each lectin construct as well as wild type plants were used in all stress assays as well as for the morphometric analysis. Particularly lines 4, 11 and 13 for Orysata (designated as Orysata (4), Orysata (11) and Orysata (13)); lines 3, 9 and 11 for OrysaEULS2 (designated as OrysaEULS2 (3), OrysaEULS2 (9) and OrysaEULS2 (11)) and lines 5, 7 and 13 for OrysaEULD1A (designated as OrysaEULD1A (5), OrysaEULD1A (7) and OrysaEULD1A (13)) were tested. Plants were subjected to different stresses being 100 and 150 mM NaCl, 50 and 100 mM mannitol and 50 and 100 µM ABA for abiotic stresses. In addition, transgenic lines were exposed to the bacterial strain *Pseudomonas syringae*. The

data generated after each stress treatment were compared for the transgenic lines and the wild type plants.

4A.3.9.1 Morphometric analysis for the transgenic lines

The morphometric analysis was conducted under normal growth conditions. At the bolting time (when 50% of seedlings start bolting), the number of rosette leaves as well as the length, width and the petiole length of the largest rosette leaf were measured for each line. At day 35, the height of the primary inflorescence was measured for each plant. When 50% of the siliques were dry, we stopped watering the plants and let them dry. Finally, the seeds for each line were harvested and weighed. All the measurements were performed for 17 plants per transgenic line per construct.

4A.3.9.2 Germination assays

The germination rate for all transgenic seeds as well as the wild type seeds was estimated under above mentioned (abiotic) stress conditions. For each line 25 seeds were sown on MS plates containing the desired stress factor (NaCl, mannitol and ABA) and three replicates were performed for each treatment. At day 6, the number of germinated (when the radicles had emerged) seeds was counted.

4A.3.9.3 Root length assays with seedlings

To perform the root length assays, 48 seeds per construct were sown in square plates containing 1.2% MS agar. Individual seeds were sown on four lines allowing 3 cm space between the lines. Afterwards, the plates were kept vertically in the growth incubator. At day 6, twelve seedlings were transferred into new square plates containing the desired concentration for each stress factor. In each plate, the seedlings were put on two lines. Subsequently, the plates were incubated vertically in the growth incubator. At day 16, pictures were taken for all plates. The length of the roots was measured by ImageJ software and the seedlings were weighed.

4A.3.9.4 Assays on adult plants

For the drought assay, the water was stopped for 12 days starting from day 26. The plants that survived were scored 6 days after re-watering.

For the NaCl assay, the plants were watered with water containing increasing concentrations of NaCl ranging from 100 mM (moderate stress) until 300 mM (severe stress) starting from day 28 every second day. Each salt concentration was used to water the plants twice. At day 38, the plants were watered with normal water. At day 42, the healthy plants were scored (plants were considered healthy if they had less than 3 yellowish or dry leaves).

Twelve plants were used for each transgenic line or the wild type, and the whole experiment was repeated 3 times.

4A.3.9.5 *Pseudomonas* assays

Pseudomonas syringae pv. Tomato strain DC3000 was used to infect the *Arabidopsis* plants. A protocol adapted from Pieterse et al. (1996) was followed. Briefly, the bacteria were grown for 2 days in 5 ml King's B medium [50 g/l peptone, 1.5 g/l KH_2PO_4, 1.5 g/l $MgSO_4.7H_2O$, 1% (v/v) glycerol, pH 7.2 adjusted with 0,5 M NaOH] at 28°C, 200 rpm. Afterwards, the OD_{600} was adapted to 0.6 – 1.0 (OD 1 corresponds to 5×10^8 cfu /ml) and the bacterial culture was shaken for 30 min. The cells were collected by centrifuging at 2500 g for 10 min, and re-suspended in 10 mM $MgSO_4$. The solution was diluted to OD_{600} 0.05 (corresponding to 2.5×10^7 cfu /ml) and 0.05% (final concentration) Silwet-77 was added. For the mock treatment, 10 mM $MgSO_4$ with 0.05% Silwet-77 was used.

The transgenic lines as well as the wild type plants were grown at 22°C with 12 dark/12 light cycle. For each line 17 plants were used for the *Pseudomonas* treatment as well as the mock treatment. This experiment was performed once.

To perform the infection assay, the plants were covered with saran wrap one day before the infection and plenty of water was added to open the stomata. The plants were sprayed with bacterial solution until run-off and covered with saran wrap for 2 days. Afterwards, pictures were taken for all plants 6 days after spraying and the healthy plants (containing less than 3 leaves with clear symptoms) were scored.

4A.3.10 Statistical analysis

All data are expressed as means ± SE or as percentages. To separate the means statistically, a two-tailed T test was applied using the Prism version 5 (GraphPad, La Jolla, CA).

4A.4 Results

4A.4.1 Localization analysis in tobacco and *Arabidopsis*

Microscopic analysis of tobacco leaves transiently transformed with a construct expressing a fusion protein of the lectin and EGFP revealed that OrysaEULS2 and OrysaEULD1A are located in the nucleus (including the nucleolus) and the cytoplasm of the parenchyma cells which is in good agreement with the absence of the signal peptide in both protein sequences. This result was observed for the constructs EGFP::OrysaEULS2, OrysaEULD1A::EGFP and EGFP::OrysaEULD1A (Fig. 4A.1a). However, no fluorescence was detected for the construct OrysaEULS2::EGFP.

Very similar results were obtained after stable transformation of *Arabidopsis* (Fig. 4A.1b). Fluorescence was detected in the nucleus (including the nucleolus) and the cytoplasm for the transgenic lines expressing the fusion constructs OrysaEULS2::EGFP and OrysaEULD1A::EGFP, but not for the constructs containing EGFP in front of the lectin sequence.
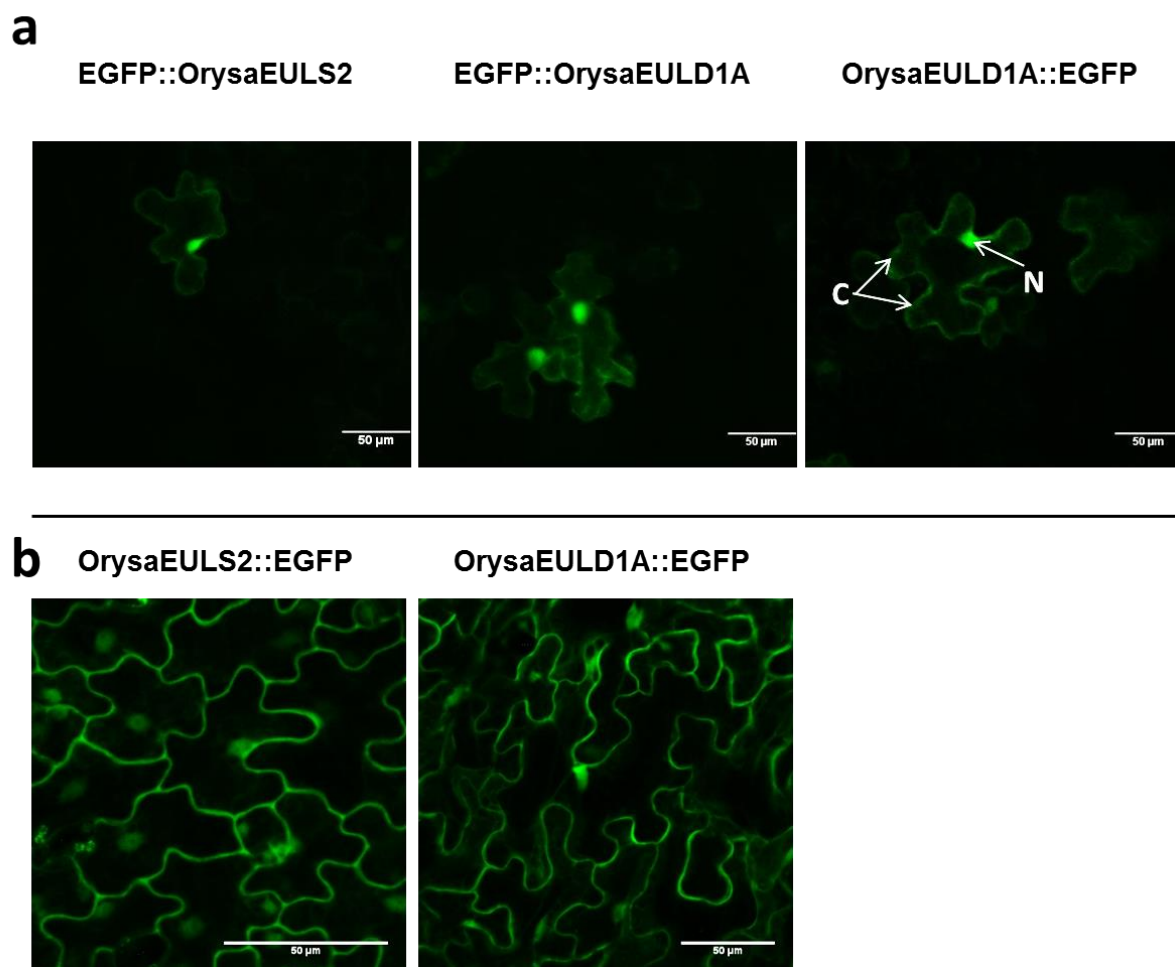
Fig.4A.1 Microscopic analysis for OrysaEULS2 and OrysaEULD1A constructs fused to EGFP. Pictures were taken after transient tobacco transformation (a) and after stable *Arabidopsis* transformation (b). Scale bars represent 50 μm. N: nucleus; C: cytoplasm.

4A.4.2 Expression of the native constructs in transgenic *Arabidopsis*

Using the Gateway technology the coding sequences of Orysata, OrysaEULS2 and OrysEULD1A were cloned under the control of the 35S promoter. After transforming the binary vectors harboring each of the lectin constructs into *Arabidopsis* plants, several independent lines (T1 lines) were analyzed at protein level (data not shown). Similarly the homozygous lines (T3 lines) were also checked by western blot analysis (Fig. 4A.2). The detection of a 23 kDa polypeptide (the exact size is 22.8 kDa) for OrysaEULS2, a 38 kDa polypeptide (the exact size is 38.5 kDa) for OrysaEULD1A and a 15 kDa polypeptide (the exact size is 15.2 kDa) for Orysata confirmed that our proteins were successfully expressed in the transgenic *Arabidopsis* lines.
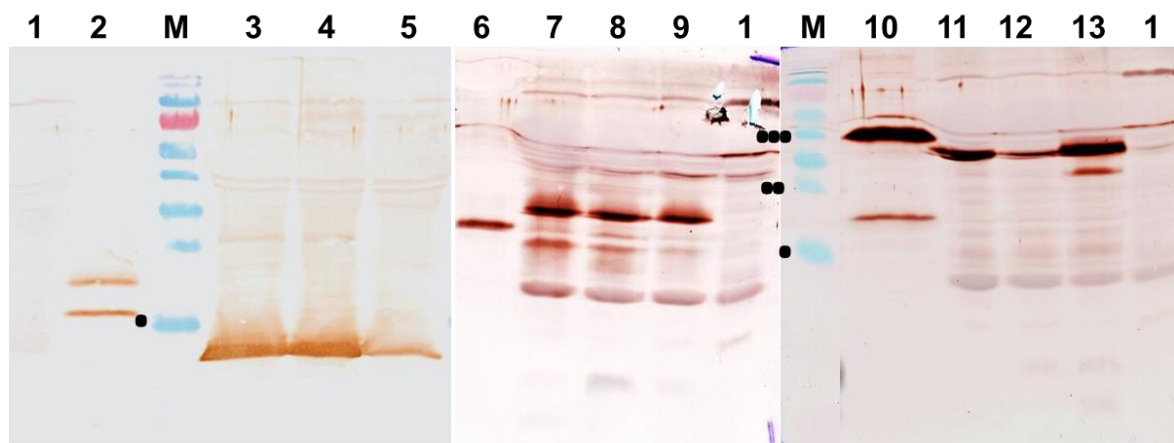
Fig. 4A.2 Western blot analysis of transgenic *Arabidopsis* lines expressing Orysata, OrysaEULS2 and OrysaEULD1A. Protein extracts were loaded following the following order: lane 1: wild type plant; lane 2: recombinant Orysata; lanes 3-5: Orysata T3 lines 4, 11 and 13; lane 6: recombinant EUL domain from OrysaEULS2; lanes 7-9: OrysaEULS2 T3 lines 3, 9 and 11; lanes; lane 10: recombinant OrysaEULD1A; lanes 11-13: OrysaEULD1A T3 lines 5, 7 and 13. In each lane 25 μg total protein was loaded for leaf extracts. Approximately 100 ng of the purified recombinant proteins (expressed in *Pichia pastoris*) were used as positive controls. M: protein ladder (Fermentas, St Leon-Rot, Germany). One, two and three dots refer to 17, 26 and 43 kDa sizes, respectively.

4A.4.2 Morphometric analysis under normal growth conditions

Various phenotypic features were analyzed to study the effect of overexpression of the genes of interest in *Arabidopsis* plants. The homozygous transgenic *Arabidopsis* plants exhibited pleiotropic effects at several stages of development. At bolting time, transgenic lines Orysata (4), Orysata (11), OrysaEULS2 (3), OrysaEULS2 (9) and OrysaEULD1A (13) showed a significantly increased number of rosette leaves. In addition, the length, the width and the petiole length of the largest rosette leaf were also increased (Fig. 4A.3 a, b, c and d respectively). Those transgenic lines (Orysata (4), Orysata (11), OrysaEULS2 (3), OrysaEULS2 (9) and OrysaEULD1A (13)) which showed increased biomass before bolting were significantly shorter (measured as height of the plant) at day 35 compared to the wild type plants (Fig. 4A.3e) and their bolting was delayed 5-10 days compared to the control plants (Fig. 4A.3g). In contrast, line OrysaEULS2 (11) which had significantly smaller rosette leaves was significantly taller compared to the wild type plants and it bolted 4 days earlier in comparison with the control (Fig. 4A.3). Interestingly, transgenic lines with increased biomass yielded a lower amount (measured by weight) of seeds compared to the wild type plants (Fig. 4A.3f). On the contrary, transgenic lines Orysata (13), OrysaEULS2 (11) and

OrysaEULD1A (5) with a size of rosette leaves comparable to the wild type plants gave a similar amount of seeds (Fig. 4A.3f).
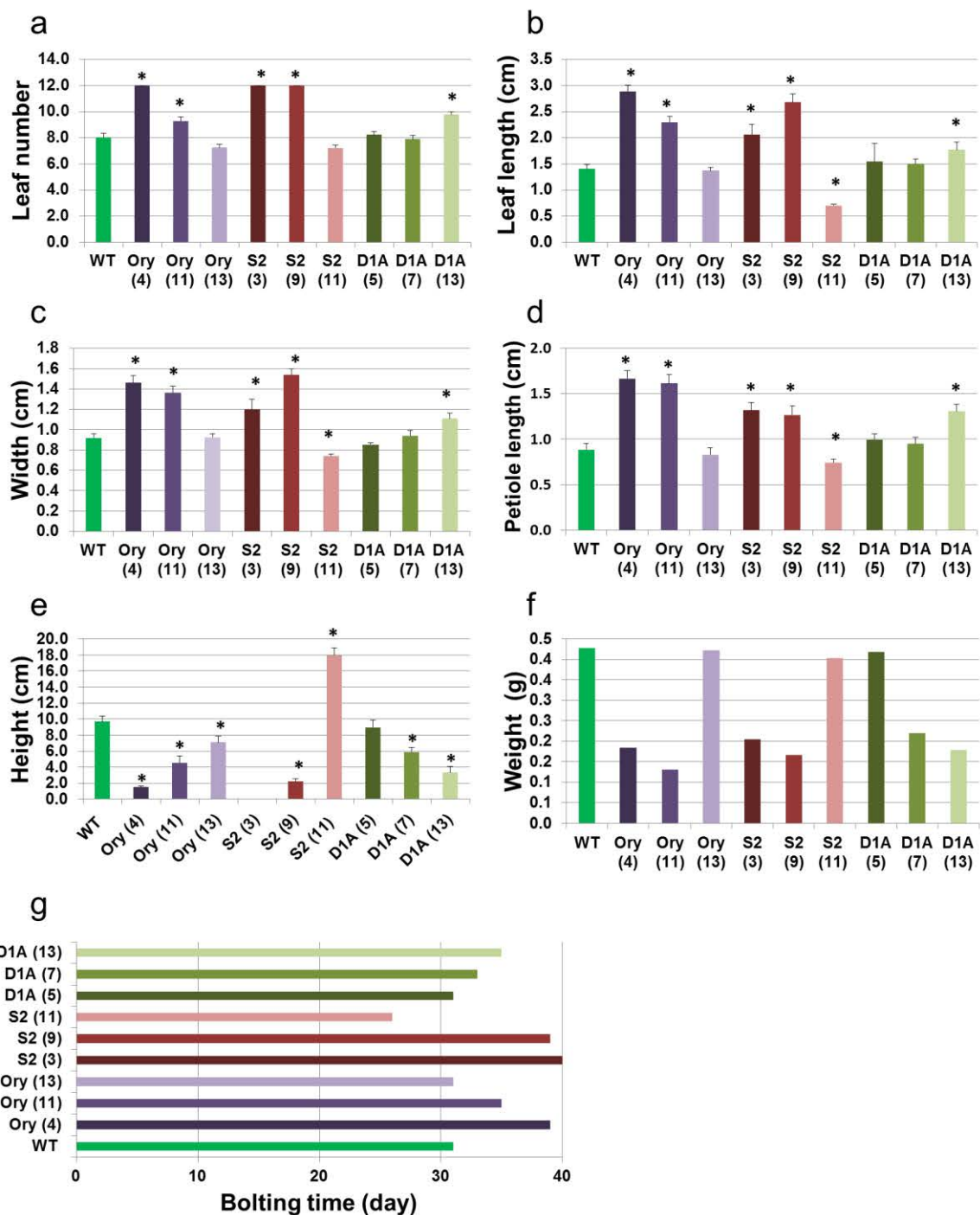

Fig. 4A.3 Legend on next page

Fig. 4A.3 Phenotypic analysis for the transgenic lines under normal growth conditions. At the bolting time the number of rosette leaves (a), the length (b), the width (c) and the petiole length (d) of the largest rosette leaves as well as the height of the primary inflorescence at day 35 (e), the weight of the seeds (f) and the bolting time (g) are shown. Bars in panels (a) to (e) represent the mean measured for 17 plants ± SE. No statistical analysis was performed in panels (f) and (g) because the bars represent a pool of all plants (n=17). In panel (e), the OrysaEULS2 (3) was not included because it did not yet shoot at day 35 (the shooting was at day 40). WT: wild type, Ory: Orysata; S2: OrysaEULS2 and D1A: OrysaEULD1A, numbers between brackets indicate different transgenic lines. Asterisks refer to significant differences in comparison with the control at P-value < 0.05

4A.4.3 Performance of seedlings for different transgenic lines in response to stress factors

The germination rate for the transgenic lines was estimated under different levels of abiotic stress as well as under the normal growth conditions. Under normal conditions, the seed germination was lower for all EUL lines tested but germination was not affected for the Orysata lines (Fig. 4A.4).

After the NaCl treatments, the germination of the Orysata lines was comparable to the wild type plants, but the OrysaEULS2 lines showed a significantly lower germination rate on medium containing 100 mM NaCl but a comparable germination to the control after treatment with 150 mM NaCl (Fig. 4A.4). The application of mannitol resulted in a significantly lower germination rate for two lines of Orysata (Orysata (11) and (13)), but only for Orsyata 11 this response was significant for both mannitol concentrations tested. Similarly two transgenic lines of OrysaEULS2 (OrysaEULS2 (9) and (11)) showed a decreased germination rate after mannitol treatments at both concentrations (Fig. 4A.4). When 50 μM ABA was applied, seed germination was lower only for Orysata line 13 whereas all OrysaEULS2 lines showed a significantly reduced germination rate. When the ABA concentration was increased to 100 μM also the Orysata lines showed a comparable germination to the wild type plants but only for OrysaEULS2 line 11 the germination was decreased significantly. Interestingly, all abiotic stresses severely suppressed the germination rate for all OrysaEULD1A lines (Fig. 4A.4).
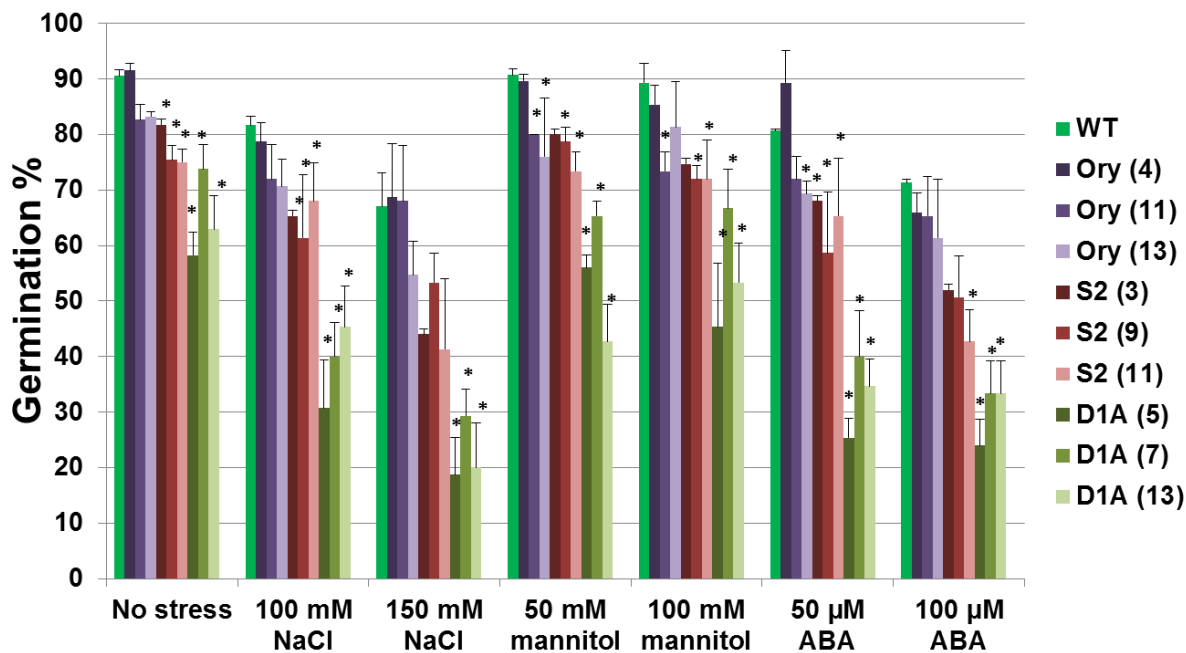
Fig.4A.4 Germination rate of transgenic lines under different levels of abiotic stress. Bars represent means ± SE based on three independent replicates, 25 seeds per replicate. WT: wild type, Ory: Orysata; S2: OrysaEULS2 and D1A: OrysaEULD1A, numbers between brackets indicate different transgenic lines. Asterisks indicate the significant differences in comparison with the control at P-value < 0.05

To observe the effect of different abiotic stresses on seedling growth of the transgenic lines and the wild type plants, the root length as well as the weight of the seedlings were determined.
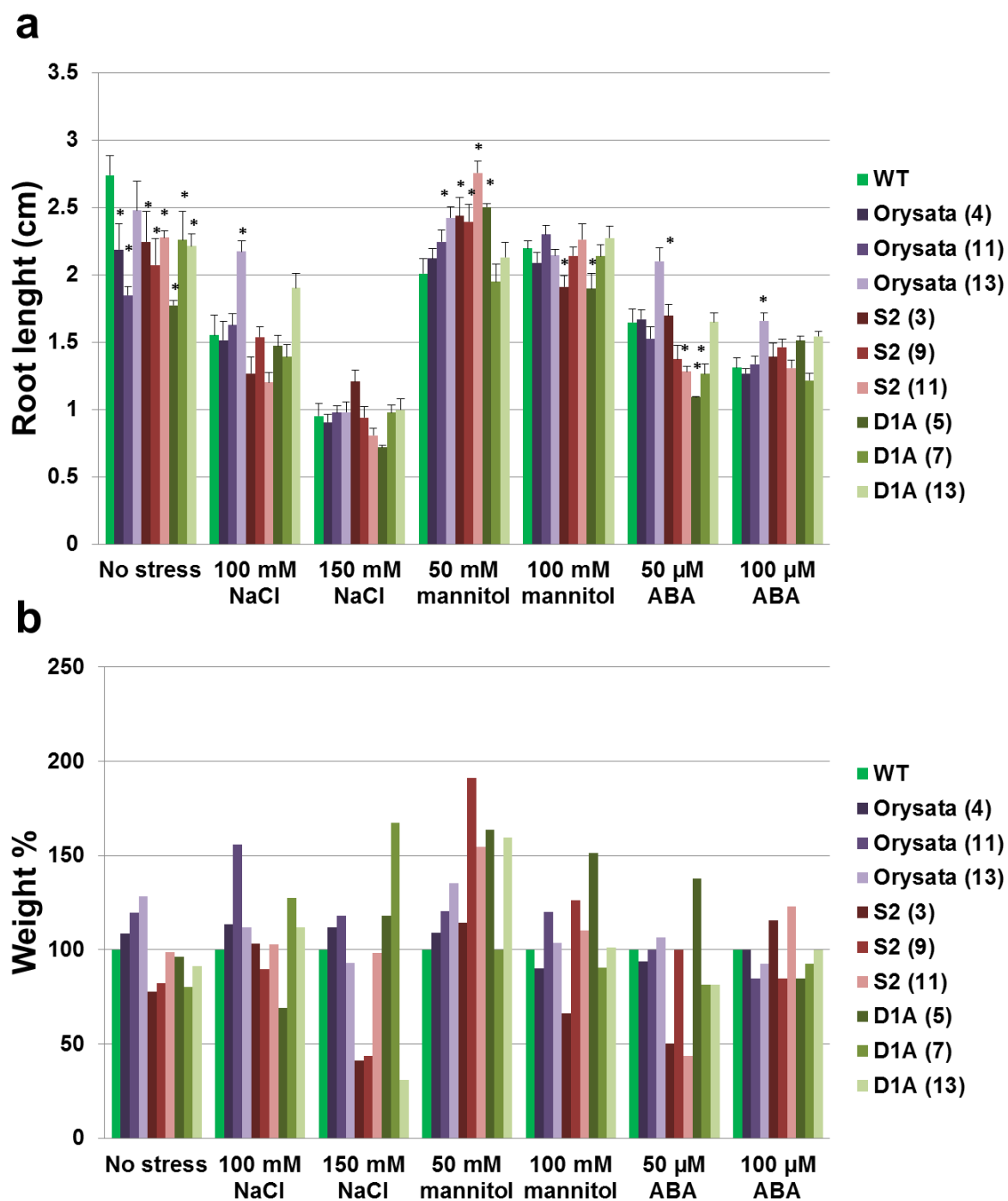
Fig. 4B.5 Performance of the transgenic lines on seedling level. Bars represent means ± SE for the root length (a) and the weight (b) of 12 seedlings under all tested conditions. WT: wild type, Ory: Orysata; S2: OrysaEULS2 and D1A: OrysaEULD1A, numbers between brackets indicate different transgenic lines. Asterisks indicate the significant differences in comparison with the control at P-value < 0.05

Under normal growth conditions, the root length for all transgenic lines was significantly shorter compared to the control except for Orysata (13) (Fig. 4A.5a). Only Orysata (13) developed longer roots compared to the wild type after NaCl treatments whereas the root length for all other transgenic lines was comparable to that of the wild type plants (Fig. 4A.5a). In addition, Orysata (13), all OrysaEULS2 lines and OrysaEULD1A (5) had longer roots compared to the control after moderate (50 mM) mannitol treatment but all lines showed comparable root length except for OrysaEULS2 (3) and OrysaEULD1A (5) which developed shorter root length in comparison with the wild type plants (Fig. 4A.5a). After ABA treatment, only Orysata (13) had longer roots compared to the control whereas OrysaEULS2 (11), OrysaEULD1A (5) and OrysaEULD1A (7) had shorter roots compared to the wild type seedlings.

The weight of the seedlings expressing Orysata was higher compared to that of the wild type plants grown under normal conditions whereas all seedlings from the transgenic lines expressing EULS2 and EULD1A showed lower weight (Fig. 4A.5b). After 100 mM NaCl treatment all Orysata lines together with OrysaD1A (7) and OrysaEULD1A (13) yielded a higher weight for the seedlings and similarly only lines 4 and 11 from Orysata as well as lines 7 and 13 from OrysaEULD1A showed higher seedling weight compared to the control plants after 150 mM NaCl treatment (Fig. 4A.5b). In addition, lines 3 and 9 from OrysaEULS2 showed a lower weight when severe NaCl stress (150 mM) was applied. After moderate (50 mM) mannitol treatment, all transgenic lines yielded seedlings with a higher weight compared to the wild type plants except for OrysaEULD1A (7) which showed a comparable weight to the control (Fig. 4A.5b). When seedlings were treated with a higher concentration of mannitol, lines 11 and 13 from Orysata, lines 9 and 11 from OrysaEULS2 and OrysaEULD1A (5) still showed a higher weight compared to seedlings of the wild type plants. After 50 μM ABA treatment, only seedlings of OrysaEULD1A (5) showed a higher weight compared to the control whereas seedlings from OrysaEULS2 lines 3 and 11 had higher weight after 100 μM ABA application (Fig. 4A.5b).

4A.4.4 Performance of adult plants from different transgenic lines in response to drought and NaCl

To elucidate the performance of transgenic lines overexpressing Orysata, OrysaEULS2 and OrysaEULD1A under drought and NaCl stress, assays were also performed on adult plants. All transgenic Orysata lines showed a significantly higher number of survived plants compared to the wild type plants after a drought period of 12 days as well as after 300 mM NaCl treatment (Fig. 4A.6). In the drought assay, transgenic lines 3 and 9 for OrysaEULS2 as well as lines 7 and 13 for OrysaEULD1A performed better compared to the wild type plants. After NaCl application, only OrysaEULS2 (9) and OrysaEULD1A (5) showed a significantly higher number of healthy plants compared to the control whereas the survival of OrysaEULD1A (7) plants was lower in comparison to the wild type plants (Fig. 4A.6).
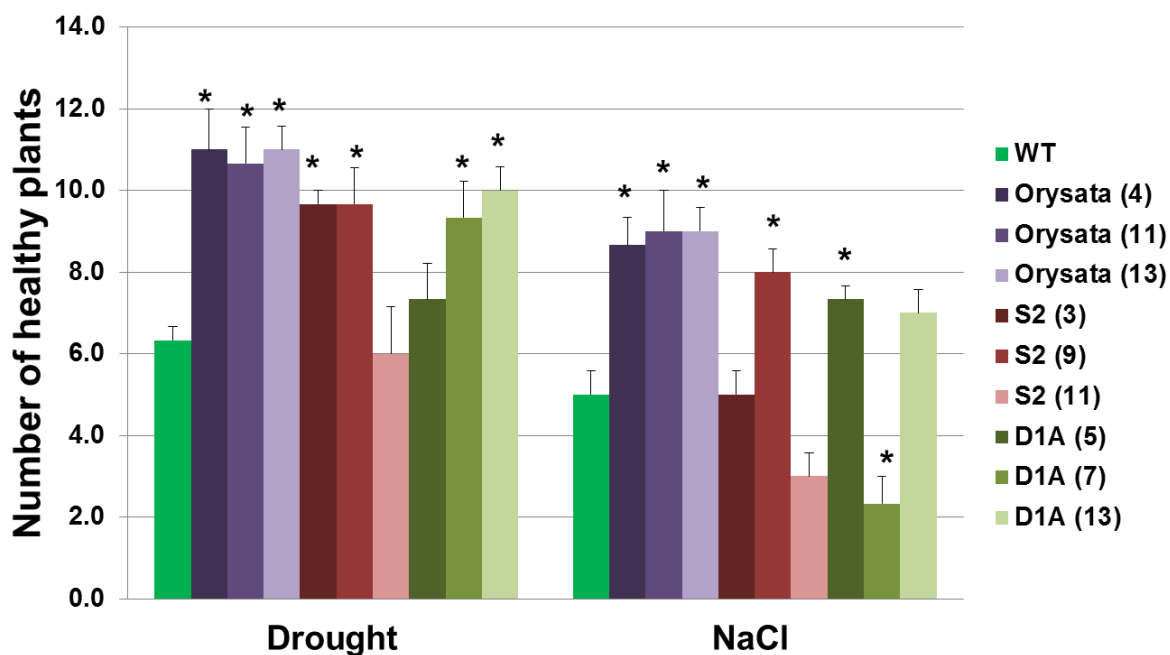
Fig. 4A.6 Drought and NaCl stress assays on adult plants. Bars represent means ± SE based on three replicates with 12 plants per replicate. WT: wild type, Ory: Orysata; S2: OrysaEULS2 and D1A: OrysaEULD1A, numbers between brackets indicate different transgenic lines. Asterisks refer to significant differences in comparison with the control at P-value < 0.05

4A.4.5 Responsiveness of different transgenic lines to *Pseudomonas* infection

To determine whether the transgenic lines overexpressing Orysata, OrysaEULS2 and OrysaEULD1A are more tolerant towards biotic stress, a bacterial pathogen *Pseudomonas*

syringae (*Pst*DC3000), which is virulent to *Arabidopsis* (Whalen et al., 1991), was selected. Infection symptoms of *Pst*DC3000 were observed on all transgenic lines from Orysata and OrysaEULS2 as well as on wild type plants on day 6 after infection of the plants. In contrast, 35% and 47% of OrysaEULD1A lines 5 and 13, respectively, remained healthy whereas the infection symptoms were seen on all OrysaEULD1A plants of line 7 (Fig. 4A.7). However, these findings have to be confirmed by repeating this experiment.
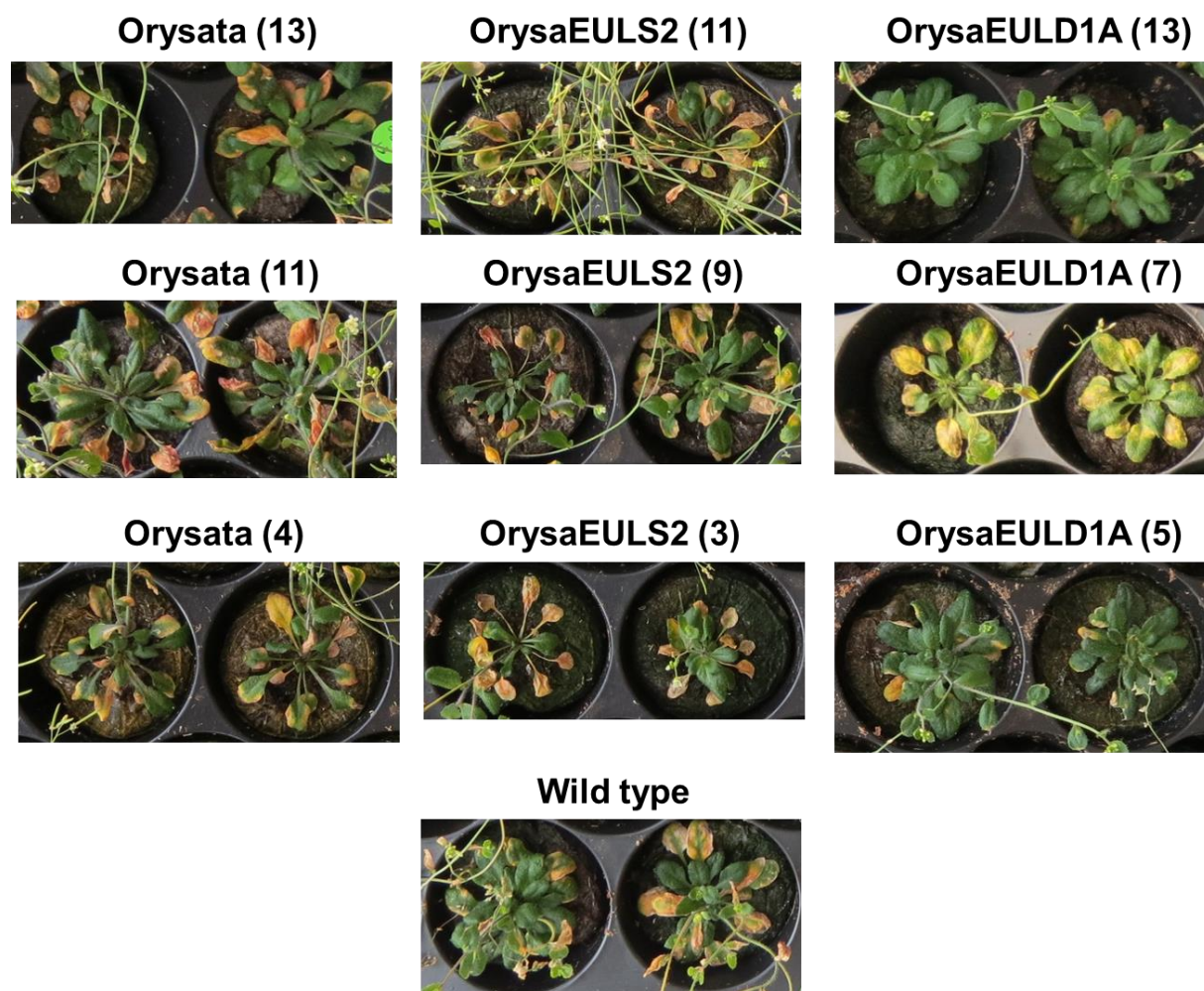
Fig. 4A.7 *Pseudomonas* infection symptoms on transgenic lines as well as on the wild type plants. Four week old plants were infected with *Pseudomonas* syringae strain DC3000 and pictures were taken 6 days after spraying

4A.5 Discussion

Several lectins have been described as stress related proteins (Zhang et al., 2000; Chen et al., 2002; Lannoo et al., 2006b; Yong et al., 2003). Since these lectins are expressed at low concentrations even after application of the stress factor, it has been suggested that these inducible lectins might play a signaling role when the plant is subjected to stress (Van Damme et al., 2008).

Rice is a very important crop because it feeds half of the world population and it serves as a crop model. In an attempt to increase the tolerance of rice (as a model for cereal crops) against the most important abiotic and biotic stresses, the effect(s) of overexpression of one lectin belonging to the jacalin-related family and two lectins belonging to the EUL family were investigated on plant performance in stress conditions. Several reports in literature have shown that overexpressing stress related genes from rice in *Arabidopsis* can increase their tolerance to abiotic and biotic stresses. For example, overexpression of topoisomerase 6 homologs (Jain et al., 2008), OsSMCP1 encoding a small protein with a single C2 domain (Yokotani et al., 2009) and OsMSR2 encoding a novel calmodulin-like protein gene (Xu et al., 2011) yields *Arabidopsis* lines with a better performance in stress situations. In this study, Orysata, OrysaEULS2 and OrysaEULDA1 were ectopically expressed in *Arabidopsis* and subjected to various abiotic and biotic stresses. These stresses were applied during seed germination, but also on the seedling and the adult stages. The performance of the transgenic lines was variable most probably due to the position effect of the transgenes in the *Arabidopsis* genome. However, despite this variability, some of the transgenic lines clearly increased the tolerance of the *Arabidopsis* plants towards abiotic and biotic stresses especially when stress was applied to adult plants (Figs. 4A.6 and 4A.7).

None of the proteins under study has a clear nuclear localization signal. Localization studies revealed that OrysaEULS2 and OrysaEULD1A are located to the nucleus (including the nucleolus) and the cytoplasm of the plant cell which is in agreement with the observation that both proteins are synthesized without a signal peptide. This localization pattern was previously also reported for EEA, the prototype of the EUL family, and the EUL type S3 from *Arabidopsis* (Van Hove et al., 2011).

The performance of all transgenic lines in the different growth and stress experiments was summarized in Table 4A.1. Morphometric analysis revealed that transgenic lines which developed a higher biomass in the early stages of development finally yielded plants that were shorter, showed a delayed flowering time and yielded less seeds (Fig. 4A.1, Table 4A.1). The short stature of the transgenic lines might be due to investing more energy to over-express the transgene compared to being used for the overall growth of the host plant. Only transgenic lines expressing OrysaEULS2 developed less biomass, had a longer stem and flowered earlier compared to the wild type plants. The overexpression of the different lectins under study produced pleiotropic effects on the growth and development of transgenic *Arabidopsis* plants, including a short(er) stature, more expanded leaves and earlier bolting. These observations suggested a possible role of these lectins in different developmental processes including flowering. Indeed, earlier studies on the promoter activity of OrysaEULD1A also suggested a role of the lectin in the expansion and cell division during the early stages of tissue development as well as in flower and silique development (Al Atalah et al., 2013; see chapter 3A). Similar pleiotropic phenotypes have previously been reported after overexpression of the topoisomerase 6 subunit A homolog from rice (Jain et al., 2008).

Under normal growth conditions, all transgenic lines expressing EULs showed a lower germination rate compared to the wild type plants whereas the germination rate for all Orysata transgenic lines was comparable to that of the control plants. As a response to different abiotic stresses, the seed germination for all transgenic lines was inhibited suggesting a stress hypersensitivity of the transgenic plants (Fig. 4A.4) especially for the lines expression OrysaEULD1A and to a lesser degree for the lines expressing OrysaEULS2 as well as Orysata (Table 4A.1). Transgenic lines overexpressing OrysaEULS2 and OrysaEULD1A exhibited hypersensitivity to ABA during seed germination, dwarfed in the early stages of growth and bolted earlier compared to the wild type plants (Table 4A.1). These phenotypes have been reported previously for the *Arabidopsis* lines that overexpress well-known ABA-dependent genes like DREB2A, a drought responsive element B2A (Sakuma et al., 2006) and AtMYB44, a transcription factor related to stress response (Jung et al., 2008) suggesting a possible role for OrysaEULS2 and OrysaEULD1A in ABA-mediated responses to abiotic stresses such as drought and salinity. These results are in line with other reports showing

that overexpression of genes, conferred an increased sensitivity to ABA on seed germination and post germination stages, could increase stress tolerance e.g. CAZFP1, a pepper zinc-finger protein gene (Kim et al., 2004a), OsbZIP23, a member of the basic leucine zipper transcription factor family (Xiang et al., 2008) and OsMSR2 (Xu et al., 2011). However, no hypersensitive effect was observed on seed germination of the Orysata lines indicating a possible role of this lectin in non ABA- mediated responses to drought and salinity.

Root elongation is a good indicator for salinity and drought tolerance in plants (Maiti 2012a; 2012b; 2012c). On seedling level, moderate concentrations of mannitol caused a positive effect on both root elongation and the weight of the seedlings for all OrysaEULS2 lines as well as one line for Orysata and OrysaEULD1A. In contrast, most of the transgenic lines showed increased tolerance towards drought and salinity when tested on adult level (Fig. 4A.6). It is generally accepted that plants normally show different tolerance levels in different developmental stages. In literature several proteins have been shown to enhance stress tolerance especially on seedling level (Jain et al., 2008; Xu et al., 2011).

After *Pseudomonas* infection, only lines 5 and 13 from OrysaEULD1A showed tolerance symptoms (Fig. 4A.7). These two transgenic lines also showed increased tolerance towards drought and salinity suggesting a possible role for this lectin in signaling after biotic and abiotic stresses. This dual tolerant role against abiotic and biotic stresses was reported by overexpressing some stress related signaling factors such as pepper CAZFP1 (Kim et al., 2004a) and barley HvRAF (Jung et al., 2007) which conferred tolerance to both stresses in *Arabidopsis*.

It can be concluded that the lectins under study can confer tolerance to drought and salinity. Based on our results we hypothesize that the EUL lectins might be involved in ABA-mediated responsiveness towards abiotic and biotic stresses. Conversely, Orysata caused tolerance to *Arabidopsis* plants via ABA non-dependent reactions towards abiotic stresses.

Arabidopsis was used in our study as a heterologous system for overexpression of the lectins mainly because of the ease of plant transformation, growth and analysis. Although several stress related rice genes have been successfully tested in *Arabidopsis* plants, showing increased tolerance towards stresses when the proteins were overexpressed in *Arabidopsis*, we cannot exclude that different results might be obtained if our lectins are overexpressed

in rice plants. Differences in stress tolerance have also been reported between rice and *Arabidopsis* plants overexpressing the *Arabidopsis* DREB1A. Overexpression of DREB1A in *Arabidopsis* resulted in improved tolerance to drought, salinity and freezing stress but with severe plant growth retardation under normal conditions of growth (Kasuga et al., 1999). On the other hand, overexpression of the same gene in rice plants improved tolerance to drought and salinity but only to a very little extent to cold stress (Oh et al., 2005). Furthermore, overexpression of some stress related genes from rice conferred resistance towards abiotic stresses when expressed in the homologous (rice) and the heterologous (*Arabidopsis*) plant systems such as OsiSAP8, a member of stress associated proteins (Kanneganti and Gupta 2008).

It is also noteworthy that *in vivo* plants are usually exposed to different stresses simultaneously and therefore there is a need to analyze the performance of transgenic lines when multiple stress factors are applied.

Table 4A.1 Performance of the transgenic lines under all tested conditions compared to the wild type. ~: wild type or comparable to the wild type; + and -: higher and lower compared to the wild type, respectively. WT: wild type, Ory: Orysata; S2: OrysaEULS2 and D1A: OrysaEULD1A, numbers between brackets indicate the different transgenic lines.

Experiment	Criterion	Wt	Ory(4)	Ory(11)	Ory(13)	S2(3)	S2(9)	S2(11)	D1A(5)	D1A(7)	D1A(13)
Morphometric analysis	Leaf number	~	+	+	~	+	+	~	~	~	+
	Leaf length	~	+	+	~	+	+	-	~	~	+
	Leaf width	~	+	+	~	+	+	-	~	~	+
	Petiole length	~	+	+	~	+	+	-	~	~	+
	Height	~	-	-	-	-	-	+	~	-	-
	Seed weight	~	-	-	~	-	-	~	~	-	-
	Bolting time	~	+	+	~	+	+	-	~	+	+
Germination	No stress	~	~	~	~	-	-	-	-	-	-
	100 mM NaCl	~	~	~	~	-	-	-	-	-	-
	150 mM NaCl	~	~	~	~	~	~	~	-	-	-
	50 mM Mannitol	~	~	-	-	~	-	-	-	-	-
	100 mM Mannitol	~	~	-	~	~	-	-	-	-	-
	50 µM ABA	~	~	~	-	-	-	-	-	-	-
	100 µM ABA	~	~	~	~	~	~	-	-	-	-
Root length /seedlings	No stress	~	-	-	~	-	-	-	-	-	-
	100 mM NaCl	~	~	~	+	~	~	~	~	~	~
	150 mM NaCl	~	~	~	~	~	~	~	~	~	~
	50 mM Mannitol	~	~	~	+	+	+	+	+	~	~
	100 mM Mannitol	~	~	~	~	-	~	~	-	~	~
	50 µM ABA	~	~	~	+	~	~	-	-	-	~
	100 µM ABA	~	~	~	+	~	~	~	~	~	~
Seedling weight /seedlings	No stress	~	+	+	+	-	-	~	~	-	-
	100 mM NaCl	~	+	+	+	~	-	~	-	+	+
	150 mM NaCl	~	+	+	~	-	-	~	+	+	-
	50 mM Mannitol	~	+	+	+	+	+	+	+	~	+
	100 mM Mannitol	~	-	+	~	-	+	+	+	-	~
	50 µM ABA	~	~	~	~	-	~	-	+	-	-
	100 µM ABA	~	~	-	-	+	-	+	-	-	+
Abiotic assays /adult plants	Drought	~	+	+	+	+	+	~	~	+	+
	NaCl	~	+	+	+	~	+	~	+	-	~
Biotic assay /adult plants	*Pseudomonas* infection	~	~	~	~	~	~	~	+	~	+

Chapter 4

Performance of transgenic lines over-expressing OrysaEULS2, OrysaEULD1A and Orysata after biotic and abiotic stress application

Chapter 4B

Insecticidal activity of Orysata, a jacalin-related lectin from rice, against biting-chewing and piercing-sucking insects

Manuscript submitted

Al Atalah B, Smagghe G, Van Damme EJM (2013) Insecticidal activity of Orysata, a jacalin-related lectin from rice, against biting-chewing and piercing-sucking insects.

4B.1 Abstract

In the past two decades research has focused on the insecticidal activity of lectins with affinity towards mannose, particularly lectins with a GNA domain. The present study reports the insecticidal activity of Orysata, a lectin from rice with mannose specificity, belonging to the family of jacalin-related lectins. So far, only two reports have described an effect for jacalin-related lectins. The effect of Orysata was investigated against three important pest insects in agriculture: the beet armyworm *Spodoptera exigua* Hübner (Lepidoptera: Noctuidae), and two aphid pests: peach-potato aphid *Myzus persicae* Sulzer and pea aphid *Acyrthosiphon pisum* (Hemiptera: Aphidoidea).

Bioassays with *S. exigua* and *M. persicae* were performed using detached leaves from transgenic tobacco lines overexpressing Orysata. The expression levels ranged between 38-71 µg/g FW, corresponding with 0.6-1.1% of total soluble protein. Intoxicated larval stages of *S. exigua* revealed significant mortality, and reductions in larval weight and a retardation of development. Similarly, feeding on leaves expressing Orysata at ≥57 µg/g caused 63% mortality of the peach-potato aphids. When pea aphids were fed on an artificial diet supplemented with different amounts of recombinant Orysata, mortality was high at relatively low lectin concentrations, the estimated 50% lethal concentration being 79 µg/ml.

In conclusion, our results demonstrated that the jacalin-related lectin Orysata is posing a strong insecticidal activity, suggesting that it can be considered as a valuable candidate to be used as a control agent against both biting-chewing and piercing-sucking pest insects.

4B.2 Introduction

The *Oryza sativa* agglutinin, abbreviated as Orysata, was the first inducible mannose-binding jacalin-related lectin detected in rice seedlings after NaCl treatment (Zhang et al., 2000). Recently, Orysata was shown to be located in the nucleus and the cytoplasm of the plant cell. SDS⁄PAGE and western blot analysis showed that the recombinant lectin exists in two molecular forms: a 23 kDa lectin glycosylated polypeptide and an 18.5 kDa unglycosylated polypeptide. Glycan array analyses revealed that Orysata interacts with high-mannose as well as with more complex N-glycan structures (Al Atalah et al., 2011; see chapter 2A).

The harmful effect of chemical insecticides on non-target organisms and the environment urged researchers to develop alternative compounds through the application of genetic engineering technologies (Ranjekar et al., 2003). The production of these so-called bio-insecticides is often based on the ectopic expression of genes encoding e.g. toxins, such as the bacterial delta-endotoxins from *Bacillus thuringiensis* (Sharma et al., 2004). Despite the successful use of this new bio-insecticide several drawbacks were also raised in the past few years. Among them are the safety concern for the consumers (Shelton et al., 2002) and the growing resistance of insects to the genetically modified crops (Sanchis and Bourguet 2008). Furthermore, *Bacillus* toxins do not affect sucking pest insects such as mirids, thrips, bugs, hoppers and aphids, and thus cannot be used to target a large part of the insect population (Malone et al., 2008). To overcome these problems, researchers focus on getting a better understanding of plant defense mechanisms. Indeed, plants develop complex defense mechanisms that allow them to successfully tolerate unfavorable conditions, including attack by insects. Therefore, the best option to enhance the tolerance of important crops is the use of insecticidal plant proteins, such as protease inhibitors and lectins, through transgenic-based pest control strategies (Jouanin et al., 1998).

Since insect glycans are mostly of the high-mannose type (Aoki et al., 2007; Schachter 2009), mannose binding lectins have received a lot of attention. In the past two decades the insecticidal properties of several lectins with mannose specificity were investigated in detail, in particular lectins with a carbohydrate recognition domain similar to the *Galanthus nivalis* (snowdrop) agglutinin

(abbreviated as GNA) (Van Damme et al., 1987). GNA itself was shown to exhibit insecticidal activity towards hemipterans, such as *Myzus persicae* (Hilder et al., 1995; Down et al., 2006) and *Acyrthosiphon pisum* (Rahbé et al., 1995), but also towards Lepidopterans, such as *Lacanobia oleracea* (Fitches et al., 2004). Similarly, the leaf and bulb lectins from garlic (*Allium sativum* L.) were shown to have insecticidal properties against the cotton leafworm *Spodoptera littoralis* (Sadeghi et al., 2007). Insect bioassays with the glucose/mannose-binding legume lectin Con A showed an effect against the pea aphid *A. pisum* (Sauvion et al., 2004a,b).

Next to plant lectins with a GNA domain or a legume lectin domain, other carbohydrate-binding domains with specificity towards high-mannose N-glycans are present especially within the family of lectins with a jacalin domain. Although all these lectins recognize mannose they definitely differ in the three-dimensional structure of the carbohydrate recognition domain, and hence in their fine specificity towards sugars. Due to the lack of available proteins, very few studies have been performed for the insecticidal activity of mannose-binding jacalin-related lectins. In the current study, the objective was to analyze the insecticidal activity of Orysata, a jacalin-related lectin from rice, against three important pest insects in agriculture with one biting-chewing Lepidopteran and two piercing-sucking aphids: the beet armyworm (*Spodoptera exigua*), the peach-potato aphid (*M. persicae*) and the pea aphid (*A. pisum*). Therefore, Orysata was overexpressed constitutively in tobacco plants under the control of cauliflower mosaic virus 35S promoter. Subsequently, bioassays with *S. exigua* and *M. persicae* were conducted on detached leaves from different transgenic lines as well as from wild type plants. In parallel experiments were performed with an artificial diet containing increasing concentrations of the purified recombinant Orysata. Our results point at the possibility of using Orysata as a pest control agent against biting-chewing and piercing-sucking insects.

4B.3 Methods and Materials

4B.3.1 Plant material and growth conditions

Nicotiana tabacum L. var. Samsun (NN) was used in this study. To obtain an *in vitro* culture of tobacco, dry seeds were surface sterilized in 95% ethanol (v/v) for 5 min

and subsequently in 5% NaOCl (v/v) for 15 min followed by thorough rinsing using sterile water. Afterwards, the sterilized seeds were sown on Murashige and Skoog (1962) medium (MS) (4.3 g/l MS micro- and macronutrients containing vitamins (Duchefa, Haarlem, The Netherlands), 30 g/l sucrose, pH 5.7 (adjusted with 0.5 M NaOH), and 8 g/l plant agar (Duchefa)) in glass jars (10 cm diameter). Plants aged four to five weeks were used for leaf disc transformation.

For the insect bioassays the sterilized seeds from wild type plants and different independent transgenic lines (further referred to as lines 5, 8, 10, 11 and 13) were sown on MS medium (containing 100 µg/ml kanamycin for the transgenic lines) in 9 cm petri dishes. After two weeks, 20 plants for each line were transferred to pot soil (20 cm diameter pots). All plants were grown in a growth chamber at 25°C with a 16/8 h light/dark photoperiod. Plants aged six weeks (and older) were used for the bioassays.

4B.3.2 Binary vectors construction

The coding sequence for Orysata (GenBank accession number CB632549) was amplified by PCR using the cDNA clone (provided by Arizona Genomics Institute: University of Arizona, Tucson, USA) encoding Orysata as a template. In order to achieve constitutive expression of Orysata the sequence was cloned in the expression vector PK7WG2.0 (Karimi et al., 2002) following the Gateway technology (Invitrogen, Carlsbad CA, USA) as described before (Al Atalah et al., 2011; see chapter 2A). Briefly, the *attB* sequences were added to the flanking regions of the Orysata coding sequence by conducting two consecutive PCR reactions using the primers EVD 545 5'AAAAAGCAGGCTTCACCATGACGCTGGTGAAGATTGGCCT3' /EVD 546 5'AGAAAGCT GGGTGTCAAGGGTGGACGTAGATGCCAATTGC3' for the first PCR reaction and EVD 2 5'GGGGACAAGTTTGTACAAAAAAGCAGGCT3' /EVD 4 5'ACCACTTT GCTCAAGAAAGCT GGGT3' for the second PCR reaction. Afterwards, the BP reaction was performed to clone the Orysata sequence into the entry vector pDONR221 (Invitrogen). Subsequently, cloning of the sequence was confirmed by sequencing the inserted fragment using the primers Donr-F 5'TCGCGTTAACGCTAGCATG3' and Seql-E 5'GTTGAATATGGCTCATAACAC3' (performed by LGC Genomics, Berlin, Germany). Finally, the LR reaction was performed to clone the coding sequence of

Orysata under the control of 35S cauliflower mosaic virus promoter in the destination vector.

4B.3.3 Tobacco transformation

The binary vector carrying the 35S::Orysata construct was mobilized into the *Agrobacterium* strain LBA4404 by tri-parental mating, using a protocol adapted from Hoekema et al. (1983). This method was carried out using *Escherichia coli* top10F cells harboring the 35S::Orysata construct (spectinomycin resistant) as the donor strain, *E. coli* DH5α with HB1001 pAK2013 (kanamycin resistant) as the helper strain and *Agrobacterium tumefaciens* LBA4404 (gentamycin resistant) as the recipient. Transformants were selected on YEB medium (consisting of 5 g/l beef extract, 5 g/l peptone, 5 g/l sucrose, 1 g/l yeast extract, and 15 g/l bacteriological agar) containing 50 μg/ml spectinomycin and 20 μg/ml gentamycin. Afterwards, the presence of 35S::Orysata construct was confirmed by PCR analysis (using primers EVD 545/EVD 546) on the plasmids extracted from several transformants. The LBA4404 strain carrying the 35S::Orysata construct was used to transform tobacco following the leaf disc transformation procedure (Horsch et al., 1985). Selection of transgenic plantlets was conducted on MS-medium containing 300 mg/l kanamycin and 100 mg/l carbenicillin.

4B.3.4 Preparation of crude extracts

Using a mortar and pestle, tobacco leaves were crushed in 20 mM un-buffered 1,3-diaminopropane using 0.5 ml buffer per g fresh weight (FW) as described previously (Sadeghi et al., 2009a). Afterwards, extracts were transferred to eppendorf tubes and centrifuged for 5 min at 13,000 g using a table micro-centrifuge and then the supernatant was transferred into new tubes. Subsequently, the protein content of the leaf extract was estimated using the Coomassie (Bradford) Protein Assay Kit (Thermo Fischer Scientific, Rockford, IL, USA) based on the Bradford dye-binding procedure (Bradford 1976) using a plant lectin (*Robinia pseudoaccacia* agglutinin) as a standard. Finally, the extracts were stored at -20°C until use.

4B.3.5 PCR analysis

PCR analysis was performed on tobacco generations 0 and 1 (T0 and T1). The genomic DNA was extracted from tobacco leaves of independent transgenic lines as well as from wild type plants following the cetyl trimethyl ammonium bromide method as described previously (Lannoo et al., 2006a). To check the quality of the genomic DNA, the Nictaba (*Nicotiana tabacum* agglutinin) sequence was amplified using the primers EVD 1 5'AAAAAGCAGGCTTCACCATGCAAGGCCAGTGGATAGCCGC3' and EVD 3 5'AGAAAGCTGGGTGTTAGTTTGGACGAATGTCGAAGCC3'. To confirm the presence of the Orysata sequence in the transgenic lines, PCR was conducted using Orysata specific primers (EVD 545/EVD 546). The following conditions were used: 2 min at 94°C followed by 25 cycles of 15 sec at 94°C, 30 sec at 55°C and 60 sec at 72°C, and a final incubation for 5 min at 72°C. The PCR amplified fragments were visualized on a 1.5% agarose gel (Invitrogen). Only those T0 lines whose genomic DNA yielded a PCR fragment of the correct size were allowed to grow into adult plants and produce seeds.

4B.3.6 Western blot analysis

Protein analysis was conducted on leaf material of T0 and T1. Crude extracts were separated by SDS–PAGE using 15% acrylamide gels under reducing conditions as described by Laemmli (1970). Subsequently, the protein samples were electro-blotted onto polyvinylidene fluoride (PVDF, 0.45 µm) transfer membranes (Biotrace[TM]PVDF, PALL, Gelman Laboratory, Ann Arbor, MI, USA). The blots were blocked in Tris buffered saline (TBS: 10 mM Tris, 150 mM NaCl and 0.1% (v⁄v) Triton X-100, pH 7.6) containing 5% (w/v) milk powder for 1 h. After blocking, blots were incubated for 1 h with a rabbit anti-Orysata antiserum as primary antibody [produced by Thermo scientific based on injecting two rabbits with recombinant Orysata (Al Atalah et al., 2011; see chapter 2A)], diluted 1⁄1000 in TBS. The secondary antibody was a 1/5000 diluted rabbit anti-mouse IgG labeled with horseradish peroxidase (Thermo scientific). Immuno-detection was achieved by placing the blots in the detection buffer consisting of 0.1 M Tris-HCl pH 7.6 containing 0.7 mM 3,3-diaminobenzidine teterahydrochloride (Sigma-Aldrich, St Louis, MO, USA) and 0.01% (v/v) H_2O_2. All washes and incubations were conducted at room temperature with gentle shaking.

4B.3.7 Semi-quantification of Orysata content in transgenic lines

Agglutination assays were performed to semi-quantify the lectin content in the transgenic lines used in this study. Rabbit erythrocytes were prepared as described previously (Van Damme et al., 1987). Agglutination assays were performed in small glass tubes by mixing 10 µl protein extract for each transgenic line (or the wild type plant), 10 µl of 1 M ammonium sulphate and 30 µl of a 10% suspension of trypsin-treated rabbit erythrocytes. Agglutination was assessed visually after incubation for 45 min at room temperature. Based on a serially diluted solution of purified recombinant Orysata of known concentration, the minimum concentration that still caused agglutination was determined (Al Atalah et al., 2011; see chapter 2A) and used to estimate the absolute lectin content in the extracts. The semi-quantification was performed on three plants randomly chosen for each line and the whole experiment was repeated twice with different batches of plants.

4B.3.8 Insect rearing

The larval stages for the beet armyworm (*S. exigua*) were reared on an agar-based artificial diet (Hakim et al., 2006). Peach-potato aphids (*M. persicae*) were selected from a continuous stock colony maintained on wild type tobacco plants in the Laboratory of Agrozoology at Ghent University, Belgium (Shahidi-Noghabi et al., 2009). Pea aphids (*A. pisum*) were maintained on young broad bean plants (*Vicia faba* L) (Sadeghi et al., 2009b). To synchronize the age of the nymphs, mature aphids were put on a bean plant. After 24 h, all neonate nymphs were used in the bioassay as described before (Sadeghi et al., 2009c). Before and during the bioassays, all insects were maintained in growth chambers at 25 ± 2°C with 65 ± 5% relative humidity and a 16/8 h light/dark photoperiod.

4B.3.9 Insect bioassays on detached leaves

Bioassays with *S. exigua* and *M. persicae* were performed on detached leaves from five independent heterozygous transgenic lines (from T1), and the wild type plants as a control.

For the *Spodoptera* bioassay, cages were constructed using polyethylene containers (9 cm in diameter, 3 cm in height) as described by Sadeghi et al. (2007). Detached

leaves were placed in the experimental cages. Subsequently, 10 newly ecdysed 2nd instar larvae of *S. exigua* were placed on the leaf in each cage. Three cages were used per line and the whole experiment was repeated three times. Several parameters were scored daily during 12 days; the larval mortality, the development (the number of larvae for each instar) and the larval weight starting from the third instar. Fresh tobacco leaves were placed once every 2 days during the first half of the experiment and daily during the second half of the experiment, and excrements were removed from the cages daily.

For the *M. persicae* bioassay, the effect of Orysata on the mortality of the peach-potato aphid nymphs was estimated as described before (Hilder et al., 1995). The detached leaves were placed upside down on water saturated cotton wool in a plastic petri dish (15 cm diameter). Afterwards, wet papers were placed on the edges of each leaf to create a walking area for the aphids. In each plate 6 nymphs were put on the leaf and 2 plates were made for each tobacco line. The whole experiment was performed twice and the data for each experiment were accumulated. During three consecutive days the nymphal mortality of the aphids was scored.

4B.3.10 Insect bioassays on artificial diet

A standard diet previously developed for *A. pisum* (Febvay et al., 1988) was used as the basal food for the aphids. A procedure described by Sadeghi et al. (2009c) was followed to prepare and handle the liquid diet as well as the feeding apparatus used to feed the aphids. Under sterile conditions, 150 µl of the artificial diet was put between two parafilm layers to form a feeding sachet. To analyze the effect of Orysata on the neonate nymphs, several concentrations of the recombinant lectin (22, 126, 213 and 637 µg/ml) were used. Since the recombinant Orysata was dissolved in 20 mM un-buffered diaminopropane, an artificial diet supplemented with relevant volumes of this buffer was used as a control. At day 0, 10 neonate nymphs were transferred onto the artificial diet and during three consecutive days, the mortality was scored and the dead insects were removed daily. For each lectin concentration, three replicates were carried out.

4B.3.11 Statistical analysis

Data were expressed as mean ± standard error (SE). For the agglutination assay, data from the two biological replicates were accumulated. Percentage reduction in weight of the larvae was calculated using the formula: % weight reduction = [(C–T)/C] * 100, where C = weight of larvae feeding on control leaves and T = weight of larvae feeding on transgenic leaves (Huang et al., 2006). All decimal percentages and numbers were rounded to the nearest integer. The 50% lethal concentration (LC_{50}) together with the 95% confidence limits and the R^2 of the sigmoid curve fitting were determined using the non-linear regression analysis in Prism version 5 (GraphPad, La Jolla, CA). To separate the means statistically, a two-tailed T test was applied using the Prism version 5. For bioassays on detached leaves, differences of the means between the transgenic lines and the wild type plants were considered significant at a P-value = 0.05 or 0.01.

4B.4 Results

4B.4.1 Expression of Orysata in transgenic tobacco

Using the Gateway technology the coding sequence of Orysata was cloned into a binary vector so that the lectin sequence was under the control of 35S promoter. After mobilizing this binary vector carrying the 35S::Orysata construct into *Agrobacterium*, it was transferred to tobacco (*N. tabacum* cv Samsun NN) by means of leaf disc transformation. Several independent T0 lines were screened for the presence of the Orysata sequence at DNA level. Furthermore expression of the lectin was also analyzed at protein level. Only those transformants that showed good lectin activity were allowed to grow into adult plants and produce seeds. In the T1 generation, the integration of Orysata sequence in the genome as well as the expression of the lectin was checked for three randomly chosen plants (Fig. 4B.1a, b). Amplification of a 468 bp fragment after PCR analysis confirmed the presence of the Orysata coding sequence in the tobacco genome. Furthermore, the detection of a 15 kDa polypeptide (the exact size is 15.2 kDa) after western blot analysis confirmed that Orysata is successfully expressed in transgenic tobacco.

The Orysata content in the leaves of the transgenic lines under study was assessed semi-quantitatively using agglutination assays with rabbit erythrocytes. The highest

180

lectin content was observed in transgenic lines 8 and 11, being 71 ± 7.7 µg/g FW (1.1% of total soluble protein) and 57 ± 16.3 µg/g FW (0.9% of total soluble protein), respectively. In contrast the lectin content was lower and comparable in transgenic lines 5, 10 and 13, being 46 ± 7.7, 40 ± 8.3 and 38 ± 4.9 µg/g FW, respectively (corresponding to approximately 0.6% of total soluble protein).
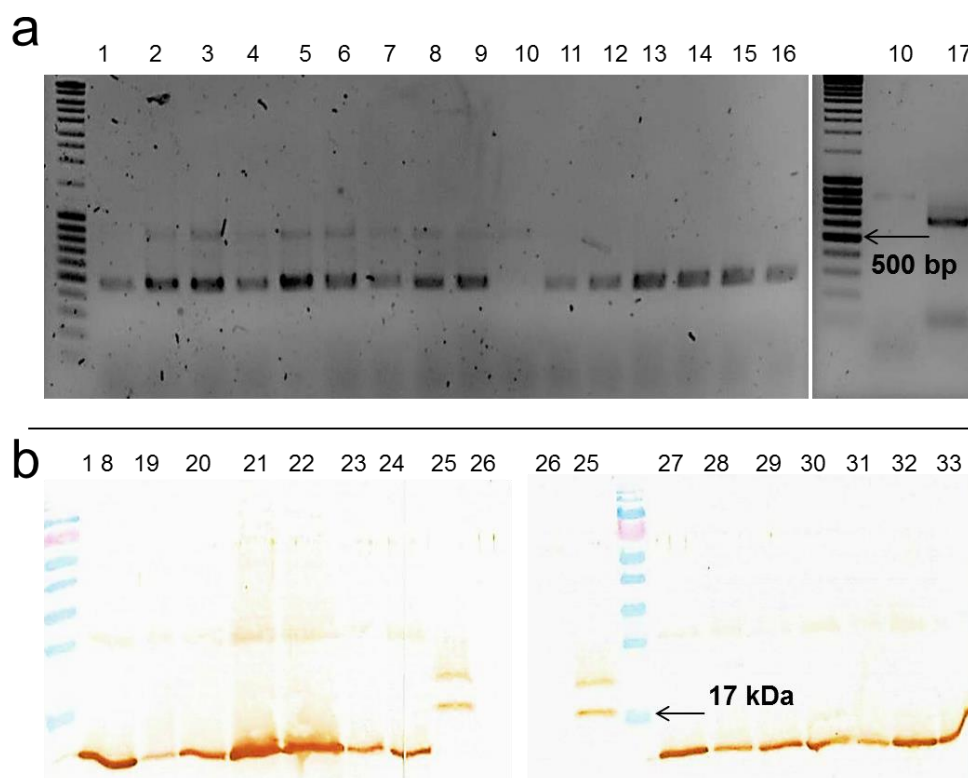
Fig. 4B.1 Analysis of the transgenic lines at DNA and protein level. Panel (a) represents gel pictures after PCR analysis on genomic DNA extracted from 3 randomly chosen plants per line using Orysata specific primers. Lanes 1-3 represent tobacco line 5; lanes 4-6 represent line 8; lanes 7-9 represent line 10; lane 10: wild type plant; lanes 11-13 represent line 11; lanes 14-16 represent line 13; lane 17: shows the Nictaba fragment amplified from wild type plant. To show the size of PCR amplicons, Mass RulerTM DNA Ladder Mix (Fermentas) was used. Expected sizes: 468 bp for Orysata and 684 bp for Nictaba. Panel (b) represents blot pictures after western blot analysis on protein samples extracted from 3 randomly chosen transgenic plants per line. Lanes 18-20 represent tobacco line 5; lanes 21-23 represent line 8; lanes 24, 27 and 28 represent line 10; lane 25: recombinant Orysata as a positive control; lane 26: wild type plant, lanes 29-31 represent line 11; lanes 32-33 represent line 13. Protein ladder (Fermentas, St Leon-Rot, Germany) was used in the blots. In each well, 30 µg protein was loaded from the crude extract and 75 ng from the recombinant lectin. Expected sizes: 15.2 kDa for the native Orysata, 18.5 and 23 kDa for the recombinant Orysata

4B.4.2 Effect of Orysata on *S. exigua* larval stages fed on transgenic lines

4B.4.2.1 *Lethal effects on S. exigua larvae*

To examine the lethal effect of Orysata on the larval stage, the mortality of larvae was assessed after feeding *S. exigua* on detached leaves from five independent transgenic lines expressing different levels of Orysata (Fig. 4B.2). Compared to wild type plants all transgenic lines caused higher larval mortality at 7, 9 and 12 days but only the larval mortality caused by line 11 was significantly different. After 7 days of feeding on line 11 the larval mortality was increased (29%) compared to feeding assays on the wild type plants (8%, P-value < 0.01). The larval mortality continued to be significantly higher (P-value < 0.05) on line 11, being 36% compared to 11% for the wild type plants on day 9. On day 12 larval mortality increased to 38% on line 11 compared to 14% for the wild type plants.
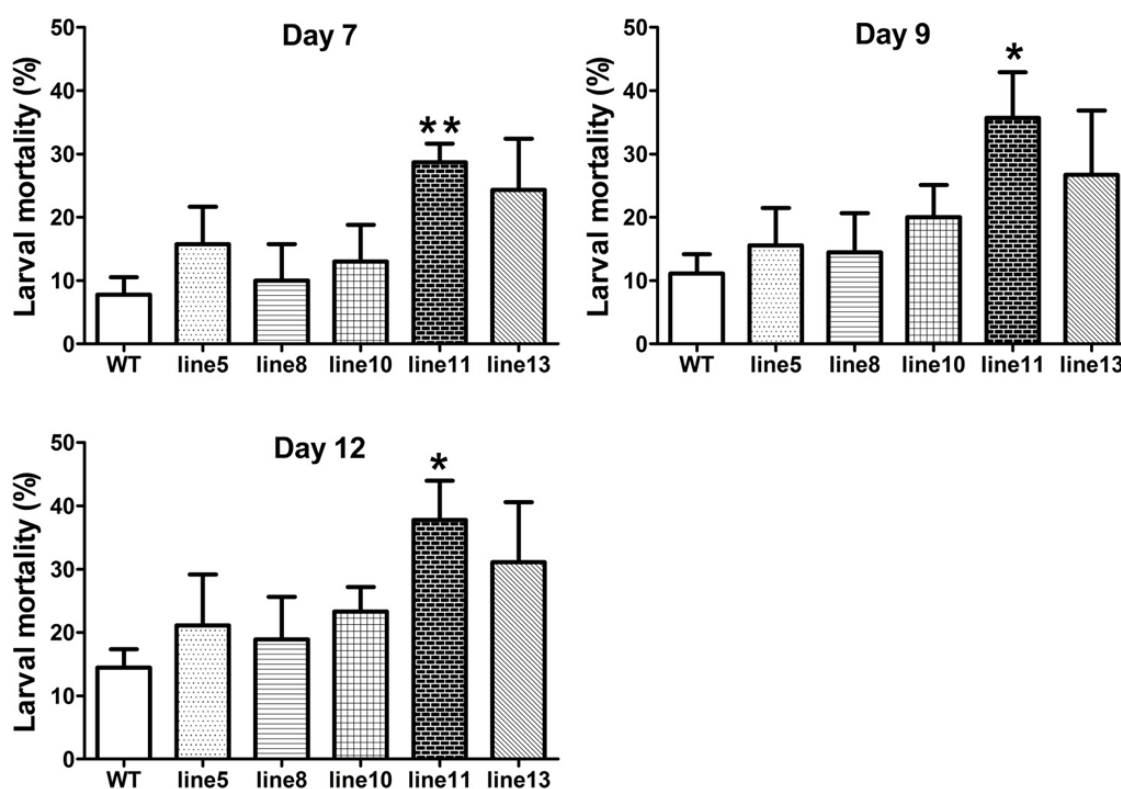
Fig. 4B.2 Larval mortality of *S. exigua* after feeding on detached leaves from five transgenic lines expressing Orysata. The larval mortality % after 7, 9, 12 days is shown. Bars represent means ± SE based on three independent replicates. One and two asterisks indicate the significant differences in comparison with the control at P-value < 0.05 and < 0.01, respectively

4B.4.2.2 *Sub-lethal effects on weight and development of S. exigua larval stages*

To analyze the sub-lethal effect of Orysata on *S. exigua* larvae, the weight was determined after 7, 9 and 12 days. A significant reduction in weight was observed for the second instar larvae fed with some of the transgenic lines (Fig. 4B.3). After 7 days, the weight of larvae fed on transgenic line 8 (39%, P-value < 0.01) and transgenic line 11 (51%, P-value < 0.01) was significantly reduced compared to larvae fed on the wild type plants. Similarly, a strong reduction in larval weight was also recorded for transgenic lines 8 and 11 after 9 days. After 12 days, the reduction of larval weight was still significant for the larvae fed on transgenic line 8 (49%, P-value < 0.05) and transgenic line 11 (47%, P-value < 0.01) compared to larvae fed on wild type leaves but for larvae fed on lines 5, 10 and 13, there was no significant weight reduction compared to the ones fed on the wild type plants.

Besides a strong weight reduction, retardation in larval development was also observed in response to feeding the larvae on detached leaves from the transgenic lines when compared to the control plants (Fig. 4B.4). After 7 days, 66% of larvae fed on wild type plants reached the fourth instar. However this percentage was much lower for larvae fed on leaves from the transgenic lines. Larvae fed on leaves of transgenic line 11 showed a significant reduction of 17% (P-value < 0.05) for fourth instar at day 7. With lines 8 and 11, at day 12 the percentage of fifth instar larvae was significantly reduced (P-value < 0.01) being 53% and 41%, respectively, compared to the wild type plants (94%). Lines 5, 10 and 13 showed no significant reduction in the number of larvae in the fourth instar (at day 7) and in the fifth instar (at day 12) compared with the larvae fed on control plants.
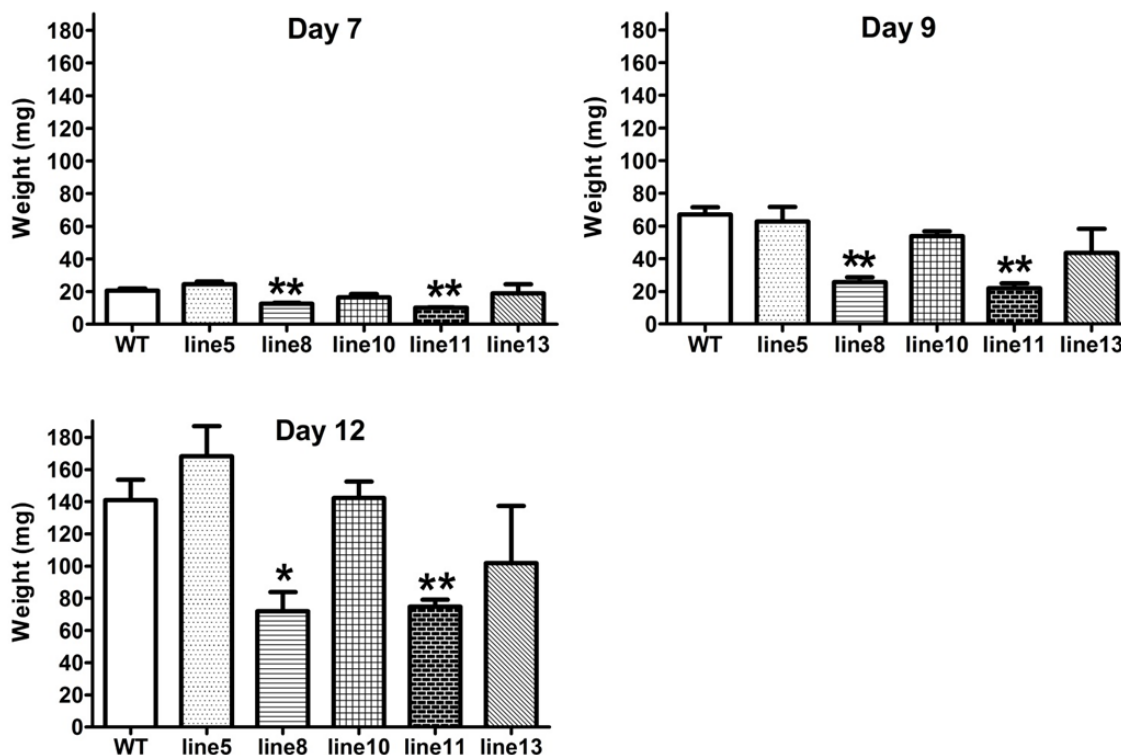
Fig 4B.3 Bioassay of *S. exigua* on transgenic lines. The effect of five transgenic lines expressing Orysata on the larval weight of *S. exigua* is shown after 7, 9, and 12 days. Bars represent means ± SE (of the average weight of one larva) based on three independent replicates. One and two asterisks indicate the significant differences in comparison with the control at P-value < 0.05 and < 0.01, respectively
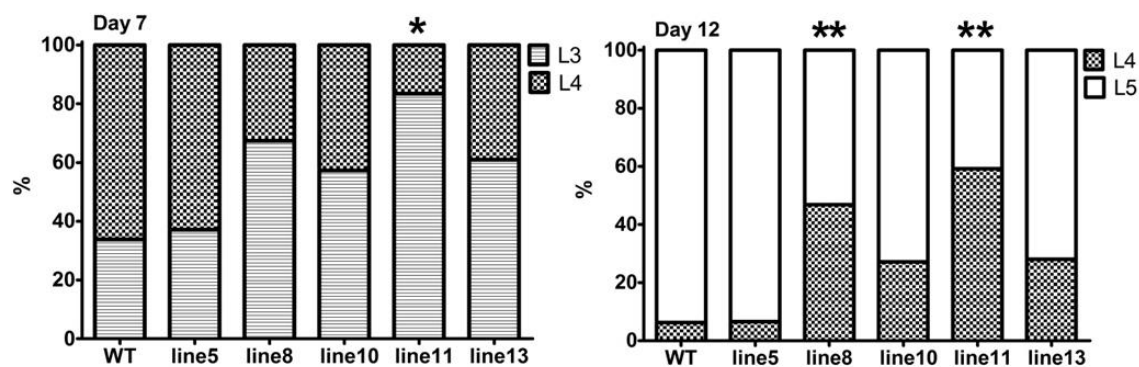
Fig 4B.4 Effect of Orysata on the development of *S. exigua* larvae. Larvae were fed on transgenic lines as well as the wild type plants. The percentages of instar 3 (L3) and instar 4 (L4) after 7 days as well as of instar 4 and instar 5 (L5) after 12 days are shown

4B.4.3 Effect of Orysata on *M. persicae* fed on transgenic lines

To analyze the lethal effect of the lectin for *M. persicae*, aphids were fed on the transgenic plants expressing Orysata. After three days the nymphal mortality was

analyzed. Mortality on the transgenic lines 8 (50%) and 11 (63%) was significantly higher compared to the control plants (13%) at P-value < 0.05 (Fig. 4B.5). Transgenic lines 5, 10 or 13 showed no significant difference in nymphal mortality compared to the wild type plants.
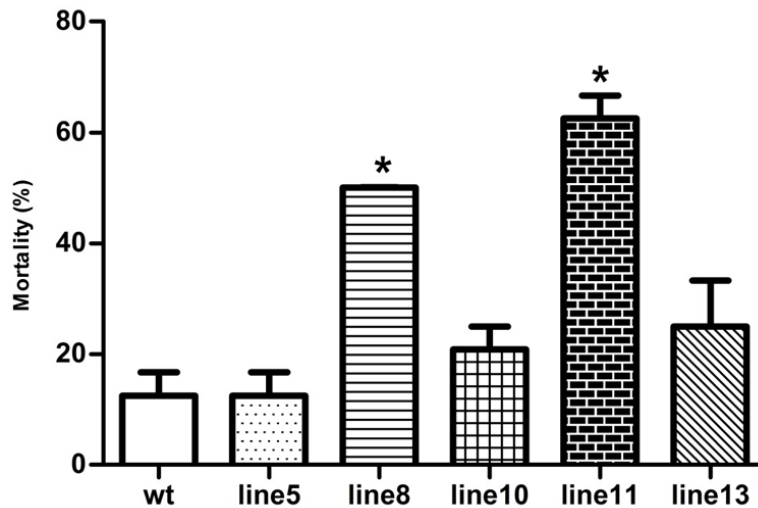
Fig 4B.5 Nymphal mortality of *M. persicae* after feeding on detached leaves from five transgenic tobacco lines over-expressing Orysata and from the wild type plants. Bars indicate the means ± SE of nymphal mortality after 3 days based on two independent replicates. One asterisk indicates the significant differences in comparison with the control at P-value < 0.05

4B.4.4 Insecticidal effect of Orysata on *A. pisum* fed on artificial diet

Neonate (<24 h) nymphs of *A. pisum* were fed on an artificial diet containing increasing concentrations of recombinant Orysata, ranging from 22, 126, 213 to 637 µg/ml. To test the insecticidal activity and determine the 50% lethal concentration of Orysata, the mortality of neonate nymphs was calculated. After 3 days, the mortality reached 23% in the assay with 22 µg/ml Orysata and this percentage increased up to 100% when 637 µg/ml Orysata was applied (Fig. 4B.6). The LC_{50} value was calculated to be 79 µg/ml (95% confidence limits: 54 to 117 µg/ml and $R^2 = 0.88$).
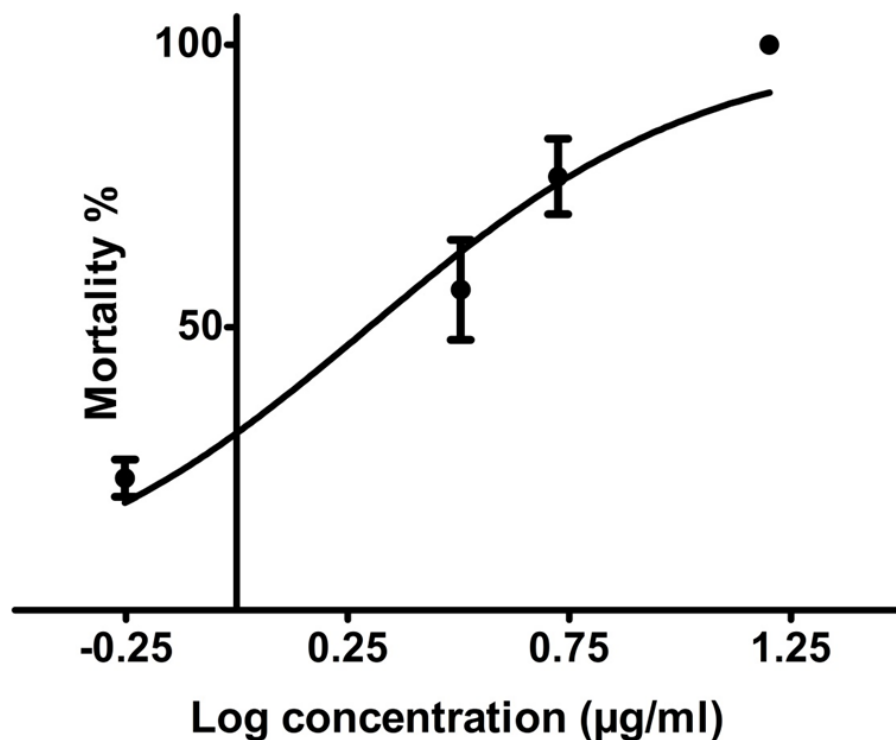
Fig 4B.6 Dose response curve of the nymphal mortality of *A. pisum* caused by feeding on recombinant Orysata for 3 days. Data are expressed as means ± SE based on three replicates with a total of 30 nymphs tested per concentration

4B.5 Discussion

Many of the classical lectins that accumulate in the vacuole are abundant proteins and have been shown to play a role in plant defense against pathogens and predators (Michiels et al., 2010; Lam and Ng 2011; Vandenborre et al., 2011b). In the last two decades several of these abundant plant lectins have been studied for insecticidal activity. Since N-linked glycan structures of the high-mannose type are most abundant in insects (Aoki et al., 2007; Schachter 2009), the focus of the studies with insects was largely on mannose-binding lectins.

At present, only few jacalin-related plant lectins have been studied for their activity on insects. Heltuba (*Helianthus tuberosus* agglutinin from *Jerusalem artichoke*) and HFR1 (Hessian fly responsive protein 1 from soft wheat) are inducible lectins with mannose specificity belonging to the family of jacalin-related lectins. Ectopically expressed Heltuba in tobacco provoked a decrease in the development and fecundity for the peach-potato aphid (*M. persicae*) (Chang et al., 2003). When

applied into the diet HFR1, recombinantly expressed in *E. coli*, showed strong insecticidal activity on the larval stage of the fruit fly (*Drosophila melanogaster*) (Subramanyamet al., 2008). In the present study, transgenic tobacco lines overexpressing a jacalin-related lectin from rice called Orysata were generated and challenged with a biting-chewing pest insect, *S. exigua*, and two piercing-sucking aphids, *M. persicae* and *A. pisum*. The bioassays with *S. exigua* and *M. persicae* were performed on detached leaves taken from the transgenic tobacco lines expressing Orysata whereas the assay with *A. pisum* was performed using an artificial diet containing increasing concentrations of the purified recombinant Orysata.

Using agglutination assays the lectin content in the different transgenic lines was calculated. A variable expression for Orysata was observed, with lectin concentrations ranging from 38 (line 13) to 71 µg/g FW (line 8). Different copies of the transgene as well as its integration position in the plant genome might be responsible for the differential expression profile of Orysata in the different transgenic tobacco lines.

The beet armyworm *S. exigua* is a dangerous pest because it has a wide range of hosts and damages important cultivated crops such as beet, tomato, cotton, corn, soybean, peanut, and pepper (Moulton et al., 2000; Senthil-Nathan et al., 2008). Once an infestation is well established, the armyworm larvae are difficult to control with insecticides (Wang et al., 2006a). Feeding of 2nd instar larvae of *S. exigua* on transgenic lines expressing Orysata revealed a significant lethal effect for transgenic line 11, causing 38% larval mortality compared to 14% in the control after 12 days (Fig. 4B.2). Although the transgenic lines showed limited lethal effects on larval mortality, they did cause clear sub-lethal effects on the larval weight and development. Throughout the experiment, tobacco lines 8 and 11 showed a significant reduction in larval weight compared to the control treatment, being 49% and 47%, respectively, after 12 days (Fig. 4B.3). Similarly, the number of fifth instar that evolved on transgenic lines 8 and 11 was significantly reduced after 12 days being 53 and 41%, respectively, compared to 94% on the wild type plants (Fig. 4B.4). Our results showed a clear correlation between the expression level of Orysata in the transgenic line and the insecticidal effect, since lines 8 and 11 with the highest concentration of Orysata (71 and 57 µg/g FW, respectively) always showed

significant differences compared to the wild type plants. Transgenic lines 5, 10 and 13 with lower lectin content showed no significant differences in larval mortality, weight reduction or development retardation. Sub-lethal effects on larval weight reduction and development retardation were previously also reported for mannose-binding lectins of the GNA type. The mannose-binding lectin from leek (APA, *Allium porrum* agglutinin) showed significant larval weight reduction of *S. littoralis* fed on transgenic tobacco lines (expressing APA with a concentration ranging from 23 to 91 µg/g FW), being 15 to 27% compared to the control after 11 days (Sadeghi et al., 2009b). In addition, the mannose-binding lectins from garlic leaves (ASAL *Allium sativum* leaf agglutinin) and bulbs (ASAII *Allium sativum* bulb agglutinin II) had a significant negative effect on the larval weight of *S. littoralis* fed on transgenic tobacco expressing ASAL and ASAII, being 42% and 30% respectively after 5 days (Sadeghi et al., 2008b). Only 23% reduction in weight was observed for the tomato moth *L. oleracea* when fed on excised leaves of transgenic potato expressing GNA (0.07% of total soluble protein) (Fitches et al., 1997). Compared to our data, the sub-lethal effects caused by Orysata were stronger than the effect of the GNA-related lectins.

Peach-potato aphid *M. persicae* is an important polyphagous pest in agriculture. This pest causes damage to crops by direct feeding and it also serves as a vehicle for more than hundred plant viruses in about thirty different families including many major crops (Kennedy et al., 1962). Similar to the caterpillar assay, transgenic lines 8 and 11 (expressing the highest concentration of Orysata) showed a significant reduction of the nymphal mortality after 3 days, being 50% and 63%, respectively compared to 13% for the wild type plants (Fig. 4B.5). This early deleterious effect on the aphid nymphs reflects a strong activity of Orysata. In general, the inhibitory effect of mannose-binding lectins such as GNA and ConA caused a reduction in fecundity and retarded development of the aphid nymphs when tested in tobacco (Hilder et al., 1995; Zhang et al., 2007), wheat (Stoger et al., 1999), potato (Gatehouse et al., 1999) and maize (Wang et al., 2005). However, the survival of the aphids fed on transgenic tobacco expressing Heltuba was not significantly affected after 11 days whereas the fecundity of the aphids was negatively affected (Chang et al., 2003). None of the transgenic lines expressing ASAL (at 0.02-0.03% of total

soluble protein) or ASAII (at 1.48-2.21% of total soluble protein) had any effect on nymphal survival of *M. nicotianae* after 7-8 days (Sadeghi et al., 2007). Later, transgenic plants expressing ASAL and ASAII did show a significant effect on the reproduction capacity of the resulting adults. Similarly, GNA expressed ectopically in maize (at 0.13-0.28% of total soluble protein) had no significant effect on the corn leaf aphid (*Rhopalosiphum maidis*) survival after 16 days but it showed a significant reduction in fecundity (Wang et al., 2005). However, when ASAL was expressed in chickpea plants (0.08-0.38% of total soluble protein), it showed negative effect on the cowpea aphid (*Aphis craccivora*) survival after 3 days (Chakraborti et al., 2009). Based on these observations, it can be concluded that Orysata has a fast and significantly negative effect on the survival of nymphs whereas most tested GNA-related mannose-binding lectins (when ectopically expressed) showed no effect on nymphal (or adult) survival for sap-sucking insects.

The pea aphid *A. pisum* is a cosmopolitan pest which damages the legume species including the forage and vegetable crops (Van Emden and Harrington 2007). Recombinant Orysata, expressed and purified from *Pichia pastoris* (Al Atalah et al., 2011; see chapter 2A), was tested at different concentrations supplemented into the artificial diet (Fig. 4B.6). The low LC_{50} value being 79 µg/ml, indicates that Orysata had a strong negative effect on *A. pisum* compared with other mannose-binding lectins. After 3 days, the LC_{50} for the aphids fed on medium containing GNA was 350 µg/ml and for the *Alium sativum* agglutinin (ASA) an LC_{50} value of 700 µg/ml was calculated. When 800 µg/ml lectin was supplemented to the artificial diet, the mortality was 76% for the GNA treatment, whereas 64% mortality was observed in the ASA treatment and only 12% mortality after APA treatment (Sadeghi et al., 2009b). In contrast 77% and 100% mortality was recorded when 213 µg/ml and 637 µg/ml, respectively, of recombinant Orysata was applied in the diet.

In rice there are two major pests, the rice brown plant hopper (*Nilaparvata lugens*) and the rice green leafhopper (*Nephotettix nigropictus*). These sap-sucking insects cause severe physiological damage to the rice plant and serve as vectors for different viruses including the tungro bacilliform virus, the grassy stunt virus and the ragged stunt virus (Saxena and Khan 1989). Bioassays based on artificial diet highlighted the toxicity of plant lectins against these insects (Powell et al., 1993), GNA being the

most toxic lectin decreasing brown plant hopper survival by 50% at a concentration of 200 µg/ml (Powell et al., 1995). In addition, GNA expressed in the transgenic rice plants decreased survival and overall fecundity of the insects, retarded insect development, and had a deterrent effect on brown plant hopper feeding (Rao et al., 1998). Furthermore, transgenic rice expressing ASAL reduced the survival and fecundity of brown plant hopper to 36% and 32%, respectively. Similarly, the survival and fecundity of green leafhopper were down to 40.5% and 29.5%, respectively, compared to control plants (Saha et al., 2006).

Our findings demonstrated that Orysata has strong insecticidal activity against both biting-chewing and piercing-sucking pest insects. We believe therefore that Orysata can help rice to cope with insect attack. However, the concentration of Orysata in rice is very low, reaching 1 µg/g FW of rice leaves exposed to 0.1 M NaCl whereas it was estimated to be 1000 times lower in the control plants (Zhang et al., 2000). Since mannose-binding lectins showed strong insecticidal activity towards brown plant hopper and green leafhopper, Orysata might be considered a good candidate to be introduced into transgenic rice to provide resistance against these sap sucking insects. Furthermore, Orysata provides a biological alternative to chemical insecticides using a transgenic based approach.

Chapter 5

General conclusions and perspectives

5.1 Inducible plant lectins

The discovery of the inducible plant lectins (Van Damme et al., 2004 a) directed our attention towards a study of the physiological role of these lectins in plants. Interestingly these inducible lectins are expressed in response to several stress conditions, such as drought, high salt, wounding, hormone treatment, and pathogen attack (Lannoo and Van Damme 2010; Vandenborre et al., 2011b; Jiang et al., 2010a). Since the lectin expression is up-regulated after stress treatment, the hypothesis was put forward that the inducible lectins might have a role in the stress physiology of the plant cell (Van Damme et al., 2004 b, 2008, 2011).

Rice (*Oryza sativa* L.) is one of the most important food crops in the world supplying 20% of daily calories to more than 3.5 billion people worldwide (Khush 2013; World Rice Statistics, http://www.irri.org). As for other crops, rice production is adversely affected by a wide range of abiotic and biotic stresses and this in turn is threatening the food security on a global scale.

In rice, several lectins belonging to different families are regulated after abiotic and biotic stresses (This work, Jiang et al., 2010a). The main goal of this PhD research was to study the possibility of exploiting some of these inducible rice lectins to improve the tolerance of crops against external stresses. Lectins from two families were selected for this study, in particular, Orysata a lectin discovered in rice seedlings exposed to NaCl treatment (Zhang et al., 2000) belonging to the jacalin-related family as well as OrysaEULS2 and OrysaEULD1A, both lectins classified in the family of EUL-related lectins (Fouquaert et al., 2009a).

5.2 Characterization of proteins under study

The first objective of this work was to characterize Orysata, OrysaEULS2 and OrysaEULD1A. In **chapter 2**, the molecular structure, biological activity and carbohydrate binding specificity for all proteins under study were analyzed in detail. Because the inducible proteins are expressed at low concentrations, it is very difficult to purify sufficient amounts of the lectin directly from the plant tissues. To overcome this problem, *Pichia pastoris* was used as a heterologous system to

produce the recombinant proteins. The carbohydrate binding specificity was refined for Orysata (**chapter 2A**) and the EUL domain of OrysaEULS2 (**chapter 2B**) using a glycan array and revealed that Orysata and OrysaEULS2 showed high affinity towards high-mannose structures. In plants, the N-glycans present on glycoproteins are rich in mannose residues (Man5-9GlcNAc2) but also contain other sugars like fucose and xylose (Gomord et al., 2005). In order to get deeper insight about the possibility for our proteins to bind to plant glycoproteins, it would be interesting to analyze our recombinant proteins on a plant glycan array which has been recently become available (Pedersen et al., 2012).

Recently, evidence has accumulated that environmental stresses such as salinity, drought, cold stress, or exposure to metals such as cadmium and zinc can change the protein glycosylation patterns (Komatsu et al., 2009; Štefanić et al., 2012) indicating the important role for glycoproteins in response to external stresses. The most common type of glycosylation on nucleocytoplasmic proteins is the addition of O-GlcNAc which is present in many different proteins located in the nucleus, nuclear pore complexes, cytosol and cytoskeleton (Lefebvre et al., 2010; Olszewski et al., 2010). Inducible lectins most probably interact with glycans present in the nucleus and the cytosol. These glycans are present either in a free form as a consequence of degradation of mis-folded proteins or in a bound form when they are attached to the surface of lipids and proteins residing in the same compartment. Therefore, we suggest analyzing the interaction of the rice lectins with the intracellular free N-glycans reported from rice cells (Maeda and Kimura 2006; Maeda et al., 2010). Analyses can be done with glycans obtained from cells grown under normal conditions as well as under stress conditions. This analysis will shed the light on the glycosylation changes and their role in the response to stress treatments.

Chapter 2B describes the construction of a three dimensional model built for the EUL domain in OrysaEULS2 . Mutational analysis revealed the importance of a tryptophan residue located in the predicted carbohydrate binding site as an essential residue for the carbohydrate activity of the EUL domain from rice. These results have to be confirmed experimentally by X-ray crystallography which is a strong method to determine the three-dimensional structure of proteins. To perform this analysis the

expression of the EUL domain from OrysaEULS2 in *P. pastoris* can be scaled up which will allow obtaining sufficient amounts of ultra-pure protein for crystallization. Resolving the three dimensional structure will be beneficial to determine the overall structure of the EUL domain in general and the carbohydrate binding site in particular.

Although amino acids predicted to be responsible for the formation of the carbohydrate binding site are identical for all EUL domains in rice (Fig. 2C.4), their fine carbohydrate binding specificity is different. Whereas OrysaEULS2 showed affinity to high-mannose structures (**chapter 2B**), OrysaEULD1A exhibited clear preferences towards galactosylated structures (**chapter 2C**). It is therefore important to perform mutational analysis by altering one or a few amino acids located in the vicinity of the carbohydrate binding site to explore which of these amino acids are responsible for the promiscuity of the sugar specificity of the EUL domain.

Until now no binding partners have been reported in rice for our proteins. Searching well known rice networks such as Database of Interacting Proteins in *Oryza Sativa* (DIPOS, http://csb.shu.edu.cn/dipos/?id=5) and *Oryza sativa* Protein-Protein Interactions Network (PRIN, http://bis.zju.edu.cn/prin/) yielded no results. However, to gain better insight into the physiological importance of our proteins, it is very important to identify their receptors in the nucleus and the cytoplasm since our proteins are localized to both cell compartments (**chapter 4A**). To perform this analysis, the recombinant proteins can be used as baits to catch the putative interacting partners through the application of pull down or co-immunoprecipitation assays. Afterwards, these binding partners can be characterized at the biochemical level and identified by amino acid sequencing and/or mass spectrometry.

5.3 Functional analyses under stress conditions

The second objective of this work was to analyze the expression of the carbohydrate-binding proteins from rice in the presence of stress applications. In **chapter 3**, this goal was covered through investigating the promoter activity for the EUL genes under study (**chapter 3A**), and conducting Q-PCR experiments to study the

expression of all studied proteins (**chapter 3B**) in the presence or absence of different stress treatments.

It is generally accepted that plants show different levels of sensitivities towards stress factors based on the developmental stage as well as the intensity of the stress that was applied. The promoter activity for OrysaEULS2 and OrysaEULD1A was analyzed during the seedling stage and after stress treatment at only one concentration for each stress factor. Therefore, it would be interesting to extend the promoter::GUS experiments by analyzing the GUS activity for more developmental stages and under weak, moderate as well as strong levels of stresses. In addition, analysis of GUS expression under the control of OrysaEULS2 and OrysaEULD1A promoters can be investigated in rice plants to study the promoter activity in their original genetic context. At a later stage, several deletions of the EUL promoters can be analyzed in an attempt to determine the most important stress responsive elements in these promoter sequences.

Our Q-PCR results indicate that rice expresses multiple lectins in response to the most important stresses that the rice culture is confronted with in its natural habitat. The up-regulation of EUL expression by 1 to 3 fold after application of abiotic stresses suggests a possible role for these proteins in stress signaling. Therefore, it would be very beneficial to complete this research by identifying the putative interacting partners. Tandem affinity purification (TAP) is an efficient system for protein complex purification and protein interaction identification (Xu et al., 2010). Purification of protein complexes by TAP was demonstrated to be effective in rice (Rohila et al., 2006). Hence, TAP constructs can be cloned and transformed into rice cells (or plants) to identify the binding partners for our proteins. In a further step, the interaction between the EUL-related proteins and these interacting partners should be confirmed by an independent technique, such as e.g. bimolecular fluorescence complementation (Kerppola 2008) where each of the proteins of interest and its putative partner can be fused to a non-functional half of a reporter protein. If the interaction occurs, the reporter protein will regain its activity, allowing us to visualize the interaction with the lectin partner.

OrysaEULS2 and OrysaEULD1A are encoded by sequences containing an N-terminal unrelated sequence next to the EUL domains. One can speculate that after translation into a chimeric protein this N-terminal domain might have a role in protein-protein interaction. Indeed after performing BLAST (http://blast.ncbi.nlm.nih.gov/) searches with these N-terminal domains (including also the linker sequence for OrysaEULD1A), several hits have been retrieved. Some of these hits were annotated in the NCBI database as stress related proteins in *Zea mays* and *Triticum aestivum*. As such, TAP should also be performed to determine protein partners for these N-terminal protein domains.

To show the importance of the protein-carbohydrate interaction involving EUL-related proteins, it would be interesting to use the EUL mutant protein, deficient in carbohydrate binding activity as a tool for some functional analyses and localization studies. Therefore native and mutant lectin constructs can be transformed into rice cells and their functionality can be compared under normal as well as under stress conditions. Furthermore, the localization patterns for both constructs can be also compared.

Q PCR analyses after biotic stress assays showed that some of the EUL proteins as well as Orysata are putative targets for rice pathogens (especially *Magnaporthe oryzae*). It is well known that ABA is a key signaling molecule in plant-pathogen interactions (Xu et al., 2013). This ABA signaling provoked by the pathogen suppresses salicylic acid (SA) and/or ethylene (ET) signaling pathways and consequently alleviates rice defense responses (Bailey et al., 2009; Jiang et al., 2010b). Therefore, it would be interesting to test the expression and regulation of our proteins after SA and ET treatments.

5.4 Tolerance towards abiotic and biotic stresses

The third objective of this research was to evaluate the performance of transgenic lines over-expressing Orysata, OrysaEULS2 and OrysaEULD1A under different abiotic and biotic stresses.

In **chapter 4A** the performance of transgenic *Arabidopsis* overexpressing the proteins under study was analyzed after abiotic and biotic stress application. Our results showed that some of the transgenic lines altered the overall growth of *Arabidopsis* plants and conferred tolerance of the plant towards the stress treatment. As a consequence, it would be interesting to overexpress our proteins in their homologous background (rice) and study the effect of overexpression of the lectins under stress conditions. Constitutively active promoters are not always desirable because overexpression of a transgene may compete for energy and building blocks for synthesis of proteins that are also required for plant growth under normal growth conditions (Kathuria et al., 2007). Several inducible and / or tissue specific promoters are available and have been studied in rice, which provides a way to express the transgene very specifically in the target tissue of interest. Expression of genes for stress tolerance could be engineered by deploying stress-inducible promoters. Many such promoters have been studied in rice (summarized in Kathuria et al., 2007). As such, our lectins can be expressed under the control of one of the stress inducible promoters in order to develop more efficient tolerant rice lines.

Chapter 4B describes the role of Orysata as a mannose binding lectin against biting-chewing and piercing-sucking insects when ectopically expressed in tobacco plants. Our results revealed that Orysata shows lethal as well as sub-lethal effects on the larval stages of *Spodoptera exigua* through reduction of larval weight and retardation of larval growth. Since the last instar larvae have to reach a certain fresh weight before entering the pupal stage (Davidowitz et al., 2003), it is important to study the effect of Orysata on pupation. Several reports in literature showed the abnormalities induced by some lectins on pupa as well as adult insects (Sadeghi et al., 2009a). Hence it will also be interesting to find out if Orysata will interfere with pupal development.

Although the aphid experiments were only performed on a short term they already showed some clear effect of Orysata on the survival of nymphs of *Myzus persicae* and *Acyrthosiphon pisum*. It will also be interesting to investigate the effect of Orysata on the fecundity and the adult survival through performing long term

experiments. All our experiments were conducted *in vitro*; it would be interesting to extend these analyses to evaluate the effect of Orysata *in vivo*. Finally, to continue this research, we suggest investigating the effect of Orysata on some important rice pests, such as the rice brown plant hopper (*Nilaparvata lugens*) and the rice green leafhopper (*Nephotettix nigropictus*). Initially these analyses could be performed on an artificial diet and later this lectin can be introduced into rice plant. Afterwards, the effect of Orysata expressed in rice can be evaluated *in vitro* and *in vivo*.

Until now, no EUL-related lectin has been tested for insecticidal activity. Hence it is important to test some EUL lectins against insects in short as well as long term experiments. This investigation could be interesting for the analysis of yet another lectin that preferentially interacts with mannose (e.g. OrysaEULS2), but could also be extended to other EULs interacting with galactose structures (e.g. OrysaEULD1A).

5.5 Mode of action of nucleocytoplasmic lectins

At present we can only speculate on the possible mode of action of the nucleocytoplasmic lectins in response to abiotic stresses since no potential receptors have been identified yet in the plant cell. After exposure of rice plants to stress conditions, a stress signal will be transmitted to the nucleus and will be recognized by some stress responsive elements present in the promoter sequences of the lectins under study, which in turn will influence the expression level of these lectins. Taking into account that the lectin sequences do not contain a signal peptide it is reasonable to assume that the lectins are synthesized on free ribosomes in the cytoplasm. It was unambiguously shown that the different EUL rice lectins as well as Orysata are partially translocated to the nucleus. Judging from the research work with ArathEULS3 from *A. thaliana* evidence has also been obtained for transport of the EUL lectin to the plasma membrane and the cell wall (J. Van Hove, unpublished data). Because of their carbohydrate binding activity these lectins can interact with different glycoconjugates, including glycoproteins, glycolipids, free glycans and polysaccharides forming lectin-glycoconjugate complexes. Since the different EULs and Orysata clearly differ from each other in their carbohydrate binding specificity multiple interactions can take place. The complexes formed might be part of a signal

transduction or regulatory mechanism recruited by the plant cell to cope with the stress effect.

The latter assumption is reasonable at least for the nucleocytoplasmic lectins that occur in very low concentrations in the cell. However, as mentioned before some nucleocytoplasmic lectins are present at high concentrations in storage tissues, such as the *Calystegia sepium* rhizomes and *Helianthus tuberosus* tubers. These storage tissues are a rich source of polysaccharides, in particular starch and other sugars like glucose, sucrose and fructose in *Calystegia* and inulin, a polymer of the monosaccharide fructose, in *Helianthus*. At present we cannot exclude that one of these polysaccharides (or simple sugars) represent the real target of the lectins which are present in these storage tissues. Until now, very few data are available describing the mode of action of nucleocytoplasmic plant lectins against biotic stresses such as fungi, bacteria and viruses. Plant lectins might have a role in the recognition of the pathogen by interacting with the glycan structures present on the surface of these pathogens, and this recognition might trigger some pathways that could help the plant to cope with pathogen attack. Bacterial glycoconjugates, such as lipopolysaccharides from the outer membrane of Gram-negative bacteria and peptidoglycan from both Gram-positive and Gram-negative bacteria (Erbs., 2008; Molinaro et al., 2009) as well as fungal glycoconjugates such as oligosaccharides derived from the polysaccharides chitin and glucan of the cell wall (Shibuya and Minami 2001) have been found to act as elicitors of plant innate immunity. Therefore, these bacterial and fungal glycoconjugates might represent the real targets for the nucleocytoplasmic lectins during the recognition process.

Binding studies using glycan arrays have shown that several nucleocytoplasmic lectins have a strong affinity for high-mannose glycans that are frequently present on insect proteins. These glycans structures occur on many digestive enzymes or transport proteins secreted in the midgut of insects or proteins embedded in the epithelial cell membrane. Therefore, these glycoproteins are potential targets for plant lectins. Indeed, it has been shown that ferritin (an important protein for iron transport) acts as a target site for the mannose binding snowdrop lectin (GNA) in the midgut of the cotton leafworm *Spodoptera littoralis* (Sadeghi et al., 2008a) and in the

midgut of the rice brown planthopper *Nilaparvata lugens* (Du et al., 2000). In addition, a membrane-bound aminopeptidase enzyme, a glycoprotein rich in mannose residues, was also targeted by both GNA and ConA (jackbean lectin) (Cristofoletti et al., 2006). Furthermore, the chitin microfibrils of the peritrophic membrane (Hakim et al., 2010) of the insect midgut might be a target for Nictaba (the leaf lectin from *Nicotiana tabacum*) since this nucleocytoplasmic lectin specifically interacts with oligomers of GlcNAc and chitin (Vandenborre et al., 2010a). Therefore, plant lectins probably act on insects through the interaction with multiple target glycoproteins (Vandenborre et al., 2010b). This is true for pest insects such as beetles, aphids and caterpillars, and for non-target insects such as honeybees (Vandenborre et al., 2011a). Secreted glycoproteins that are targeted by plant lectins in the midgut will tend to cluster, resulting in large complexes with a molecular size that is too big to pass through the peritrophic membrane. These macromolecular insect protein-lectin complexes may prevent larval enzymes to diffuse back across the peritrophic membrane for being recycled in the digestive system, resulting in leakage of digestive enzymes. Obviously, this will have a negative impact on the nutritional system of the insect.

In conclusion, the first objective of this PhD research representing the characterization of Orysata, OrysaEULS2 and OrysaEULD1A was achieved. Our results showed for the first time that the EUL proteins are true lectins with a promiscuous carbohydrate-binding specificity. The second objective, namely the expression analysis of the EUL proteins as well as Orysata in response to stress treatments was also accomplished. Based on our results, it can be concluded that rice expresses multiple carbohydrate binding proteins in response to external stresses. We hypothesize that our proteins play a role in sensing the external stress and behave as stress signaling elements. In addition, some of our proteins are targeted by rice pathogens suggesting a very important role for these proteins in pathways required during defense aginst these pathogens. The third objective of this study namely, evaluation of the performance of transgenic lines over-expressing Orysata, OrysaEULS2 and OrysaEULD1A under different abiotic and biotic stresses was also achieved to a large extent, but need to be extended in the future. We could show

that over-expression of our proteins confers *Arabidopsis* plants tolerance against some environmental stresses. Taking all these results together, we believe that this work contributed significantly to a better understanding of the role of inducible lectins from rice in the stress physiology of the plant. This understanding will help in unrevealing the mechanisms and pathways that inducible nucleocytoplasmic lectins are involved in and in turn can be used to improve the stress tolerance of crops.

Summary/ Samenvatting

Summary

Plant lectins group all proteins that possess at least one non-catalytic domain which recognizes and binds reversibly certain carbohydrate structures. These carbohydrate-binding proteins are widespread in the plant kingdom. In the past, research was focused on lectins that are expressed at high concentrations especially in storage tissues where they locate in the vacuolar or extracellular compartment of the plant cell. For many of these lectins it was shown that they play a role in plant defense. During the last decade evidence has accumulated that plants also synthesize certain carbohydrate-binding proteins after exposure to a variety of abiotic and biotic stresses. These "inducible" proteins are present in the nucleus and the cytoplasm of the plant cell. Consequently, the inducible proteins are suggested to be involved in the carbohydrate-protein interactions taking place inside the plant cell and therefore fulfill a signaling role in stress physiology.

In the frame of this PhD thesis, several inducible proteins of rice were selected. *Oryza sativa* agglutinin abbreviated as Orysata, is an inducible protein that belongs to the family of jacalin-related lectins. Orysata consists exclusively of the jacalin domain with a molecular weight (MW) of 15.2 KDa. This lectin was detected in rice seedlings after exposing them to NaCl treatment. In addition to Orysata, two proteins belonging to the family of EUL-related lectins were investigated. OrysaEULS2 represents the S type of the EUL family and consists of an EUL domain preceded by a 56 AA unrelated N-terminal sequence with a total MW of 22.7 kDa. OrysaEULD1A represents the D type of the EUL family and consists of two almost identical EUL domains (91% sequence similarity and 72% sequence identity) tandemly arrayed, separated by a 23 AA linker domain and preceded by an 19 AA N-terminal sequence with a total MW of 38.5 kDa.

In the **first chapter**, an overview of current literature on lectinology and the role of lectins in response to the most important abiotic and biotic stresses is given.

To study the possibility of using the selected proteins in increasing the stress tolerance of crops, we first had to characterize the lectins. **Chapter 2** describes the characterization of Orysata, OrysaEULS2 and OrysaEULD1A. Since the inducible

proteins are expressed at very low concentrations and only after the application of stress, it is difficult and in most cases even impossible to purify sufficient amounts of these proteins from plant tissues. Therefore, the heterologous system, *Pichia pastoris*, was used to express all proteins under study. Subsequently the recombinant proteins were purified using a combination of ion exchange chromatography and metal affinity chromatography on a Ni–Sepharose column. In **chapter 2A**, glycan array analysis showed that Orysata interacts with high-mannose as well as with more complex N-glycan structures. In **chapter 2B**, a three-dimensional model was constructed for the EUL domain within OrysaEULS2. Subsequently, mutation analysis for some important amino acids revealed that the tryptophan residues play a role in the formation of the carbohydrate binding site. OrysaEULS2 shows affinity towards high-mannose structures as concluded from glycan array analysis and thus it was proven for the first time that OrysaEULS2 is a true lectin. In **chapter 2C**, OrysaEULD1A was shown to have lectin activity since it agglutinated red blood cells and its carbohydrate binding activity was directed towards galactose related compounds.

To gain more knowledge with respect to the promoter sequences of the EUL lectins, the promoter activity of OrysaEULS2 and OrysaEULD1A was investigated. In **chapter 3A**, the GUS gene was cloned under the control of each of the EUL promoters and transformed into *Arabidopsis* plants. The GUS staining pattern was analyzed after different abiotic stress treatments and during several developmental stages. GUS activity was detected for the OrysaEULD1A promoter in the young true leaves but not for the OrysaEULS2 promoter. It was shown that the activity of OrysaEULD1A promoter was increased after ABA and mannitol treatments but decreased after NaCl treatment.

To further investigate the role of the EUL proteins, Q-PCR experiments were conducted on rice seedlings exposed to several abiotic and biotic stresses (**chapter 3B**). The Q-PCR analysis revealed that the expression of the whole set of rice EUL genes is stress related. After ABA and NaCl treatments, all EULs as well as Orysata were up-regulated in the roots. A more variable expression was observed in the shoots after biotic stress treatments. In response to *Xanthomonas oryzae* infection,

only EULs from the D type (except OrysaEULD1A) as well as Orysata were up-regulated at 4 and 8 dpi, whereas no effect was observed on the expression of EULs from the S type. Furthermore, our results showed that *Magnaporthe oryzae* targeted the expression of some of the EULs, in particular OrysaEULS3, OrysaEULD1A and OrysaEULD2 as well as Orysata.

In an attempt to analyze the possibility that EULs (and Orysata) might increase the tolerance against abiotic and biotic stresses, the lectins under study were over-expressed and the performance of the transgenic lines overexpressing the lectins was investigated after different selected stresses. **Chapter 4A** describes the overexpression of Orysata, OrysaEULS2 and OrysaEULD1A in *Arabidopsis* plants and the selection of one-insertion homozygous lines. Since plants show different levels of tolerance based on the stress intensity and the developmental stage, the experiments were performed during different developmental stages in particular germinated seeds, seedlings and adult plants subjected to moderate and severe levels of stress. Our results indicate that the lectins under study can confer tolerance to drought and salinity especially on the adult plants. In **chapter 4B**, the coding sequence of Orysata was expressed constitutively in tobacco plants. Afterwards, five heterozygous lines were challenged against biting-chewing (*Spodoptera exigua*) and piercing-sucking (*Myzus persicae* and *Acyrthosiphon pisum*) insects. Transgenic tobacco lines overexpressing Orysata induced mortality and caused reductions in larval weight as well as a retardation of the development of *S. exigua*. In addition, Orysata showed lethal effects on nymphal survival of *M. persicae and A. pisum*. Our findings demonstrated that Orysata has strong insecticidal activity suggesting a possible role for this lectin as a control agent against both biting-chewing and piercing-sucking pest insects.

In **chapter 5**, the results of the whole PhD research are concluded and perspectives to continue the work are proposed. Overall, this PhD research considerably increased our understanding on the role of the nucleocytoplasmic lectins in response to external stresses. Our results point out that Orysata, OrysaEULS2 and OrysaEULD1A can confer tolerance against some important abiotic and biotic stresses that crops normally encounter. Despite that, further work is still needed to

confirm the role of these lectins in the response to stress treatments and investigate the possibility for using these lectins to help crops cope with external stresses through the application of transgenic based strategies.

Samenvatting

Plantlectines omvatten alle eiwitten die minstens één niet-katalytisch domein bezitten dat bepaalde koolhydraatstructuren kan herkennen en er op reversibele wijze mee kan binden. Deze koolhydraat-bindende eiwitten zijn wijdverspreid in het plantenrijk. In het verleden was het onderzoek toegespitst op lectines die voornamelijk in opslagweefsels in hoge concentraties tot expressie gebracht worden, en zich daar in de vacuole of het extracellulair gedeelte van de plantencel bevinden. Voor vele van deze lectines is aangetoond dat ze een rol spelen in de verdediging van de plant. In de loop van het laatste decennium werd aangetoond dat planten ook bepaalde koolhydraat-bindende eiwitten synthetiseren na blootstelling aan verschillende vormen van abiotische en biotische stress. Deze "induceerbare" proteïnen bevinden zich in de nucleus en het cytoplasma van de plantencel. Bijgevolg wordt vermoed dat deze induceerbare proteïnen een rol spelen in de koolhydraat-eiwit interacties die plaatsvinden in de plantencel en dus een signaalfunctie vervullen in de stressfysiologie.

In het kader van dit doctoraatsproefschrift werden verscheidene induceerbare proteïnen geselecteerd. *Oryza sativa* agglutinine, afgekort Orysata, is een induceerbaar proteïne dat behoort tot de familie van jacaline-gerelateerde lectines. Orysata bestaat enkel uit het jacaline domein met een moleculaire massa van 15.2 kDa. Dit lectine werd gedecteerd in rijstzaailingen na blootstelling aan NaCl behandeling. Naast Orysata, werden twee eiwitten onderzocht die behoren tot de EUL-gerelateerde familie. OrysaEULS2, met een totaal moleculair gewicht van 22.7 kDa, vertegenwoordigt het S-type binnen de EUL familie en bestaat uit één EUL domein dat voorafgegaan wordt door een 56 aminozuren (AZ) lange, ongerelateerde N-terminale sequentie. OrysaEULD1A vertegenwoordigt het D-type binnen de EUL familie en bestaat uit twee bijna identieke EUL domeinen (91% sequentiegelijkenis en 72% sequentieïdentiteit) die in tandem gepositioneerd zijn, worden gescheiden zijn door een 23 AZ lang verbindingsdomein en worden voorafgegaan door een 19 AZ lang N-terminale sequentie. De totale moleculaire massa van OrysaEULD1A bedraagt 38.5 kDa.

In het **eerste hoofdstuk** wordt een overzicht gegeven van de huidige literatuur in verband met lectinologie en de rol van lectines in de respons op de meest belangrijke abiotische en biotsche stressvormen.

Om de mogelijkheid te bestuderen om de geselecteerde eiwitten in te zetten om de stresstolerantie van voedselgewassen te verhogen, werden eerst de lectines gekarakteriseerd. **Hoofstuk 2** beschrijft de karakterisatie van Orysata, OrysaEULS2 en OrysaEULD1A. Aangezien de induceerbare proteïnen slechts op een heel laag niveau tot expressie komen en bovendien alleen na toediening van stress, is het moeilijk en in de meeste gevallen zelfs onmogelijk om voldoende hoeveelheden van deze eiwitten uit plantenweefsels te isoleren. Daarom werd het heterologe expressiesysteem *Pichia pastoris* gebruikt om alle bestudeerde eiwitten tot expressie te brengen. Vervolgens werden de recombinante eiwitten opgezuiverd door gebruik te maken van een combinatie tussen ion-uitwisselingschromatografie en metaal affiniteitschromatografie. In **hoofdstuk 2A** toonde glycaan array analyse aan dat Orysata interageert met mannose-rijke en meer complexe N-glycaanstructuren. In **hoofstuk 2B** werd een drie-dimensioneel model geconstrueerd voor het EUL domein van OrysaEULS2. Vervolgens toonde mutatie-analyse van enkele belangrijke aminozuren aan dat de tryptofaan residu's een belangrijke rol spelen in de vorming van de koolhydraatbindingsplaats. Uit de glycaan array analyse werd besloten dat OrysaEULS2 affiniteit vertoont tegenover mannose-rijke structuren en bijgevolg werd voor de eerste keer bewezen dat OrysaEULS2 een echt lectine is. In **hoofdstuk 2C**, werd aangetoond dat OrysaEULD1A lectine activiteit vertoont, aangezien het rode bloedcellen agglutineert. De koolhydraatbindingsactiviteit was gericht tegen galactose-gerelateerde verbindingen.

Om een beter inzicht te verwerven in de promoter sequenties van de EUL lectines, werd de promoteractiviteit van OrysaEULS2 en OrysaEULD1A onderzocht. In **hoofdstuk 3A**, werd het GUS gen gekloneerd onder de controle van elk van de EUL promoters en getransformeerd in *Arabidopsis* planten. Het GUS kleuringspatroon werd geanalyseerd na verschillende abiotische stress behandelingen en gedurende

verscheidene ontwikkelingsstadia. GUS activiteit werd gedetecteerd voor de OrysaEULD1A promoter in de jonge echte bladeren, maar dit was niet het geval voor de OrysaEULS2 promoter. Verder werd aangetoond dat de activiteit van de OrysaEULD1A promoter steeg na ABA en mannitol behandeling, maar daalde na een NaCl behandeling.

Om de rol van de EUL proteïnen verder te onderzoeken, werden qPCR experimenten uitgevoerd op rijstzaailingen die werden blootgesteld aan verscheidene vormen van abiotische en biotische stress (**Hoofdstuk 3B**). De qPCR analyse toonde aan dat de expressie van de hele set EUL genen uit rijst stress-gerelateerd is. Na ABA en NaCl behandelingen, werd de expressie van alle EULs alsook Orysata opgereguleerd in de wortels. Een meer variabele expressie werd vastgesteld in de scheuten na abiotische stress behandelingen. In de respons tegen *Xanthomonas oryzae* infectie, werden alleen EULs van het D type (behalve OrysaEULD1A) alsook Orysata opgereguleerd op 4 en 8 dagen na infectie, terwijl geen effect werd waargenomen op de expressie van de S-type EULs. Bovendien toonden onze resultaten aan dat *Magnaporthe oryzae* de expressie van sommige EULs beïnvloedde, in het bijzonder OrysaEULS3, OrysaEULD1A en OrysaEULD2 alsook Orysata.

In een poging om de mogelijkheid na te gaan dat EULs (en Orysata) de tolerantie van planten tegen abiotische en biotische vormen van stress kunnen verhogen, werden de onderzochte lectines tot over-expressie gebracht en werden de prestaties van deze transgene overexpressie-plantenlijnen onderzocht na blottstelling aan verschillende vormen van stress. **Hoofstuk 4A** beschrijft de overexpressie van Orysata, OrysaEULS2 en OrysaEULD1A in *Arabidopsis* planten en de selectie van homozygote lijnen met één insertie. Aangezien planten verschillende niveau's van tolerantie vertonen, gebaseerd op de intensiteit van de stress en het ontwikkelingstijdstip, werden de experimenten uitgevoerd gedurende verschillende ontwikkelingsstadia. In het bijzonder kiemings-, zaailing- en volwassen stadia werden onderworpern aan matige en zware stressniveau's. Onze resultaten tonen dat de onderzochte lectines tolerantie tegen droogte en zout kunnen bieden, vooral aan de volwassen planten. In **hoofdstuk 4B**, werd de coderende sequentie van Orysata constitutief tot expressie gebracht in tabaksplanten. Nadien werden vijf

heterozygote lijnen getest tegen bijtend-kauwende (*Spodoptera exigua*) en stekend-zuigende (*Myzus persicae* en *Acyrtosiphon pisum*) insecten. Transgene tabaksplanten die Orysata tot overexpressie brengen, induceerden sterfte en veroorzaakten vermindering in het larvale gewicht en groeivertraging bij *S. exigua*. Daarnaast vertoonde Orysata dodelijke effecten op nymfen van *M. persicae* en *A. pisum*. Onze bevindingen toonden aan dat Orysata een sterke insecticidale activiteit vertoont, wat een mogelijke rol van dit lectine als een beschermingsmiddel tegen zowel bijtend-kauwende als stekend-zuigende plaaginsecten suggereert.

In **hoofdstuk 5** worden de resultaten van het hele doctoraatsproefschrift besloten en worden perspectieven om het werk verder te zetten voorgesteld.Globaal genomen heeft dit doctoraatsproefschrift ons inzicht in de rol van de nucleocytoplasmatische lectinen in de respons tegen externe vormen van stress aanzienlijk vergroot. Onze resultaten wijzen erop dat Orysata, OrysaEULS2 en OrysaEULD1A tolerantie kunnen bieden tegen de meest belangrijke vormen van abiotische en biotische stress die gewassen normaal ondervinden. Niettegenstaande is er meer onderzoek nodig om de rol van deze lectines in de respons tegen stressbehandelingen te bevestigen en de mogelijkheid te onderzoeken om via toepassing van transgenese-gebaseerde strategiën deze lectines in te zetten in de bescherming van voedselgewassen tegen externe vormen van stress.

Addendum

Suppl. File 3A.1 Promoter sequences for OrysaEULS2, OrysaEULS3, and OrysaEULD1A. Different promoter elements are indicated. In red: start codon. In green: TATA box. In yellow: CAAT boxes. Underlined: 5' untranslated regions. In grey: transcription start site (TSS). The nucleotide A in the start codon (ATG) is designated as +1

OrysaEULS2 promoter

```
-1913 ATAGTGTTACGCTAAAGCATGCCCACATGCCCACAAACGAGCCATAGTGC
-1863 CCGTCCGAGAGAGAGAGAGAGAGAGAGCGGGGGGGATAGGGAGAGGAGAGG
-1813 AGAGGAGAGGTGGAGGTGTGAAGGGAGAAGATGAGTCCCACCTTTTTTTT
-1763 CCTTACCGGATGCAACTTAACCATAGTGGAAAAAACCGGACCGGTAGATG
-1713 ACCCAGTCACCCTTCTGGTTCACCGGTCTCGGACAGTTTGACCATGAACC
-1663 AGCCGATCGAACCGCAGTTTACTAAGTGCATGCATGTAACGTATGTATTT
-1613 CAATGCTAACACATACTCCATCCGTTTCAAAATGTTTGACACCGTTGACT
-1563 TTTTAGCACATGTTTGACCGTTCGTCTTATTCAAAAAAATTATGAAATAT
-1513 GTAAACTATATGTGTACATGAGAGTATATTTAACAATAAATCAAATGAT
-1463 ATGAAAGAATAAATAATTACTTAATTTTTTTAAATAAGACGAATGATCA
-1413 AACACGTACTAAAAAGTCAACGGTGTCAAACATTTTGAAACGGAGAGAGT
-1363 ACATCCGATTGGTAGGATGGCCACGGGAGTTGCCATGCATCCCATTGACC
-1313 AGGGATGAACCGGCCATCGTGTGGCCCATTCATTCATTGACTTGCCTTAG
-1263 TCCGGCTGGACCAAAGGCGGTTTATTGCCCGGTTCATCACAAAACCGGCC
-1213 GCCAACCTCTGGTTCCGATTTTTTTCCGGTTAAACTGTCAGTCCGTTCCA
-1163 GTTTTTTTTCCACTATGCAATTAACGCCATGTTGTAGAAAACCACCGTAC
-1113 AACTTGCCAATAAGGTAGCTTCCTTCTTCTTTCTTTTTTTTTTTGAGAA
-1063 CTAGGTTAGCTTCCTTCTAGAGCTCCTTGTACGACCTCGTTGGAAAACTC
-1013 ATTGTTGGACCTAATTCATCACTCTTTTTATTTCAAACTAAAGGGTTTCC
-963  TTTTTAGACTTTTCTAATTGTTCCAAACTAATAGAGGTGGCATCAATCAA
-913  ATTTGCTCTTTCTTGTCCTATCTAAGAGTATCTTGATGAAACTCTTATCT
-863  TTATCGACTTTCGTAATCAAAATTTTAGTTTTTTCAAGCTGCTCAATGGT
-813  CACTTTACGACATGATTAGTGTTATGCTCGGTATTATTGTAGTTGAGGGA
-763  GAATAAGTGTAACTTCCGTTCATGATCGGCCCGTTAAGAATTTAGACCCA
-713  TTAGTCCCACCCCCAACTCCCTTTTCTTTTCTTCCTTTTTGGCCAAATTT
-663  TCAAATTTCTGGTGGATACACTGCACGCCCGGGGAATTTTCTGTCGAAAA
-613  ATTGTATAGGTAATCCCTAATAAATTTCTAAATGTGCTTGGCTGATTCCT
-563  TTTCTTCTCGATCGATCCCGTTATTCCTTTTTAATCTTTTCTTGTGCGGA
-513  ATTGCAGCACTTTTTCTCTGCTTGATCACAACGCAGCAAGGACATGAGAA
-463  TACGTGAGATTCTCCGGGCGATGGAGCTGCACCGTGTGCTTTGCTATCGT
-413  CTCGCGTCGCGCTCGCCTGCAATAATACGGCAGCCAGATCGAGATGGTTC
-363  CCAAAACGACTAGAAACGGAATCACGTAAAGCACCATGGATGGATGGAAT
-313  CACCTTTTCCCCAGGAAAGCGCAAGTGTGGAATGGCATCATCTATCCTAC
-263  CAAGTTCCTAAGCCCCTGACCATGCAGTGCACGAATAATGCACATCGATC
-213  AGCTATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGATAGATAGCATCTC
-163  GATCAGTGCTACTGATCAATCGATCATCCCCTATAAATCCCCCATCCCAT
-113  TGCCATGCATAGAGGCAGCACATCAACGATCGAACGAACAACGCACTCTG
-63   CAGCTAGCTATACAATCTCTGTTTAATTAGCTAGGTAGCTAGCGAGTCAT
-13   AAATCAAGAAGCCATGGACT
```

OrysaEULS3 promoter

```
-1911 ATCCTCGAGGGGAAGGGCCCAGAGCTTCTAATCCTACACGGCAACAACAC
-1961 TTTATACTTGTGTATAATTTCTCTTCGTTTCTGAGTTCATGGCTATCTTT
-1811 GTCTCTCTTATCTTCTCCCTTTTGCTATCTCTATATTTGTGATTGCCATG
-1761 GAAATACAGAAGACGAGGGGCCGCTGATGTATTGACTTGTGAGTTCTAGT
-1711 TTGGGGGCTTTCGTTTGCAAGGATTTTCATTTGATTGATCTTCTTGATTT
-1661 GTTATTTTGGATCGAAGCTGGCTACCTTCATGCCGTGTTGCATACCATAC
-1611 TTCAGTCTGGTCGGTCAGACCGATGATCGAGACCGAAGTATCGATTAACC
-1561 GAACTTTCGGTTTGTATATCTGTGAGGACAAGTTCTGTCAT`CAAT`TCAAA
-1511 ATGTTAACATAATAAATTGTGTTAAAATTATTGCTAAAATTCTGTGAAAA
-1461 TTGTCCTAAATATTGTGTTCTGTCAACTTTTACAAAATTTGTCAACATTT
-1411 AACTTGAAATACGCACAAATTTAACTTGGAACATGTCCAGTGTTGATGCC
-1361 AAAATTATTGATTT`CAAT`AGTAGGGGTATATTTGTCCATTCACCACACAT
-1311 TTAACACCGTTACCATGAGTTAACAGAGTAGAGGCAGACGGCTGCTTCAC
-1261 TTAAAAGTAGGGGCAAAATCACAAACCCAAAAAACAGGAGCAAAATCAT
-1211 AATTAAAC`TATAAA`ATAATGGTAGGAACA`CAAT`TACCCTAGTTTCTACTA
-1161 GTATTACTCATGAAATTACATACAGAAATTCTTGTATATTTTATTTTAGT
-1111 TACATATATAAGACATATACGCGCATACTCCACCCA`G`ATACATCTCTCTA
-1061 ATAACACATGCTTAAAAGATACGAAAACCAGCAACTCACTATGGGAATGT
-1011 CTAGATCGACTTTAAACGTCTGAAATCACTTTTCTATTTCATTGTGTGAT
-961  AATGTAATATTTCTCACAACCATTTACTATTAAACAAAGTTCCTTCTTTT
-911  CAAAATGGCTCAA`CAAT`AAATTATTCTAGCTAACAACAAAAATCATAATA
-861  TATATATCATAAAAAATATCTAGTATCACTTATTTGCTTATTTGTTATTC
-811  ACTCGTTTGAATCAAATGTTAGAGTACTCTATGATATTAAAAGATATTGT
-761  GGGA`CAAT`TTGTACAAGCGGCAAAAG`CAAT`AGTGGGACGGTAGATTGGCT
-711  GTCATCGTCACTACATAATTTCAAAGGAGTACAGAAAGGGTAGCATAGAT
-661  ATTTAAAATCGAGGAAATACGAATGGCTGATTATTTTAGACGTGTTTAGG
-611  GGTGAATTTGGATTGATTTAGATGTTGATGGTGTGAATAACCCCAAGTTG
-561  CAGGGATGG`CAAT`GGGGCCCCGTTCCCCGTCACCCGGTGGGGAATTCCTC
-511  TACTAGGGTGCCAAAACAGGAGAGAAATAATCCCCATGGGGATGTTAACG
-461  GGCGAAATTCATACCCATCGGGTGAGCGGGGATGGGGACCCGGTAACTGC
-411  CTGCCTGTCATGTGGGTCCTATCTTTCTCACACACATATAGCCTGTGTGT
-361  ACGTCTATTTCTTCAAGCAAGGAGACTAGCTAGCTAGTTACTCCACATCT
-311  GTATCTGTACGCATGCAAAGTAAACTAGTCCTATTTATTTGTAAATCGTG
-261  TGTCTCGTTCTCGTCGATCTCCATCCTAGCACTGTGATAGCTCATGCTTG
-211  GATTGAGGGATTGATTGATTTCTTCTCTGAACTTTTTGTGATTGAGCCCA
-161  CTGTAACTGTAGCGATTACCGGGGGGTATTATGGGCCGTGGGTACCCGCGC
-111  GGGGCCGGGGACGGGACGGGAGACC`CAAT`CATAAATTGGAGGCGCCGGCC
-61   ATCTCTCCTCTAGACCTCCTGACTCTAATAAATCTCCTCCATTCCCCCAA
-11   ATCCATCCATC`ATG`GAGTTC
```

OrysaEULD1A promoter

```
-1896 CAATTTAGCAAATGCAACATCATCTGCTCCAACGTCCACAAACTCCACAG
-1846 GACAACCTACACATAGCAATTACTACTCCTAATTGTAAGGTGAGTTTAAG
-1796 GACAGGCTTCACCTACTCCTCTAAAGTTATCACTGGTTATTTATAAGCTC
-1746 TCTACAAAAGATCTATGTCAGTATATTTTCCTACTTAGAAGGTATTAAAT
-1696 GAAGAGAGAGAGTAAAGCTATCTACTAACTTAGAGATAATCTATAGAGAA
-1646 AAACGAGGCAATGCATGAGAGAGCTATAGATACAAATGTAGACATACTAT
-1596 TAAGGTGGTTTACTATTAATCGAGTCTATTGCTGAGATGTATATGTTTTA
-1546 TAGAGAGTACCTTATTTTACCATTGCGGGAGCTATAAGGGTAGAACGGTT
-1496 GAAACAACGGTACTATACTACTACAAACCAGCAACGCGTTGCTATCTGCA
-1446 CAGCACTCCAAATGTTCGCGAAAACTTGACTTGTTTTCGTTTGAGAAAGT
-1396 TAAAGAGTCGAAACTTTTCATCCTGACAAGTTGAGAGAGTAGTACGTGAC
-1346 AACAAAAATCAAAGGGAGATGCATGTAGCTAACCCGTTTTCTGCCTGCTG
-1296 CATTGTTGTTAAGCACAACAACAAAGCTTAAGCCTGTAAAAGCATTGCAA
-1246 AAGGCAGGCGACGCGTCGATTAATTAGGTGTTGAAATCTGGAACGTTCTT
-1196 GGCCATACAGGCGCTGCGTGGCCGTGATATAGAGCAGCTTGCCGGGTAAA
-1146 AAGATTATTCTCTTATCAGTTAACAACTTTTTTACCACCTTTCATGCATC
-1096 ATGCATCATTCAGGTCAGATTAGATGTGCTCATTGGAAACGGACAGGGAT
-1046 GATCGAACCATCCATCCATGGGCGATCCATGGTGATTCAGAAAAGTAAAA
-996  CTGCCTTTGTTTCTTGTGTAGCACAGGGATGCAAACGTACAAAACGACCG
-946  TCACAAGGGTCATCTCTAGCTGGAATCACATCTCTTAGCTTGTTTTATGG
-896  ATTTGCGTCAGGTGCACGCTTCAATCCTGGGCCACGGCATGCATTTCGAT
-846  CACCTCGTCCTTCGTGGCGACATACCTTTCGATTTTGGGGAAAACAATGG
-796  TCGAAATTCCCGGACAATGGAAAGCGAGCAAGCAGCATCTCTCTTTGTTC
-746  TTTTCTTAAGCAAAGATTGTTTTTTTTCCCCTTAATTTTTATTTCGGGTGA
-696  AAGAGTTTATGCTTTTCTGAAAAAGAAAGTGTTGTTTTACAAGTTGGAAG
-646  GGATCCACATTCTCTATCATAATGGCAGCACTACCATTACAAGCACTGAC
-596  GTGTGGAACAATAGGAGTATATGACCCACATACCTGTCACGGTAATGGTA
-546  GTGATATTATGTTATCCGGTTGTAAGAAATCAGGCTGCTCCAAAAATACT
-496  CAATGAAGTTGGTGTACTGCTCTAGTCTTCCACTGAGTAAGTAGAGCAAT
-446  TAACTGTGCACCAGTAATCAAATCTTTGCATTTACTCGCACATTTAGTAA
-396  TGCAAATCTCAGCACAATGAATTACTGGTACTAGCATCCAAGTCAAGTTC
-346  ATGGAACCGAATCTGAGAGTACATTAAACACCAAAGATATATCATATCAT
-296  ATCAAATATTCAAATCAAACGGCTAGAATGCAAAACATACACGCACGCAT
-246  GTCACGCGTTTGAGCCAGTCGAGGACCACCGGTGGGCCCAGCTGTCAGCG
-196  AACGTGGGAGTTGGCACAGGAGAGCGCGTCATCGAGCCGAGCCGATCGAT
-146  CAGTGGGTGAGGGTGAGAGTGGCCGGCCTTGGCCTTCTATAAAACGTGGC
-96   ACCTCTCGCTCCTCCTCGCATCATCGCTCGCAGTCGCAGCACAAAAACAT
-46   TTCTCTCGCGAGTCGCGAGTTGCACCGATCGAATTAAGGGTCGAAGATGT
```

Suppl. File 3A.2 Stress regulatory elements in the promoter region of OrysaEULS2, OrysaEULS3 and OrysaEULD1A. *In silico* promoter analyses were performed using the PLACE database. Frequency refers to the number of times each core sequence is present in the promoter region. S2: OrysaEULS2; S3: OrysaEULS3 and D1A: OrysaEULD1A. Promoter elements related to certain stresses are indicated in different colors.

Light	Drought	Cold	ABA	Gibberellin	Salt
Pathogen	freezing	Anaerobic	Auxin	brassinosteroid	

Stress regulatory elements							
Function	**Element**	**Element core sequence**	**Frequency**			*	+
			S2	S3	D1A		
Response to light	MYCCONSENSUSAT	CANNTG	1	3	5	6	Both
	GATABOX	GATA	3	6	4	1	Both
	GT1CONSENSUS	GRWAAW	12	0	6	1	Both
	SORLIP1AT	GCCAC	2	0	4	1	Both
	IBOXCORE	GATAA	0	0	1	1	Both
Response to drought	MYCCONSENSUSAT	CANNTG	1	3	5	6	Both
	DRE2COREZMRAB17	ACCGAC	0	1	0	4	Mono
	DRECRTCOREAT	RCCGAC	0	1	0	3	Both
Response to cold	MYCCONSENSUSAT	CANNTG	1	3	5	6	Both
	DRE2COREZMRAB17	ACCGAC	0	1	0	4	Mono
	DRECRTCOREAT	RCCGAC	0	1	0	3	Both
Response to ABA	MYCCONSENSUSAT	CANNTG	1	3	5	6	Both
	WRKY71OS	TGAC	7	3	8	2	Both
	RYREPEATVFLEB4	CATGCATG	2	0	0	1	Both
	ARFAT	TGTCTC	0	2	0	3	Both
	DRE2COREZMRAB17	ACCGAC	0	1	0	4	Mono
	AGCBOXNPGLB	AGCCGCC	0	0	0	2	Both

Response to gibberellin	WRKY71OS	TGAC	7	3	8	2	Both
	GARE2OSREP1	TAACGTA	1	0	0	1	Mono
	CAREOSREP1	CAACTC	1	1	2	1	Mono
	MYBGAHV	TAACAAA	0	2	0	1	Both
	GARE1OSREP1	TAACAGA	0	1	0	1	Mono
Response to salt	DRE2COREZMRAB17	ACCGAC	0	1	0	4	Mono
	DRECRTCOREAT	RCCGAC	0	1	0	3	Both
	AGCBOXNPGLB	AGCCGCC	0	0	0	2	Both
	MYCCONSENSUSAT	CANNTG	1	3	5	6	Both
Response to pathogen	BIHD1OS	TGTCA	1	3	4	1	Mono
Response to freezing	MYCCONSENSUSAT	CANNTG	1	3	5	6	Both
Response to anaerobic conditions	ANAERO3CONSENSUS	TCATCAC	2	0	0	1	Both
	ANAERO2CONSENSUS	AGCAGC	1	1	3	1	Both
	ANAERO1CONSENSUS	AAACAAA	0	1	1	1	Both
Response to auxin	ARFAT	TGTCTC	0	2	0	3	Both
Brassinosteroid	ARFAT	TGTCTC	0	2	0	3	Both

*: Times repeated in different stresses. +: Element was found only in sequences from monocotyledonous (Mono) plants or in sequences of both dicotyledonous and monocotyledonous (Both) plants
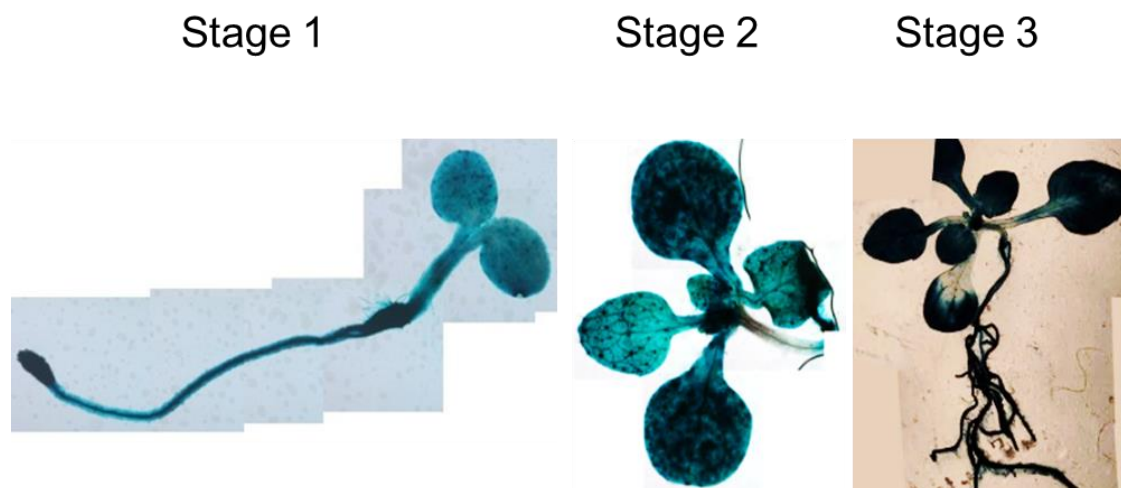
Suppl. Fig 3A.1 Histochemical analysis of GUS activity in *Arabidopsis* seedlings harboring a 35S promoter::GUS construct. Images were taken using a Nikon eclipse TE2000-e epi-fluorescence Microscope (Nikon Benelux, Brussels, Belgium) and a Leica microscope (Leica, Nussloch, Germany)
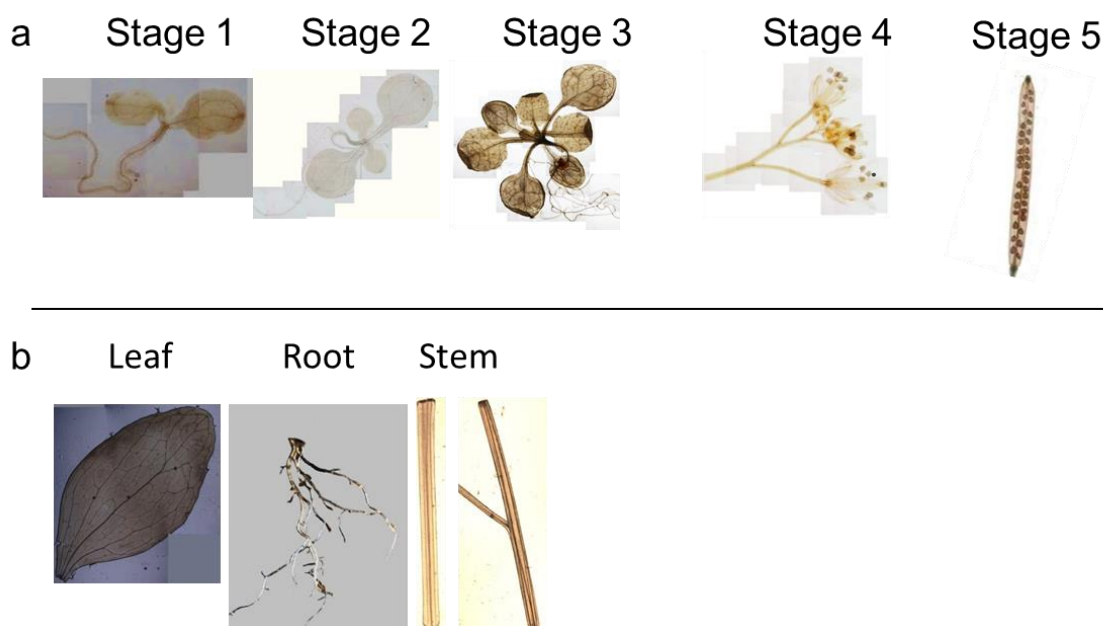
Suppl. Fig. 3A.2 Histochemical analysis of GUS activity in *Arabidopsis* seedlings. No GUS staining was detected for *Arabidopsis* plants harboring the OrysaEULS2 promoter. Non-treated and stress-treated plant tissues did not show any GUS activity. Some representative pictures are shown for each plant stage (a). GUS staining was absent in the leaf, stem and root samples analyzed for developmental stages 4 and 5 of *Arabidopsis* plants expressing each of the EUL promoter constructs. Some representative pictures are shown for *Arabidopsis* plants harboring the OrysaEULD1A promoter at developmental stage 5 (b). Images were taken using a Nikon eclipse TE2000-e epi-fluorescence Microscope (Nikon Benelux, Brussels, Belgium) and a Leica microscope (Leica, Nussloch, Germany)

Suppl. Table 3B.1 Table showing all PCR primers used in this study

Primers	Name and sequence (5' to 3')	Name and sequence (5' to 3')
Gene of interest	Forward	Reverse
OrysaEULS2	EVD898/AGCAGTACGGCGGGTATG	EVD899/GTTCCCCTCGATGCTCAC
OrysaEULS3	EVD896/CTTCCCACGACTACGCTCA	EVD897/GTACTCGGATCCACCGTAGTG
OrysaEULD1A	EVD902/GAGGCACAGCAACAAGATCA	EVD903/CTCTGGGTTGTAGGGAACCA
OrysaEULD1B	EVD1089/CCGTGATCTGTGGAGTTGG	EVD1090/GCAGGACTCGAGAAAACGAC
OrysaEULD2	EVD1091/TCGAGAGACCGTCAACAAAA	EVD1092/GGACACGCAACAGTAACACG
Reference genes	Forward	Reverse
EXP (LOC_Os03g27010)	EVD910/TGTGAGCAGCTTCTCGTTTG	EVD911/TGTTGTTGCCTGTGAGATCG
EIF5C (LOC_Os11g21990,1	EVD912/CACGTTACGGTGACACCTTTT	EVD913/GACGCTCTCCTTCTTCCTCAG
UBQ5 (LOC_Os01g22490.1)	CTTAGGCGTAGGCTCCTGTTC	AGAGGTGATGCTAAGGTGTTCA
EEF1a (LOC_Os03g08020.1)	GGCTGTTGGCGTCATCAAGA	CCGTGCACAAAACTACCACTT
Positive control gene	Forward	Reverse
Orysata	EVD966/CGAAATAATGTTCCATGGTGTT	EVD967/TGTACTACGGATCGGTGCAA
PR1A (LOC_Os07g03710.1)	EVD1109/CGTACGTATGCTGGTGAGAA	EVD1110/CTAAGCAAATACGGCTGACAGT
cDNA quality	Forward	Reverse
Ubiquitine5 (AK061988)	EVD635/ACCACTTCGACCGCCACTACT	EVD636/ACGCCTAAGCCTGCTGGTT

Suppl. Table 3B.2 The fold change for genes under study after all treatments. Plus (+) and minus (-) signs represent up-regulation and down-regulation, respectively. Significant differences are indicated in bold.

Assay	Abiotic					
Tissue	Shoot			Root		
Treatment	ABA	NaCl	Mannitol	ABA	NaCl	Mannitol
OrysaEULS2	**+1.73**	-1.25	+1.40	**+2.76**	**+4.09**	-1.31
OrysaEULS3	-1.10	-1.21	-1.16	**+2.43**	**+1.21**	-1.04
OrysaEULD1A	**+1.23**	**+1.27**	+1.04	**+1.92**	**+1.40**	+1.12
OrysaEULD1B	+1.25	**+2.96**	-1.02	**+4.17**	**+2.74**	-1.05
OrysaEULD2	**+1.92**	**+2.59**	+1.00	**+2.22**	**+1.90**	+1.04
Orysata	**+2.16**	**+1.90**	**+2.50**	**+20.97**	**+4.03**	+1.18
Assay	Biotic					
Treatment	*Xanthomonas*			*Magnaporthe*		
Time point	2dpi	4 dpi	8 dpi	1 dpi	2 dpi	
OrysaEULS2	-1.04	-1.30	-1.41	-1.76	-1.21	
OrysaEULS3	-1.07	-1.10	+1.15	**-4.46**	-1.73	
OrysaEULD1A	+1.05	-1.06	**-2.71**	**-4.20**	-1.10	
OrysaEULD1B	+1.06	**+2.14**	**+9.35**	-1.27	+1.01	
OrysaEULD2	+1.00	**+2.57**	**+5.01**	**-2.41**	**+1.87**	
Orysata	-1.06	**+3.53**	**+4.62**	**-5.11**	**-5.45**	
PR1A	+3.30	+1.39	+1.63	**+5.88**	**+53.91**	

References

Ahn IP, Kim S, Kang S, Suh SC, Lee YH (2005) Rice defense mechanisms against *Cochliobolus miyabeanus* and *Magnaporthe grisea* are distinct. Phytopathol 95: 1248-1255

Al Atalah B, Fouquaert E, Van Damme EJM (2013) promoter analysis for three types of EUL-related rice lectins in transgenic *Arabidopsis*. Plant Mol Biol Rep 31: 1315–1324

Al Atalah B, Fouquaert E, Vanderschaeghe D, Proost P, Balzarini J, Smith DF, Rougé P, Lasanajak Y, Callewaert N, Van Damme EJM (2011) Expression analysis of the nucleocytoplasmic lectin 'Orysata' from rice in *Pichia pastoris*. FEBS J 278: 2064–2079

Al Atalah B, Rougé P, Smith DF, Proost P, Lasanajak Y, Van Damme EJM (2012) Expression analysis of a type S2 EUL-related lectin from rice in *Pichia pastoris*. Glycoconj J 29: 467–479

Allen RD (1995) dissection of oxidative stress tolerance using transgenic plants. Pant Physiol 107: 1049-1054

Amano K, Takase M, Ando A, Nagata Y (2003) Production of functional lectin in *Pichia pastoris* directed by cloned cDNA from *Aleuria aurantia*. Biosci Biotechnol Biochem 67: 2277-2279

Aoki K, Perlman M, Lim JM, Cantu R, Wells L, Tiemeyer M (2007) Dynamic developmental elaboration of N-linked glycan complexity in the *Drosophila melanogaster* embryo. J Biol Chem 282: 9127–9142

Arndt JW, Gu J, Jaroszewski L, Schwarzenbacher R, Hanson MA, Lebeda FJ, Stevens RC (2005) The structure of the neurotoxin-associated protein HA33/A from *Clostridium botulinum* suggests a reoccurring β-trefoil fold in the progenitor toxin complex. J Mol Biol 346: 1083-1093

Asakura T, Hirose S, Asatsuma S, Nanjo Y, Nakaizumi T, Itoh K, Hori H, Komatsu S, Mitsui T (2007) Proteomic characterization of tissue expansion of rice scutellum stimulated by abscisic acid. Biosci Biotechnol Biochem 71: 1260-1268

Asselbergh B, De Vleesschauwer D, Höfte M (2008) Global switches and fine-tuning-ABA modulates plant pathogen defense. Mol Plant-Microbe Interact 21: 709–719

Babu PR, Sekhar AC, Ithal N, Markandeya NG, Reddy AR (2002) Annotation and BAC/PAC localization of non-redundant ESTs from drought-stressed seedlings of an indica rice. J Genet 81: 25–44

Bailey TA, Zhou XJ, Chen JP, Yang Y (2009) Role of ethylene, abscisic acid and MAP kinase pathways in rice blast resistance. Advances in Genetics, Genomics and Control of Rice Blast Disease: Springer Netherlands pp 185-190

Baker B, Zambryski P, Staskawicz B, Dinesh-Kumar SP (1997) Signaling in plant-microbe interactions. Sci 276: 726-733

Balzarini J (2006) Inhibition of HIV entry by carbohydrate-binding proteins. Antiviral Res 71: 237-247

Barondes SH (1988) Bifunctional properties of lectins: lectins redefined. Trends Biochem Sci 13: 480–482

Baumgartner P, Harper K, Raemaekers RJM, Durieux A, Gatehouse AMR, Davies HV, Taylor MA (2003) Large-scale production and purification of recombinant *Galanthus nivalis* agglutinin (GNA) expressed in the methylotrophic yeast *Pichia pastoris*. Biotechnol Lett 25: 1281–1285

Baumgartner P, Harper K, Raemaekers RJM, Durieux A, Gatehouse AMR, Davies HV, Taylor MA (2002) Large-scale production and purification and characterization of recombinant *Phaseolus vulgaris* phytohemagglutinin E form expressed in the methylotrophic yeast *Pichia pastoris*. Protein Expr Purif 26: 394–405

Bause E (1983) Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. Biochem J 209: 331–336

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28: 235-242

Bertaux C, Daelemans D, Meertens L, Cormier EG, Reinus JF, Peumans WJ, Van Damme EJM, Igarashi Y, Oki T, Schols D, Dragic T, et al. (2007) Entry of hepatitis C virus and human immunodeficiency virus is selectively inhibited by carbohydrate-binding agents but not by polyanions. Virol 366: 40-50

Bezerra WM, da Silveira Carvalho CP, Moreira RA, Grangeiro TB (2006) Establishment of a heterologous system for the expression of *Canavalia brasiliensis* lectin: a model for the study of protein splicing. Genet Mol Res 5: 216-223

Bienert GP, Schjoerring JK, Jahn TP (2006) Membrane transport of hydrogen peroxide. Biochim Biophys Acta 1758: 994–1003

Binzel ML, Hess FD, Bressan RA, Hasegawa PM (1988) Intracellular compartmentation of ions in salt adapted tobacco cells. Plant Physiol 86: 607–614

Biodiversity Institute of Ontario (2006) All Leps Barcode of Life Project. http://www.lepbarcoding.org/ [accessed 22 July 2009]

Blackman PG, Davies WJ (1985) Root-to-shoot communication in maize plants of the effects of soil drying. J Exp Bot 36: 39–48

Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, et al. (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. Proc Nat Acad Sci USA 101: 17033-17038

Blum A (2005) Drought resistance, water-use efficiency, and yield potential - are they compatible, dissonant, or mutually exclusive? Australian J Agric Research 56: 1159-1168

Blumwald E (2000) Sodium transport and salt tolerance in plants. Curr Opin Cell Biol 12: 431-434

Blumwald E, Grover A (2006) Salt tolerance. In: Halford NG (Ed.) Plant biotechnology: current and future uses of genetically modified crops. John Wiley and Sons Ltd. UK pp 206–224

Bohnert HJ, Nelson DE, Jensen RG (1995) Adaptations to environmental stresses. Plant Cell 7: 1099–1111

Boonjung H and Fukai S (1996) Effects of soil water deficit at different growth stages on rice growth and yield under upland conditions. 2. Phenology, biomass production and yield. Field Crop Res 48: 47-55

Bota J, Flexas J, Medrano H (2004) Is photosynthesis limited by decreased Rubisco activity and RuBP content under progressive water stress? New Phytol 162: 671–681

Botos I, Wlodawer A (2005) Proteins that bind high-mannose sugars of the HIV envelope. Prog Biophys Mol Biol 88: 233–282

Bouman BAM, Peng S, Castaòeda AR, Visperas RM (2005) Yield and water use of irrigated tropical aerobic rice systems. Agric Water Management 74: 87-105

Bourne Y, Houlès-Astoul C, Zamboni V, Peumans WJ, Menu-Bouaouiche L, Van Damme EJM, Barre A, Rougé P (2002) Structural basis for the unusual carbohydrate-binding specificity of jacalin towards galactose and mannose. Biochem J 364: 173-180

Bourne Y, Roig-Zamboni V, Barre A, Peumans WJ, Houlès-Astoul C, Van Damme EJM, Rougé P (2004) The crystal structure of the *Calystegia sepium* lectin reveals a novel quaternary arrangement of lectin subunits with a β-prism fold. J Biol Chem 279: 527-533

Boyd WC, Reguera RM (1949) Studies on haemagglutinins present in seeds of some representatives of the family Leguminoseae. J Immunol 62: 333–339

Boyd WC, Shapleigh E (1954) Specific precipitating activity of plant agglutinins (lectins). Sci 119: 419

Boyer JS (1982) Plant productivity and environment. Sci 218: 443–448

Boyer JS, Westgate ME (2004) Grain yields with limited water. J Exp Bot 55: 2385–2394

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

Bray EA, Bailey-Serres J, Weretilnyk E (2000) Responses to abiotic stresses. In: Gruissem W, Buchannan B, Jones R (Ed.) Biochemistry and molecular biology of plants. American Society of Plant Physiologists; Rockville pp 1158–1249

Bunn-Moreno MM, Campos-Neto A (1981) Lectin(s) extracted from seeds of *Artocarpus integrifolia* (jackfruit): potent and selective stimulator(s) of distinct human T and B cell function. J Immunol 127: 427–429

Cabuslay GS, Ito O, Alejar AA (2002) Physiological evaluation of responses of rice (*Oryza sativa* L.) to water deficit. Plant Sci 163: 815–827

Cao F, Yoshioka K, Desveaux D (2011) The roles of ABA in plant-pathogen interactions. J Plant Res 124: 489–499

Carpentier SC, Witters E, Laukens K, Van Onckelen H, Swennen R, Panis B (2007) Banana (*Musa* spp.) as a model to study the meristem proteome: acclimation to osmotic stress. Proteomics 7: 92–105

Cereghino JL, Cregg JM (2000) Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol Rev 24: 45–66

Carraway KL, Hull SR (1989) O-Glycosylation pathway for mucin-type glyco-proteins. Bioessays 10: 117-121

Chakraborti D, Sarkar A, Mondal HA, Das S (2009) Tissue specific expression of potent insecticidal, *Allium sativum* leaf agglutinin (ASAL) in important pulse crop, chickpea (*Cicer arietinum* L.) to resist the phloem feeding *Aphis craccivora*. Transgenic Res 18: 529–544

Chang T, Chen L, Chen S, Cai H, Liu X, Xiao G, Zhu Z (2003) Transformation of tobacco with genes encoding *Helianthus tuberosus* agglutinin (HTA) confers resistance to peach-potato aphid (*Myzus persicae*). Transgenic Res 12: 607–614

Chao DY, Luo YH, Shi M, Luo D, Hong HX (2005) Salt-responsive genes in rice revealed by cDNA microarray analysis. Cell Res 15: 796-810

Chaves MM, Flexas J, Pinheiro C (2009) Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. Ann Bot 103: 551–560

Chaves MM, Oliveira MM (2004) Mechanisms underlying plant resilience to water deficits: prospects for water-saving agriculture. J Exp Bot 55: 2365-2384

Chen X, Shang J, Chen D, Lei C, Zou Y, Zhai W, Liu G, Xu J, Ling Z, Cao G, et al. (2006) A B-lectin receptor kinase gene conferring rice blast resistance. Plant J 46: 794–8

Chen Y, Peumans WJ, Hause B, Bras J, Kumar M, Proost P, Barre A, Rougé P, Van Damme EJM (2002) Jasmonic acid methyl ester induces the synthesis of a cytoplasmic/nuclear chitooligosaccharide binding lectin in tobacco leaves. FASEB J 16: 905–907

Cheng Y, Qi Y, Zhu Q, Chen X, Wang N, Zhao X, Chen H, Cui X, Xu L, Zhang W (2009) New changes of plasma-membrane-associated proteome of rice roots under salt stress. Proteomics 9: 3100-3114

Chinnusamy V, Jagendorf A, and Zhu J K (2005) Understanding and improving salt tolerance in plants. Crop Sci 45: 437–448

Chisholm ST, Parra MA, Anderberg RJ, Carrington JC (2001) *Arabidopsis* RTM1 and RTM2 genes function in phloem to restrict long-distance movement of tobacco etch virus. Plant Physiol 127: 1667-1675

Claes B, Dekeyser R, Villarroel R, Van den Bulcke M, Bauw G, Montagu MV, Caplan A (1990) Characterization of a rice gene showing organic-specific expression in response to salt stress and drought. Plant Cell 2: 19–27

Clough DJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium* mediated transformation of *Arabidopsis thaliana*. Plant J 16: 735-743

Cristofoletti PT, de Sousa FA, Rahbé Y, Terra WR (2006) Characterization of a membrane-bound aminopeptidase purified from *Acyrthosiphon pisum* midgut cells. A major binding site for toxic mannose lectins. FEBS J 273: 5574–5588

Cushman JC, Bohnert HJ (2000) Genomic approaches to plant stress tolerance. Curr Opin Plant Biol 3: 117–124

Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR (2010) Abscisic acid: Emergence of a core signaling network. Annu Rev Plant Biol 61: 651–679

Dai LY, Liu XL, Xiao YH, Wang GL (2007) Recent advances in cloning and characterization of disease resistance genes in rice. J Integr Plant Biol 49: 112–119

Daly R, Hearn MTW (2005) Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. J Mol Recognit 18: 119–138

Dam TK (2011) Fine specificities of two lectins from *Cymbosema roseum* seeds: a lectin specific for high-mannose oligosaccharides and a lectin specific for blood group H type II trisaccharide. Glycobiol 21: 925–933

Davidowitz G, D'amico LJ, Nijhout HF (2003) Critical weight in the development of insect body size. Evolution and Development 5: 188-197

Davies WJ, Wilkinson Loveys SB (2002) Stomatal control by chemical signalling and the exploitation of this mechanism to increase water use efficiency in agriculture. The New Phytol 153: 449-460

De Souza Filho GA, Ferreira BS, Dias JMR, Queiroz KS, Branco AT, Bressan-Smith RE, Oliveira JG, Garcia AB (2003) Accumulation of SALT protein in rice plants as a response to environmental stresses. Plant Sci 164: 623–628

De Vleesschauwer D, Gheysen G, Höfte M (2013) Hormone defense networking in rice: tales from a different world. Trends Plant Sci 18: 555-565

De Vleesschauwer D, Yang Y, Cruz CV, Höfte M (2010) Abscisic acid-induced resistance against the brown spot pathogen *Cochliobolus miyabeanus* in rice involves MAP kinase-mediated repression of ethylene signaling. Plant Physiol 152: 2036–2052

Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ Thon M, Kulkarni R, Xu JR, Pan H, et al. (2005) The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature 434: 980–986

Delporte A, Lannoo N, Vandenborre G, Ongenaert M, Van Damme EJM (2011) Jasmonate response of the *Nicotiana tabacum* agglutinin promoter in *Arabidopsis thaliana*. Plant Physiol Biochem 49: 843-851

Demiral T, Turkan I (2006) Exogenous glycinebetaine affects growth and proline accumulation and retards senescence in two rice cultivars under NaCl stress. Environ Exp Bot 56: 72–79

Dhlamini Z, Spillane C, Moss JP, Ruane J, Urquia N, Sonnino A (2005) Status of research and application of crop biotechnologies in developing countries. Rome, Italy: food and agriculture oganization of the United Nations natural resources management and environment department.

Dooki AD, Mayer-Posner FJ, Askari H, Zaiee AA, Salekdeh GH (2006) Proteomic responses of rice young panicles to salinity. Proteomics 6: 6498-6507

Down RE, Fitches EC, Wiles DP, Corti P, Bell HA, Gatehouse GA, Edwards JP (2006) Insecticidal spider venom toxin fused to snowdrop lectin is toxic to the peach-potato aphid, *Myzus persicae* (Hemiptera: Aphididae) and the rice brown planthopper, *Nilaparvata lugens* (Hemiptera: Delphacidae). Pest Manag Sci 62: 77-85

Du J, Foissac X, Carss A, Gatehouse AMR, Gatehouse JA (2000) Ferritin acts as the most abundant binding protein for snowdrop lectin in the midgut of the rice brown planthopper (*Nilaparvata lugens*). Insect Biochem Mol Biol 30: 297–305

Elfstrand M (1898) Über blutkörperchenagglutinierende Eiweisse. In: Görberdorfer Veröffentlichungen a. Band I pp 1–159. Kobert R (Ed.) Enke, Stuttgart, Germany.

Elias SB, Brust PF, Koutz PJ, Waters AF, Harpold MM, Gingeras TR (1985) Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast *Pichia pastoris*. Mol Cell Biol 5: 1111-1121

Erbs G, Silipo A, Aslam S, De Castro C, Liparoti V, Flagiello A, Pucci P, Lanzetta R, Parrilli M, Molinaro A, et al. (2008) Peptidoglycan and muropeptides from pathogens *Agrobacterium* and *Xanthomonas* elicit innate immunity: Structure and activity. Chem Biol 15: 438–448

Fan LM, Zhao Z, Assmann SM (2004) Guard cells: A dynamic signaling model. Curr Opin Plant Biol 7: 537–546

Farooq M, Wahid A, Basra SMA, Din ID (2009) Improving water relations and gas exchange with brassinosteroids in rice under drought stress. J Agron Crop Sci 195: 262–269

Febvay G, Delobel B, Rahbé Y (1988) Influence of the amino acid balance on the improvement of an artificial diet for a biotype of *Acyrthosiphon pisum* (Homoptera: Aphididae). Can J Zool 66: 2449-2453

Fettke J, Hejazi M, Smirnova J, Höchel E, Stage M, Steup M (2009) Eukaryotic starch degradation: integration of plastidial and cytosolic pathways. J Exp Bot 60: 2907-2922

Fettke J, Nunes-Nesi A, Fernie AR, Steup M (2011) Identification of a novel heteroglycan-interacting protein, HIP 1.3, from *Arabidopsis thaliana*. J Plant Physiol 168: 1415-1425

Fitches E, Edwards MG, Mee C, Grishin E, Gatehouse AMR, Edwards JP, Gatehouse JA (2004) Fusion proteins containing insect-specific toxins as pest control agents: snowdrop lectin delivers fused insecticidal spider venom toxin to insect haemolymph following oral ingestion. J Insect Physiol 50: 61-71

Fitches E, Gatehouse AMR, Gatehouse JA (1997) Effects of snowdrop lectin (GNA) delivered via artificial diet and transgenic plants on the development of tomato moth (*Lacanobia oleracea*) larvae in laboratory and glasshouse trials. J Insect Physiol 43:727-739

Flexas J, Bota J, Loreto F, Cornic G, Sharkey TD (2004) Diffusive and metabolic limitations to photosynthesis under drought and salinity in C3 plants. Plant Biol 6: 269-279

Flowers TJ, Colmer TD (2008) Salinity tolerance in halophytes. Tansley Review New Phytol 179: 945-963

Flowers TJ, Duque E, Hajibagheri MA, McGonigle TP, Yeo AR (1985) The effect of salinity on leaf ultrastructure and net photosynthesis of 2 varieties of rice- further evidence for a cellular-component of salt-resistance. New Phytol 100: 37-43

Fouquaert E, Hanton SL, Brandizzi F, Peumans W, Van Damme EJM (2007) Localization and topogenesis studies of cytoplasmic and vacuolar homologs of the *Galanthus nivalis* agglutinin. Plant Cell Physiol 48: 1010-1021

Fouquaert E, Peumans WJ, Smith DF, Proost P, Savvides S, Van Damme EJM (2008) The old *Euonymus europaeus* agglutinin represents a novel family of ubiquitous plant proteins. Plant Physiol 147: 1316-1324

Fouquaert E, Peumans WJ, Vandekerckhove TTM, Ongenaert M, Van Damme EJM (2009a) Proteins with an *Euonymys* lectin-like domain are ubiquitious in Embryophyta. BMC Plant Biol 9:136

Fouquaert E, Smith DF, Peumans WJ, Proost P, Balzarini J, Savvides SN, Van Damme EJM (2009b) Related lectins from snowdrop and maize differ in their carbohydrate-binding specificity. Biochem Biophys Res Commun 380: 260-265

Fouquaert E, Van Damme EJM (2012) Promiscuity of the *Euonymus* carbohydrate-binding domain. Biomolecules 2: 415–434

Fujimoto Z, Kuno A, Kaneko S, Kobayashi H, Kusakabe I, Mizuno H (2002) Crystal structures of the sugar complexes of *Streptomyces olivaceoviridis* E-86 xylanase: sugar binding structure of the family 13 carbohydrate binding module. J Mol Biol 316: 65-78

Fukai S and Cooper M (1995) Development of drought-resistant cultivars using physiomorphological traits in rice. Field Crop Res 40: 67-86

Fukuda A, Nakamura A, Tagiri A, Tanaka H, Miyao A, Hirochika H, Tanaka Y (2004) Function, intracellular localization and the importance in salt tolerance of a vacuolar Na^+/H^+ antiporter from rice. Plant Cell Physiol 45: 146-159

Funakoshi Y, Suzuki T (2009) Glycobiology in the cytosol: the bitter side of a sweet world. Biochim Biophys Acta 1790: 81–94

Gaboriaud C, Bissery V, Benchetrit T, Mornon JP (1987) Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. FEBS Lett 224: 149-155

Gao JP, Chao DY, Lin HX (2007) Understanding abiotic stress tolerance mechanisms: Recent studies on stress response in rice. J Integ Plant Biol 49: 742-750

Garcia AB, Engler JA, Claes B, Villarroel R, Van Montagu M, Gerats T, Caplan A (1998) The expression of the salt-responsive gene salT from rice is regulated by hormonal and developmental cues. Planta 207: 172-180

Garcia AB, Engler JD, Iyer S, Gerats T, VanMontagu M, Caplan AB (1997) Effects of osmoprotectants upon NaCl stress in rice. Plant Physiol 115: 159-169

Gatehouse AMR, Davison GM, Stewart JN, Gatehouse LN, Kumar A, Geoghegan IE, Birch ANE, Gatehouse JA (1999) Concanavalin A inhibits development of tomato moth (*Lacanobia oleracea*) and peach-potato aphid (*Myzus persicae*) when expressed in transgenic potato plants. Mol Breed 5: 153-165

Gepstein S, Grover A, Blumwald E (2006) Producing biopharmaceuticals in the desert: building an abiotic stress tolerance in plants for salt, heat and drought. In: Knablein J, Muller RH (Ed.) Modern biopharmaceuticals: Wiley-VCH Verlag GmbH. Weinhaum pp 967–994

Geshi N, Brandt A (1998) Two jasmonate-inducible myrosynase-binding proteins from *Brasica napus* L. seedlings with homology to jacalin. Planta 204: 295-304

Gilson MK, Honig BH (1987) Calculation of electrostatic potential in an enzyme active site. Nat 330: 84-86

Gingeras TR, Higuchi R, Kricka LJ, Lo YM, Wittwer CT (2005) Fifty years of molecular (DNA/RNA) diagnostics. Clin Chem 51: 661-671

Giri J, Vij S, Dansana PK ,Tyagi AK (2011) Rice A20∕AN1 zinc-finger containing stress-associated proteins (SAP1∕11) and a receptor-like cytoplasmic kinase (OsRLCK253) interact via A20 zinc-finger and confer abiotic stress tolerance in transgenic *Arabidopsis* plants. New Phytol 191: 721-732

Godfray HC, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM, Toulmin C (2010) Food security: the challenge of feeding 9 billion people. Sci 327: 812-818

Goldstein IJ, Hughes RC, Monsigny M, Osawa T, Sharon N (1980) What should be called a lectin? Nat 285: 66

Gomord V, Chamberlain P, Jefferis R, Faye L (2005) Biopharmaceutical production in plants: problems, solutions and opportunities. Trends Biotechnol 23: 559-565

Grassi G, Magnani F (2005) Stomatal, mesophyll conductance and biochemical limitations to photosynthesis as affected by drought and leaf ontogeny in ash and oak trees. Plant Cell Environ 28: 834-849

Grattan R, Zeng L, Shannon M, Roberts S (2002) Rice is more sensitive to salinity than previously thought. Calif Agric 56: 189-195

Hakim RS, Blackburn M, Corti P, Gelman D, Goodman C, Elsen K, Loeb M, Lynn D, Smagghe G (2006) Growth and mitogenic effects of arylphorin *in vivo* and *in vitro*. Arch Insect Biochem Physiol 64: 63-73

Hakim RS, Baldwin K, Smagghe G (2010) Regulation of midgut growth, development, and metamorphosis. Ann Rev Entomol 55: 593–608

Haliwell B (1986) Oxygen free radicals and iron in relation to biology and medicine: some problems and concepts. Arch Biochem Biophys 246: 501–514

Hao Q, Van Damme EJM, Hause B, Barre A, Chen Y, Rougé P, Peumans WJ (2001) Iris bulbs express Type 1 and Type 2 ribosome inactivating proteins with unusual properties. Plant Physiol 125: 866–876

Harrison SJ, Mott EK, Parsley K, Aspinall S, Gray JC, Cottage A (2006) A rapid and robust method of identifying transformed *Arabidopsis thaliana* seedlings following floral dip transformation. Plant Methods 2: 19

Hartner FS, Glieder A (2006) Regulation of methanol utilisation pathway genes in yeasts. Microb Cell Fact 5: 39

Hauser F, Waadt R, Schroeder, Julian I (2011) Evolution of abscisic acid synthesis and signaling mechanisms. Curr Biol 21: R346–R355

Hazen SP, Pathan MS, Sanchez A, Baxter I, Dunn M, Estes B, Chang H-S, Zhu T, Kreps JA, Nguyen HT (2005) Expression profiling of rice segregating for drought tolerance QTLs using a rice genome array. Funct Integr Genomics 5: 104-116

Helenius A, Aeb M (2004) Roles of N-linked glycans in the endoplasmic reticulum. Annu Rev Biochem 73: 1019-1049

Hetherington AM, Woodward FI (2003) The role of stomata in sensing and driving environmental change. Nature 424: 901–908

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database. Nucleic Acids Res 27: 297-300

Hilder VA, Powell KS, Gatehouse AMR, Gatehouse JA, Gatehouse LN, Shi Y, Hamilton WDO, Merryweather A, Newell CA, Timans JC, et al. (1995) Expression of snowdrop lectin in transgenic tobacco results in added protection against aphids. Transgenic Res 4: 18–25

Hirano K, Teraoka T, Yamanaka H, Harashima A, Kunisaki A, Takahashi H, Hosokawa D (2000) Novel mannose-binding rice lectin composed of some isolectins and its relation to a stress inducible salT gene. Plant Cell Physiol 41: 258–267

Hirasawa T (1999) Physiological characterization of rice plant for tolerance of water deficit, in genetic improvement of rice for water-limited environments (Ed.) by Ito O, O'Toole JC and Hardy B. International Rice Research Institute, Los Baños. Philippines pp 89-98

Hirayama T, Shinozaki K (2010) Research on plant abiotic stress responses in the post-genome era: past, present and future. Plant J 61: 1041–1052

Hoekema A, Hirsch P, Hooykass P (1983) A binary plant vector strategy based on separation of the vir and TDNA regions of the *Agrobacterium tumefaciens* Ti plasmid. Nat 303: 179-180

Hogenhout S, Ammar E, Whitfield A, Redinbaugh M (2008) Insect vector interactions with transmitted viruses. Ann Rev Phytopathol 46: 327-359

Hong CY, Chao YY, Yang MY, Cheng SY, Cho SC, Kao CH (2009) NaCl-induced expression of glutathione reductase in roots of rice (*Oryza sativa* L.) seedlings is mediated through hydrogen peroxide but not abscisic acid. Plant Soil 320: 103–115

Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT (1985) A simple and general method for transferring genes into plants. Sci 227: 1229–1231

Houlès-Astoul C, Peumans WJ, Van Damme EJM, Barre A, Bourne Y, Rougé P (2002) The size, shape and specificity of the sugar-binding site of the jacalin-related lectins is profoundly affected by the proteolytic cleavage of the subunits. Biochem J 367: 817–824

Huang F, Leonard BR, Gable RH (2006) Comparative susceptibility of European corn borer, southwestern corn borer, and sugarcane borer (Lepidoptera: Crambidae) to Cry1Ab protein in a commercial *Bacillus thuringiensis* corn hybrid. J Econ Entomol 99: 194–202

Huang P, Ju HW, Min JH, Zhang X, Kim SH, Yang KY, Kim CS (2013) Overexpression of L-type lectin-like protein kinase 1 confers pathogen resistance and regulates salinity response in *Arabidopsis thaliana*. Plant Sci 203-204: 98-106

Imanishi S, Kito-Nakamura K, Matsuoka K, Morikami A, Nakamura K (1997) A major jasmonate-inducible protein of sweet potato, ipomoelin, is an ABA-independent wound- inducible protein. Plant Cell Physiol 38: 643-652

Inthapan P, Fukai S (1988) Growth and yield of rice cultivars under sprinkler irrigation in south-eastern Queensland. 2. Comparison with maize and grain sorghum under wet and dry conditions. Aust J Exp Agric 28: 243–248

IRGSP (2005) The map-based sequence of the rice genome. Nat 436: 793–800

Ito Y, Katsura K, Maruyama K, Taji T, Kobayashi M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2006) Functional analysis of rice DREB1/CBF-type transcription factors involved in cold-responsive gene expression in transgenic rice. Plant Cell Physiol 47: 141–153

Jacobs PP, Geysens S, Vervecken W, Contreras R, Callewaert N (2009) Engineering complex-type N-glycosylation in *Pichia pastoris* using GlycoSwitch technology. Nat Protoc 4: 58-70

Jahic M, Gustavsson M, Jansen AK, Martinelle M, Enfors SO (2003) Analysis and control of proteolysis of a fusion protein in *Pichia pastoris* fed-batch processes. J Biotechnol 102: 45–53

Jain M, Nijhawan A, Tyagi AK, Khurana JP (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. Biochem Biophys Res Commun 345: 646-651

Jain M, Tyagi AK, Khurana JP (2006) Overexpression of putative topoisomerase 6 genes from rice confers stress tolerance in transgenic *Arabidopsis* plants. FEBS J 273: 5245–5260

Jain M, Tyagi AK, Khurana JP (2008) Constitutive expression of a meiotic recombination protein gene homolog, OsTOP6A1, from rice confers abiotic stress tolerance in transgenic *Arabidopsis* plants. Plant Cell Rep 27: 767–778

Jaleel CA, Manivannan P, Lakshmanan GMA, Gomathinayagam M, Panneerselvam R (2008) Alterations in morphological parameters and photosynthetic pigment responses of *Catharanthus roseus* under soil water deficits. Colloids Surf B Biointerfaces 61: 298–303

Jefferson RA (1987) Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol Biol Rep 5: 387-405

Jiang C, Shimono M, Sugano S, Kojima M, Yazawa K, Yoshida R, Inoue H, Hayashi N, Sakakibara H, Takatsuji H (2010b) Abscisic acid interacts antagonistically with salicylic acid signaling pathway in rice–*Magnaporthe grisea* interaction. Mol Plant Microbe Interact 23: 791–798

Jiang JF, Han Y, Xing LJ, Xu YY, Xu ZH, Chong K (2006) Cloning and expression of a novel cDNA encoding a mannose-specific jacalin-related lectin from *Oryza sativa*. Toxicon 47: 133-139

Jiang JF, Xu YY, Chong K (2007) Overexpression of OsJAC1, a lectin gene, suppresses the coleoptile and stem elongation in rice. J Integr Plant Biol 49: 230-237

Jiang SY, Ma Z, Ramachandran S (2010a) Evolutionary history and stress regulation of the lectin superfamily in higher plants. BMC Evolutionary Biol 10:79

Jonaliza CL, Grenggrai P, Boonrat J, Theerayut T (2004) Quantitative trait loci associated with drought tolerance at reproductive stage in rice. Plant Physiol 135: 384-399

Jongdee B, Pantuwan G, Fukai S and Fischer K (2006) Improving drought tolerance in rainfed lowland rice: An example from Thailand. Agric Water Manag 80: 225- 240

Jouanin L, Bonadé-Bottino M, Girard C, Morrot G, Giband M (1998) Transgenic plants for insect resistance. Plant Sci 131: 1–11

Jung C, Seo JS, Han SW, Koo YJ, Kim CH, Song SI, Nahm BH, Choi YD, Cheong JJ (2008) Overexpression of AtMYB44 enhances stomatal closure to confer abiotic stress tolerance in transgenic *Arabidopsis*. Plant Physiol 146: 623–635

Jung J, Won SY, Suh SC, Kim H, Wing R, Jeong Y, Hwang I, Kim M (2007) The barley ERF-type transcription factor HvRAF confers enhanced pathogen resistance and salt tolerance in *Arabidopsis*. Planta 225: 575–588

Kanneganti V, Gupta AK (2008) Overexpression of OsiSAP8, a member of stress associated protein (SAP) gene family of rice confers tolerance to salt, drought and cold stress in transgenic tobacco and rice. Plant Mol Biol 66: 445–462

Karimi M, Inzé D, Depicker A (2002) GATEWAY^{TM} vectors for *Agrobacterium*-mediated plant transformation. Trends Plant Sci 7: 193-195

Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1999) Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nat Biotechnol 17: 287–291

Kathuria H, Giri J, Tyagi H, Tyagi AK (2007) Advances in transgenic Rice Biotechnology. Crit Rev Plant Sci 26: 65–103

Kawano Y, Shimamoto K (2013) Early signaling network in rice PRR-mediated and R-mediated immunity. Current Opin Plant Biol 16: 496-504

Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai k, Galbraitha D, Bohnert HJ (2001) Gene expression profiles during the initial phase of salt stress in rice. Plant Cell 13: 889–905

Ke Y, Han G, He H, Li J (2009) Differential regulation of proteins and phosphoproteins in rice under drought stress. Biochem Biophys Res Commun 379: 133–138

Kennedy JS, Day MF, Eastop VF (1962) A conspectus of aphids as vectors of plant viruses. In: commonwealth institute of entomology. London pp 114

Kerppola TK (2008) Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. Nat Protoc 1: 1278-1286

Khatun S, Flowers TJ (1995) Effects of salinity on seed set in rice. Plant Cell Environ 18: 61–67

Khurana R, Kapoor S, Tyagi AK (2013) Spatial and temporal activity of upstream regulatory regions of rice anther-specific genes in transgenic rice and *Arabidopsis*. Transgenic Res 22: 31-46

Khush, GS (2013) Strategies for increasing the yield potential of cereals: case of rice as an example. Plant Breeding doi:10.1111/pbr.1991

Kim SH, Hong JK, Lee SC, Sohn KH, Jung HW, Hwang BK (2004a) CAZFP1, Cys2/His2-type zinc-finger transcription factor gene functions as a pathogen-induced early defense gene in *Capsicum annuum*. Plant Mol Biol 55: 883–904

Kim ST, Cho KS, Yu S, Kim SG, Hong JC, Han CD, Bae DW, Nam MH, Kang KY (2003) Proteomic analysis of differentially expressed proteins induced by rice blast fungus and elicitor in suspension-cultured rice cells. Proteomics 3: 2368–2378

Kim ST, Kim SG, Hwang DH, Young Kang SY, Koo SC, Cho MJ, Kang KY (2004b) Expression of a salt-induced protein (SALT) in suspension-cultured cells and leaves of rice following exposure to fungal elicitor and phytohormones. Plant Cell Rep 23: 256–262

Kimura Y, Kamamoto M, Maeda M, Okano M, Yokoyama M, Kino K (2005) Occurrence of Lewis a epitope in N-glycans of a glycoallergen, Jun a 1, from mountain cedar (*Juniperus ashei*) pollen. Biosci Biotechnol Biochem 69: 137–144

Kittur FS, Lalgondar M, Yu HY, Bevan DR, Esen A (2007) Maize β-glucosidase-aggregating factor is a polyspecific jacalin-related chimeric lectin, and its lectin domain is responsible for β-glucosidase aggregation. J Biol Chem 282: 7299-7311

Kocourek J, Horejsi V (1983) A note on the recent discussion on definition of the term 'lectin'. In: Lectins: Biology, Biochemistry and Clinical Biochemistry. Vol. 3 pp 3–6. Bøg- Hansen TC, Spengler G A. Eds. Walter de Gruyter, Berlin, Germany

Koga H, Dohi K, Mori M (2004) Abscisic acid and low temperatures suppress the whole plant-specific resistance reaction of rice plants to the infection of *Magnaporthe grisea*. Physiol Mol Plant Pathol 65: 3–9

Koles K, Lim JM, Aoki K, Porterfield M, Tiemeyer M, Wells L, Panin V (2007) Identification of N-glycosylated proteins from the central nervous system of *Drosophila melanogaster*. Glycobiol 17:1388–1403

Komatsu S, Yamada E, Furukuwa K (2009) Cold stress changes the concanavalin A-positive glycosylation pattern of proteins expressed in the basal parts of rice leaf sheaths. Amino Acids 36: 115-123

Koshte VL, Van Dijk W, Van Der Stelt ME, Aalberse RA (1990) Isolation and characterization of BanLec-I, a mannoside binding lectin from *Musa paradisiac* (banana). Biochem J 272: 721-726

Krieger E, Koraimann G, Vriend G (2002) Increasing the precision of comparative models with YASARA NOVA - a self-parameterizing force field. Proteins 47: 393-402

Kristensen NP, Scoble MJ, Karsholt O (2007) Lepidoptera phylogeny and systematics: the state of inventorying moth and butterfl y diversity. Zootaxa 1668: 699–747

Kuriakose B, Arun V, Gnanamanickam SS, Thomas G (2009) Tissue-specific expression in transgenic rice and *Arabidopsis thaliana* plants of GUS gene driven by the 50 regulatory sequences of an anther specific rice gene YY2. Plant Sci 177: 390-397

Kyndt T, Denil D, Haegeman A, Trooskens G, Bauters L, Criekinge WV, De Meyer T, Gheysen G (2012) Transcriptional reprogramming by root knot and migratory nematode infection in rice. New Phytol. 196: 887-900

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nat 227: 680-685

Lafitte HR, Price AH, Courtois B (2004) Yield response to water deficit in an upland rice mapping population: associations among traits and genetic markers. Theor Appl Genet 109: 1237–1246

Lafitte HR, Yongsheng G, Yan S, Li ZK (2007) Whole plant responses, key processes, and adaptation to drought stress: the case of rice. J Exp Bot 58: 169–175

Lam SK, Ng TB (2011) Lectins: production and practical applications. Appl Microbiol Biotechnol 89: 45–55

Landsteiner K, Raubitschek H (1907) Beobachtungen über Hämolyse und Hämagglutination. Zentralbl Bakteriol Parasitenk Infektionskr Hyg Abt 45: 660–667

Lannoo N, Peumans WJ, Van Damme EJM (2006b) The presence of jasmonate-inducible lectin genes in some but not all *Nicotiana* species explains a marked intragenus difference in plant responses to hormone treatment. J Exp Bot 57: 3145-3155

Lannoo N, Peumans WJ, Van Pamel E, Alvarez R, Xiong T-C, Hause G, Mazars C, Van Damme EJM (2006a) Localization and *in vitro* binding studies suggest that the cytoplasmic/nuclear tobacco lectin can interact in situ with high-mannose and complex N-glycans. FEBS Lett 580: 6329–6337

Lannoo N, Van Damme EJM (2010) Nucleocytoplasmic plant lectins. Biochim Biophys Acta 1800: 190–201

Lannoo N, Vandenborre G, Miersch O, Smagghe G, Wasternack C, Peumans, W.J, Van Damme EJM (2007a) The jasmonate-induced expression of the *Nicotiana tabacum* leaf lectin. Plant Cell Physiol 48: 1207–1218

Lannoo N, Vervecken W, Proost P, Rougé P, Van Damme EJM (2007b) Expression of the nucleocytoplasmic tobacco lectin in the yeast *Pichia pastoris*. Protein Expr Purif 53: 275-282

Laroy W, Contreras R, Callewaert N (2006) Glycome mapping on DNA sequencing equipment. Nat Protoc 1: 397-405

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemistry of protein structures. J Appl Cryst 26: 283-291

Lauriano JA, Ramalho JC, Lidon FC, Matos M (2004) Peanut photosynthesis under drought and re-watering. Photosynthetica 42: 37–41

Lawlor DW, Cornic G (2002) Photosynthetic carbon assimilation and associated metabolism in relation to water deficits in higher plants. Plant Cell Environ 25: 275–294

Lee BM, Park YJ, Park DS, Kang HW, Kim JG, Song ES, Park IC, Yoon UH, Hahn JH, Koo BS, et al. (2005) The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. Nucleic Acids Res 33: 577–586

Lee KS, Choi WY, Ko JC, Kim TS, Gregorio GB (2003) Salinity tolerance of japonica and indica rice (*Oryza sativa* L.) at the seedling stage. Planta 216: 1043–1046

Lee RH, Wang CH, Huang LT & Chen SC (2001) Leaf senescence in rice plants: cloning and characterization of senescence up-regulated genes. J Exp Bot 52: 1117– 1121

Lefebvre T, Dehennaut V, Guinez C, Olivier S, Drougat L, Mir AM, Mortuaire M, Michalski JC (2010) Dysregulations of the nutrient/stress sensor O-GlcNAcylation are involved in the etiology of cardiovascular disorders, type-2 diabetes and Alzheimer's disease. Biochim Biophys Acta 1800: 67–79

Leonard, R (2002) The presence of Lewis A epitopes in *Arabidopsis thaliana* glycoconjugates depends on an active α4-fucosyltransferase gene. Glycobiology 12: 299–306

Lerouge P, Cabanes-Macheteau M, Rayon C, Fischette-Laine AC, Gomord V, Faye L (1998) N-glycoprotein biosynthesis in plants: recent developments and future trends. Plant Mol Biol 38: 31-48

Leung J, Giraudat J (1998) Abscisic acid signal transduction. Annu Rev Plant Physiol 49: 199–222

Li Q, Chen F, Sun L, Zhang Z, Yang Y, He Z (2006) Expression profiling of rice genes in early defense responses to blast and bacterial blight pathogens using cDNA microarray. Physiol Mol Plant Pathol 68: 51–60

Liao M, Li Y, Wang Z (2009) Identification of elicitor-responsive proteins in rice leaves by a proteomic approach. Proteomics 9:2809–2819

Liu JK, Liao DQ, Oane R, Estenor L, Yang XE, Li ZC and Bennett J (2006) Genetic variation in the sensitivity of anther dehiscence to drought stress in rice. Field Crop Res 97: 87-100

Liu W, Liu J, Ning Y, Ding B, Wang X, Wang Z, Wang GL (2013) Recent progress in understanding PAMP- and effector-triggered immunity against the rice blast fungus *Magnaporthe oryzae*. Mol Plant 6: 605-620

Loreto F, Tricoli D, Marco GD (1995) On the relationship between electron transport rate and photosynthesis in leaves of the C4 plant *Sorghum bicolor* exposed to water stress, temperature changes and carbon metabolism inhibition. Aust J Plant Physiol 22: 885–892

Loris R, Hamelryck T, Bouckaert J, Wyns L (1998) Legume lectin structure. Biochim Biophys Acta 1383: 9-36

Losey JE, Eubanks MD (2000) Implications of pea aphid host-plant specialization for the potential colonization of vegetables following post-harvest emigration from forage crops. Environ Entomol 29: 1283-1288

Ma L, Zhang H, Sun L, Jiao Y, Zhang G, Miao C, Hao F (2012) NADPH oxidase AtrbohD and AtrbohF function in ROS dependent regulation of Na^+/K^+ homeostasis in *Arabidopsis* under salt stress. J Exp Bot 63: 305–317

Ma QH (2013) Monocot chimeric jacalins: a novel subfamily of plant lectins. Crit Rev Biotechnol Early Online: 1–7. DOI: 10.3109/07388551.2013.793650

Ma QH, Tian B, Li YL (2010) Overexpression of a wheat jasmonate-regulated lectin increases pathogen resistance. Biochimie 92: 187–193

Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM (2005) Heterologous protein production using the *Pichia pastoris* expression system. Yeast 22: 249–270

Macedo MLR, Damico DCS, Freire MGM, Toyama MH, Marangoni S, Novello JC (2003) Purification and characterization of an Nacetylglucosamine-binding lectin from *Koelreuteria paniculata* seeds and its effect on the larval development of

Callosobruchus maculatus (Coleoptera: Bruchidae) and *Ephestia kuehniella* (Lepidoptera: Pyralidae). J Agric Food Chem 51: 2980-2986

Mackauer M, Way MJ (1976) *Myzus persicae* Sulz. an aphid of world importance. In: Studies in Biological Control (Ed.) Delucchi VL. Cambridge University Press pp 51-119

Maclean JL, Dave DC, Hardy B, Hettel GP (2002) Rice almanac. CAB International, Wallingford

Maeda M, Kimura M, Kimura Y (2010) Intracellular and extracellular free N-glycans produced by plant cells: occurrence of unusual plant complex-type free N-glycans in extracellular spaces. J Biochem 148: 681–692

Maeda M, Kimura Y (2006) Glycoform analysis of N-glycans linked to glycoproteins expressed in rice culture cells: predominant occurrence of complex type N-glycans. Biosci Biotechnol Biochem 70: 1356-63

Mahajan S, Tuteja N (2005) Cold, salinity and drought stresses: An overview. Arch Biochem Biophys 444: 139–158

Maiti RK (2012a) Potential morpho-physiological trait selection criteria for biotic and abiotic stress resistance in crops. International Journal of Bio-resource and Stress Management. Editorial 3(2), i-ii

Maiti RK (2012b) Root responses are indicators for salinity and drought stress in crops. International Journal of Bio-resource and Stress Management Editorial 3(3), i-ii

Maiti RK (2012c) A low cost technologyfor increasing crop productivity under sustainable Agriculture. International Journal of Bio-resource and Stress Management. Editorial 3(4), i-iii

Malone LA, Gatehouse AMR, Barratt BIP (2008) Beyond Bt: alternative strategies for insect-resistant genetically modified crops. In: Romeis J, Shelton AM, Kennedy GG (Ed.) Integration of insect-resistant genetically modified crops within IPM programs. Springer. The Netherlands pp 357–417

Manschadi AM, Christopher J, devoil P and Hammer GL (2006) The role of root architectural traits in adaptation of wheat to water-limited environments. Funct Plant Biol 33: 832-837

MansWeld TJ, Atkinson CJ (1990) Stomatal behaviour in water stressed plants, in: R.G. Alscher, J.R. Cumming (Ed.), Stress Responses in Plants: Adaptation and Acclimation Mechanisms, Wiley-Liss. New York pp 241–264

Maras M, Callewaert N, Piens K, Claeyssens M, Martinet W, Dewaele S, Contreras H, Dewerte I, Penttilä M, Contreras R (2000) Molecular cloning and enzymatic characterization of a Trichoderma reesei 1,2-α-D-mannosidase. J Biotechnol 77: 255-263

Mas MT, Smith KC, Yarmush DL, Aisaka K, Fine RM (1992) Modeling the anti-CEA antibody combining site by homology and conformational search. Proteins Struc Func Genet 14: 483-498

McCord JM (2000) The evolution of free radicals and oxidative stress. Am J Med 108: 652–659

Melo F, Devos D, Depiereux E, Feytmans E (1997a) ANOLEA: a www server to assess protein structures. Proc Int Conf Intell Syst Mol Biol 5: 187-190

Melo, N (1997b) Identification of the human Lewis A carbohydrate motif in a secretory peroxidase from a plant cell suspension culture (*Vaccinium myrtillus* L.). FEBS Lett 415: 186–191

Meng CM, Zhang TZ, Guo WZ (2009) Molecular cloning and characterization of a novel *Gossypium hirsutum* L. bHLH gene in response to ABA and drought stresses. Plant Mol Biol Rep 27: 381–387

Michiels K, Van Damme EJM, Smagghe G (2010) Plant-insect interactions: what can we learn from plant lectins? Arch Insect Biochem Physiol 73: 193–212

Mittler R (2006) Abiotic stress, the field environment and stress combination. Trends Plant Sci 11: 15–19

Mittler R, Blumwald E (2010) Genetic engineering for modern agriculture: challenges and perspectives. Annu Rev Plant Biol 61: 443–462

Mohammadi-Nejad G, Singh RK, Arzani A, Rezaie AM, Sabouri H, Gregorio GB (2010) Evaluation of salinity tolerance in rice genotypes. Int J Plant Prod 4: 199–207

Molina J, Landa A, Bautista G, Martinez M, Cordoba F (2004) Molecular association of lectin and beta-glucosidase in corn coleoptile. Biochim Biophys Acta 1674: 299-304

Molinaro A, Newman MA, Lanzetta R, Parrilli M (2009) The structures of lipopolysaccharides from plant-associated gram-negative bacteria. Eur J Org Chem 2009: 5887–5896

Monzavi-Karbassi B, Luo P, Cunto-Amesty G, Jousheghany F, Pashov A, Weissman D, Kieber-Emmons T (2004) Fucosylated lactosamines participate in adhesion of HIV-1 envelope glycoprotein to dendritic cells. Arch. Virol 149: 75-91

Moons A, Bauw G, Prinsen E, Montagu MV, Van der Straeten D (1995) Molecular and physiological responses to abscisic acid and salts in roots of salt-sensitive and salt-tolerant indica rice varieties. Plant Physiol 107: 177–186

Moons A, Gielen J, Vandekerckhove J, Van Der Straeten D, Gheysen G, Van Montagu M (1997a) An abscisic-acid and salt-stress-responsive rice cDNA from a novel plant gene family. Planta 202: 443-454

Moons A, Prinsen E, Bauw G, Van Montagu M (1997b) Antagonistic effects of abscisic acid and jasmonates on salt stress-inducible transcripts in rice roots. Plant cell 9: 2243-2259

Moradi F, Ismail AM (2007) Responses of photosynthesis, chlorophyll fluorescence and ROS scavenging system to salt stress during seedling and reproductive stages in rice. Ann Bot 99: 1161–1173

Morris RT, O'Connor TR, Wyrick JJ (2008) Osiris: an integrated promoter database for *Oryza sativa* L. Bioinformatics 24: 2915–2917

Moulton JK Pepper DA, Dennehy TJ (2000) Beet armyworm (*Spodoptera exigua*) resistance to spinosad. Pest Manag Sci 56: 842–848

Munns R (2005) Genes and salt tolerance: Bringing them together. New Phytol 167: 645–663

Munns R, Tester M (2008) Mechanisms of salinity tolerance. Annu Rev Plant Biol 59: 651–681

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio-assays with tobacco tissue cultures. Physiol Plant 15: 473-497

Murasugi A, Asami Y, Mera-Kikuchi Y (2001) Production of recombinant human bile salt stimulated lipase in *Pichia pastoris*. Protein Expr and Purif 23: 282–288

Murasugi A, Tohma-Aiba Y (2003) Production of native recombinant human midkine in the yeast, *Pichia pastoris*. Protein Expr Purif 27: 244–252

Nahar K, Kyndt T, Nzogela YB, Gheysen G (2012) Abscisic acid interacts antagonistically with classical defense pathways in rice-migratory nematode interaction. New Phytol 196: 901–913

Nakagawa R, Yasokawa D, Okumura Y, Nagashima K (2000) Cloning and sequence analysis of cDNA coding for a lectin from *Helianthus tuberosus* callus and its jasmonate-induced expression. Biosci Biotechnol Biochem 64: 1247–1254

Nakamura-Tsuruta S, Uchiyama N, Peumans WJ, Van Damme EJM, Totani K, Ito Y, Hirabayashi J (2008) Analysis of the sugar binding specificity of mannose-binding type jacalin-related lectins by frontal affinity chromatography – an approach to functional classification. FEBS J 275: 1227-1239

Nakamura K, Inoue M, Yoshiie T, Hosoi K, Kimura Y (2008) Changes in structural features of free N-glycan and endoglycosidase activity during tomato fruit ripening. Biosci Biotechnol Biochem 72: 2936–2945

Nambara E, Marion-Poll A (2005) Abscisic acid biosynthesis and catabolism. Annu Rev Plant Biol 56: 165–185

Narsai R, Ivanova A, Sophia NG, Whelan W (2010) Defining reference genes in *Oryza sativa* using organ, development, biotic and abiotic transcriptome datasets. BMC Plant Biol 10: 56

Negrao S, Courtois B, Ahmadi N, Abreu I, Saibo N, Oliveira MM (2011) Recent updates on salinity stress in rice: from physiological to molecular responses. Crit Rev Plant Sci 30: 329–377

Neuman PM (1993) Rapid and reversible modifrications of extension capacity of cell walls in elongating maize leaf tisses reponding to root addition and removal of NaCl. Plant Cell Environ 16: 1107–1114

Nguyen HT, Babu RC and Blum A (1997) Breeding for drought resistance in rice: physiology and molecular genetics considerations. Crop Sci 37: 1426-1434

Nicholls A, Sharp KA, Honi, B (1991) Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins Struc Func Genet 11: 281-296

Nie X, Bressan RA, Hasegawa PM, Pardo JM (1995) Ion homeostasis in NaCl stress environments. Plant Physiol 109: 735–742

Nino-Liu DO, Ronald PC, Bogdanove AJ (2006) *Xanthomonas oryzae* pathovars: model pathogens of a model crop. Mol Plant Pathol 7: 303-324

Nolan T, Hands RE, Bustin SA (2006) Quantification of mRNA using real-time RT-PCR. Nature Protocols 1: 1559–1582

Notenboom V, Boraston AB, Williams SJ, Kilburn DG, Rose DR (2002) High-resolution crystal structures of the lectin-like xylan binding domain from *Streptomyces lividans* xylanase 10A with bound substrates reveal a novel mode of xylan binding. Biochem 41: 4246-4254

Nothaft H, Szymanski CM (2010) Protein glycosylation in bacteria: sweeter than ever. Nat Rev Microbiol 8: 765-778

O'Toole JC, Baldia EP (1982) Water deficits and mineral uptake in rice. Sci 22: 1144–1150

O'Toole JC, Cruz RT (1980) Response of leaf water potential, stomatal resistance, and leaf rolling to water stress. Plant Physiol 65: 428–432

O'Toole JC, Moya TB (1981) Water deficits and yield in upland rice. Field Crop Res 4: 247–259

Oh SJ et al. (2005) *Arabidopsis* CBF3/DREB1A and ABF3 in transgenic rice increased tolerance to abiotic stress without stunting growth. Plant Physiol 138: 341–351

Oliveira C, Felix W, Moreira RA, Teixeira JA, Domingues L (2008) Expression of frutalin, an a-d-galactose-binding jacalin-related lectin, in the yeast *Pichia pastoris*. Protein Expr Purif 60: 188–193

Oliveira CFR, Luz LA, Paiva PMG, Coelho LCBB, Marangoni S, Macedo MLR (2011) Evaluation of seed coagulant *Moringa oleifera* lectin (cMoL) as a bioinsecticidal tool with potential for the control of insects. Process Biochem 46: 498-504

Olszewski NE, West CM, Sassi SO, Hartweck LM (2010) O-GlcNAc protein modification in plants: evolution and function. Biochim Biophys Acta 1800: 49–56

Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek R L, Lee Y, Zheng L, et al. (2007) The TIGR rice genome annotation resource: improvements and new features. Nucleic Acids Res 35: 846-851

Pang CH, Wang BS (2008) Oxidative stress and salt tolerance in plants. In Progress in Botany, Lüttge U, Beyschlag W, Murata J (Ed.) (Berlin, Germany: Springer) pp 231–245

Pantuwan G, Fukai S, Cooper M, Rajatasereekul S and O'Toole JC (2002) Yield response of rice (*Oryza sativa* L.) genotypes to drought under rainfed lowlands: 2. Selection of drought resistant genotypes. Field Crop Res 73: 169-180

Parbery DG (1996) Trophism and the ecology of fungi associated with plants. Biol Rev Camb Philos Soc 71: 473–527

Passioura JB (2007) The drought environment: physical, biological and agricultural perspectives. J Exp Botan 58: 113-117

Passioura JB, Munns R (2000) Rapid environmental changes that affect leaf water status induce transient surges or pauses in leaf expansion rate. Aust J Plant Physiol 27: 941–948

Paterson AH, Freeling M, Sasaki T (2005) Grains of knowledge: genomics of model cereals. Genome Res 15:1643-50; PMID: 16339361; http://dx.doi.org/10.1101 /gr.3725905

Pedersen HL, Fangel JU, McCleary B, Ruzanski C, Rydahl MG, RaletMC, Farkas V, Schantz LV, Marcus SE, Andersen MCF, et al. (2012) Versatile high resolution oligosaccharide microarrays for plant glycobiology and cell wall research. J Bio Chem 287: 39429–39438

Petryniak J, Pereira ME, Kabat EA (1977) The lectin of *Euonymus europeus*: purification, characterization, and an immunochemical study of its combining site. Arch Biochem Biophys 178: 118-134

Peumans WJ, Hause B, Van Damme EJM (2000b) The galactose-binding and mannose-binding jacalin-related lectins are located in different subcellular compartments. FEBS Lett 477: 186–192

Peumans WJ, Stinissen HM (1983) *Gramineae* lectins: occurrence, molecular biology, and physiological function. In: Chemical Taxonomy, Molecular Biology and Function of Plant Lectins. Goldstein IJ, Etzler ME (Ed.) Alan R. Liss Inc. New York pp 99–116

Peumans WJ, Van Damme EJ (1995) Lectins as plant defense proteins. Plant Physiol 109: 347–352

Peumans WJ, Winter HC, Bemer V, Van Leuven F, Goldstein IJ, Truffa-Bachi P, Van Damme EJM (1997) Isolation of a novel plant lectin with an unusual specificity from *Calystegia sepium*. Glycoconj J 14: 259-265

Peumans WJ, Zhang W, Barre A, Houlès Astoul C, Balint-Kurti PJ, Rovira P, Rougé P, May GD, Van Leuven F, Truffa-Bachi P, et al. (2000) Fruit-specific lectins from banana and plantain. Planta 211: 546-554

Pfaffl MW, Horgan GW, Dempfle L (2002) Relative Expression Software Tool (REST(c)) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 30: e36

Pieterse CMJ, Van Wees SCM, Hoffland E, Van Pelt JA, Van Loon LC (1996) systemic resistance in *Arabidopsis* induced by biocontrol bacteria 1s independent of salicylic acid accumulation and pathogenesis-related gene expression. Plant Cell 8: 1225-1237

Ponder JW, Richards FM (1987) Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. J Mol Biol 193: 775-791

Powell KS, Gatehouse AMR, Hilder VA, Gatehouse JA (1993) Antimetabolic effects of plant lectins and plant and fungal enzymes on the nymphal stages of two important rice pests, *Nilaparvata lugens* and *Nephotettix nigropictus*. Entomol Exp Appl 66: 119–126.

Powell KS, Gatehouse AMR, Hilder VA, Van Damme EJM, Peumans WJ, Boonjawat J, Horsham K, Gatehouse JA (1995) Different antimetabolic effects of related lectins towards nymphal stages of *Nilaparvata lugens*. Entomol Exp Appl 75: 61–65

Preston A, Mandrell RE, Gibson BW, Apicella MA (1996) The lipooligosaccharides of pathogenic gram-negative bacteria. Crit Rev Microbiol 22: 139-180

Price AH (2002) QTLs for root growth and drought resistance in rice, in Molecular techniques in crop improvement (Ed.) by Mohan Jain S, Brar DS and Ahloowalia BS. Kulwer Academic Publisher. Norwell, MA, USA pp 563-584

Price AH, Cairns JE, Horton P, Jones HG, Griffiths H (2002) Linking drought-resistance mechanisms to drought avoidance in upland rice using a QTL approach: progress and new opportunities to integrate stomatal and mesophyll responses. J Exp Bot 53: 989–1004

Priem B, Gitti R, Bush CA, Gross KC (1993) Structure of ten free N-glycans in ripening tomato fruit. Plant Physiol. 102: 445–458

Puckridge DW, O'Toole JC (1980) Dry matter and grain production of rice, using a line source sprinkler in drought studies. Field Crop Res 3: 303–319

Qin QM, Zhang Q, Zhao WS, Wang YY & Peng YL (2003) Identification of a lectin gene induced in rice in response to *Magnaporthe grisea* infection. Acta Bot Sin 45: 76–81

Qiu Q, Guo Y, Quintero FJ, Pardo JM, Schumaker KS, Zhu JK (2004) Regulation of vacuolar Na$^+$/H$^+$ exchange in *Arabidopsis thaliana* by the salt-overly-sensitive (SOS) pathway. J Biol Chem 279: 207–215

Raemaekers RJM, de Muro L, Gatehouse JA & Fordham- Skelton AP (1999) Functional phytohemagglutinin (PHA) and *Galanthus nivalis* agglutinin (GNA) expressed in *Pichia pastoris*. J Biochem 265: 394–403

Rahbé Y, Sauvion N, Febvay G, Peumans WJ, Gatehouse AMR (1995) Toxicity of lectins and processing of ingested proteins in the pea aphid *Acyrthosiphon pisum*. Entomol Exp Appl 76: 143-155

Ramachandran GN, Sasisekharan V (1968) Conformation of polypeptides and proteins. Adv Protein Chem 23: 283-438

Ramamurthy M, Nina F (2003) Stress response, cell death and signalling: The many faces of reactive oxygen species. Plant Physiol 119: 56 – 68

Ranathunge K, Kotula L, Steudle E, Lafitte R (2004) Water permeability and reflection coefficient of the outer part of young rice roots are differently affected by closure of water channels (aquaporins) or blockage of apoplastic pores. J Exp Bot 55: 433–447

Ranjekar PK, Patankar A, gupta V, Bhatnagar R, Bentur J, Kumar PA (2003) Genetic engineering of crop plants for insect resistance. Curr Sci 84: 321-29

Rao KV, Rathore KS, Hodges TK, Fu X, Stoger E, Sudhakar D, Williams S, Christou P, Bharathi M, Bown DP, et al. (1998) Expression of snowdrop lectin (GNA) in transgenic rice plants confers resistance to rice brown planthopper. Plant J 15:469-477

Raval S, Gowda SB, Singh DD, Chandra NR (2004) A database analysis of jacalin-like lectins: sequence-structure-function relationships. Glycobiol 14: 1247–1263

Rendic D, Wilson IBH, Paschinger K (2008) The glycosylation capacity of insect cells. Croatica Chemica Acta 81: 7–21

Renkonen KO (1948) Studies on hemagglutinins present in seeds of some representatives of Leguminoseae. Ann Med Exp Biol Fenn 26: 66–72

Riccardi F, Gazeau P, Jacquemot MP, Vincent D, Zivy M (2004) Deciphering genetic variations of proteome responses to water deficit in maize leaves. Plant Physiol Biochem 42: 1003–1011

Rico A, Preston GM (2008) *Pseudomonas syringae* pv. tomato DC3000 uses constitutive and apoplast-induced nutrient assimilation pathways to catabolize

nutrients that are abundant in the tomato apoplast. Mol Plant Microbe Interact 21: 269–282

Rodríguez M, Canales E, Borrás-Hidalgo O (2005) Molecular aspects of abiotic stress in plants. Biotecnología Aplicada 22: 1-10

Rohila JS, Chen M, Chen S, Chen J, Cerny R, Dardick C, Canlas P, Xu X, Gribskov M, Kanrar S, et al. (2006) Protein–protein interactions of tandem affinity purification-tagged protein kinases in rice. Plant J 46: 1–13

Rougé P, Peumans WJ, Barre A, Van Damme EJM (2003) A structural basis for the difference in specificity between the two jacalin-related lectins from mulberry (*Morus nigra*) bark. Biochem Biophys Res Commun 304: 91-97

Rutenber E, Katzin BJ, Ernst S, Collins EJ, Mlsna D, Ready MP, Robertus JD (1991) Crystallographic refinement of ricin to 2.5-Å. Proteins 10: 240-250

Sadeghi A, Broeders S, De Greve H, Hernalsteens JP, Peumans WJ, Van Damme EJM, Smagghe G (2007) Expression of garlic leaf lectin under the control of the phloem-specific promoter Asus1 from *Arabidopsis thaliana* protects tobacco plants against the tobacco aphid (*Myzus nicotianae*). Pest Manag Sci 63: 1215–1223

Sadeghi A, Smagghe G, Broeders S, Hernalsteens JP, De Greve H, Peumans WJ, Van Damme EJM (2008b) Ectopically expressed leaf and bulb lectins from garlic (*Allium sativum* L.) protect transgenic tobacco plants against cotton leafworm (*Spodoptera littoralis*). Transgenic Res 17: 9–18

Sadeghi A, Smagghe G, Proost P, Van Damme EJM (2008a) Ferritin acts as a target site for the snowdrop lectin (GNA) in the midgut of the cotton leafworm *Spodoptera littoralis*. Insect Sci 15: 513–519

Sadeghi A, Smagghe G, Jurado-jácome E, Peumans WJ, Van Damme EJM (2009a) Laboratory study of the effects of leek lectin (APA) in transgenic tobacco plants on the development of cotton leafworm *Spodoptera littoralis* (Lepidoptera: Noctuidae). Eur J Entomol 106: 21–28

Sadeghi A, Van Damme EJM, Michiels K, Kabere A, Smagghe G (2009b) Acute and chronic insecticidal activity of a new mannose-binding lectin from *Allium porrum* against *Acyrthosiphon pisum* via an artificial diet. Can Entomol 141: 1-7

Sadeghi A, Van Damme EJM, Smagghe G (2009c) Evaluation of the susceptibility of the pea aphid, *Acyrthosiphon pisum*, to a selection of biorational insecticides using an artificial diet. J Insect Sci 9: 1-8

Saha P, Majumder P, Dutta I, Ray T, Roy SC, Das S (2006) Transgenic rice expressing *Allium sativum* leaf lectin with enhanced resistance against sap-sucking insect pests. Planta 223: 1329-1343

Sakuma Y, Maruyama K, Osakabe Y, Qin F, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2006) Functional analysis of an *Arabidopsis* transcription factor, DREB2A, involved in drought-responsive gene expression. Plant Cell 18: 1292–1309

Sana TR, Fischer S, Wohlgemuth G, Katrekar A, Jung KH, Ronald PC, Fiehn O (2010) Metabolomic and transcriptomic analysis of the rice response to the bacterial blight pathogen *Xanthomonas oryzae* pv. *oryzae*. Metabolomics 6: 451–465

Sanchis V, Bourguet D (2008) *Bacillus thuringiensis*: applications in agriculture and insect resistance management. Rev Agron Sustain Dev 28: 11–20

Santner A, Calderon-Villalobos LI, Estelle M (2009) Plant hormones are versatile chemical regulators of plant growth. Nat Chem Biol 5: 301–307

Sasaki T, Matsumoto T, Antonio BA, Nagamura Y (2005) From mapping to sequencing, post-sequencing and beyond. Plant Cell Physiol 46: 3–13

Sastry MVK, Banarjee P, Patanjali SR, Swamy MJ, Swarnalatha GV, Surolia A (1986) Analysis of the saccharide binding to *Artocarpus integrifolia* lectin reveals specific recognition of T-antigen (β-D-Gal(1,3)D-GalNAc). J Biol Chem 261: 11726-11733

Sauvion N, Charles H, Febvay G, Rahbé Y (2004a) Effects of jackbean lectin (ConA) on the feeding behavior and kinetics of intoxication of the pea aphid, *Acyrthosiphon pisum*. Entomol Exp Appl 10: 34–44

Sauvion N, Nerdon C, Febvay G, Gatehouse AMR, Rahbé Y (2004b) Binding of the insecticidal lectin Concanavalin A in pea aphid, *Acyrthosiphon pisum* (Harris) and induced effects on the structure of midgut epithelial cells. J Insect Physiol 50: 1137–1150

Saxena RC, Khan ZR (1989) Factors affecting resistance of rice varieties to planthopper and leafhopper pests. Agri Zool Rev 3: 97–132

Schachter H (2009) Paucimannose N-glycans in *Caenorhabditis elegans* and *Drosophila melanogaster*. Carbohydr Res 344: 1391–1396

Schouppe D, Rougé P, Lasanajak Y, Barre A, Smith DF, Proost P, Van Damme EJM (2010) Mutational analysis of the carbohydrate binding activity of the tobacco lectin. Glycoconj J 27: 613-623

Schouppe D, Van Damme EJM (2011) Plant lectins. in: Brooks S, Appelmelk B, Rudd P, (Eds.), Glycosylation in diverse cell systems—challenges and new frontiers in experimental glycobiology. In Essential Reviews in Experimental Biology 4: 45-71

Schroder LA, Glick BS, Dunn WA (2007) Identification of pexophagy genes by restriction enzyme-mediated integration. Methods Mol Biol 389: 203-218

Schroeder JI, Allen GJ, Hugouvieux V, Kwak JM, Waner D (2001a) Guard cell signal transduction. Annu Rev Plant Physiol Plant Mol Biol 52: 627–658

Schroeder JI, Kwak JM, Allen GJ (2001b) Guard cell abscisic acid signalling and engineering drought hardiness in plants. Nature 410: 327–330

Senthil-Nathan S, Choi MY, Paik CH, Kalaivani K (2008) The toxicity and physiological effect of goniothalamin, a styryl-pyrone, on the generalist herbivore, *Spodoptera exigua* Hubner. Chemosphere 72: 1393-1400

Seo YS, Chern M, Bartley LE, Han M, Jung KH, Lee I, Walia H, Richter T, Xu X, Cao P, et al. (2011) Towards establishment of a rice stress response interactome. PLoS Genet7: e1002020. doi: 10.1371/journal.pgen.1002020

Serraj R, Kumar A, McNally KL, Slamet-Loedin I, Bruskiewich R, Mauleon R, Cairns J, Hijmans RJ (2009) Improvement of drought resistance in rice. Adv Agron 103: 41–99

Sesma A, Osbourn AE (2004) The rice leaf blast pathogen undergoes developmental processes typical of root-infecting fungi. Nat 431: 582–586

Shahidi-Noghabi S, Van Damme EJM, Smagghe G (2009) Expression of *Sambucus nigra* agglutinin (SNA-I') from elderberry bark in transgenic tobacco plants results in enhanced resistance to different insect species. Transgenic Res 18: 249-259

Shahmuradov IA, Gammerman AJ, Hancock JM, Bramley PM, Solovyev VV (2003) PlantProm: a database of plant promoter sequences. Nucleic Acids Res 31: 114-117

Sharma HC, Sharma KK, Crouch JH (2004) Genetic transformation of crops for insect resistance: Potential and limitations. Crit Rev Plant Sci 23: 47-72

Sharp RE, Poroyko V, Hejlek LG, Spollen WG, Springer GK, Bohnert HJ, Nguyen HT (2004) Root growth maintenance during water deficits: physiology to functional genomics. J Exp Bot 55: 343-351

Shelton AM, Zhao JZ, Roush RT (2002) Economic, ecological, food safety, and social consequences of the deployment of Bt transgenic plants. Annu Rev Entomol 47: 845–881

Shibuya N, Minami E (2001) Oligosaccharide signalling for defense responses in plant. Physiol Mol Plant Pathol 59: 223–233

Shinjo A, Araki Y, Hirano K, Arie T, Ugaki M, Tohru T (2011) Transgenic rice plants that over-express the mannose-binding rice lectin have enhanced resistance to rice blast. J Gen Plant Pathol 77: 85-92

Shiyou Lu, Hongya Gu, Xiaojing Yuan, Xiaoming Wang, Ai-Min Wu, Lijia Qu, Jin-Yuan Liu (2007)The GUS reporter-aided analysis of the promoter activities of a rice metallothionein gene reveals different regulatory regions responsible for tissue-specific and inducible expression in transgenic *Arabidopsis*. Transgenic Res 16:177–191

Singh DD, Saikrishnan K, Kumar P, Surolia A, Sekar K, Vijayan M (2005) Unusual sugar specificity of banana lectin from *Musa pardisiaca* and its probable evolutionary origin. Crystallographic and modelling studies. Glycobiol 15: 1025-1032

Smith DF, Song X, Cummings RD (2010) Use of glycan microarrays to explore specificity of glycan-binding proteins. In: Minoru Fukuda (Ed.) Meth Enzymol Vol. 480, Burlington: Academic Press pp 417-444

Song WY, Chen LL, Pi LY, Gardner J, Wang B, Zhai WX, Zhu LH, Fauquet C, Ronald P (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. Sci 270: 1804-1806

Sreekrishna K, Brankamp RG, Kropp KE, Blankenship DT, Tsay JT, Smith PL, Wierschke JD, Subramaniam A, Birkenberger LA (1997) Strategies for the optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. Gene 190: 55–62

Sreenivasulu N, Sopory SK, Kavi Kishor PB (2007) Deciphering the regulatory mechanisms of abiotic stress tolerance in plants by genomic approaches. Gene 388: 1–13

Sreenivasulu N, Varshney RK, Kavi Kishor PB, Weschke W (2004) Tolerance to abiotic stress in cereals: a functional genomics approach. In: Gupta PK, Varshney RK (Ed.) Cereal Genomic pp 483–514

Sripinyowanich S, Klomsakul P, Boonburapong B, Bangyeekhun T, Asami T, Gu H, Buaboocha T, Chadchawan S (2013) Exogenous ABA induces salt tolerance in indica rice (*Oryza sativa* L.): The role of OsP5CS1 and OsP5CR gene expression during salt stress. Environ Exp Bot 86: 94-105

Stanley P, Cummings RD (2009) Structures common to different glycans. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME (Ed.) Essentials of Glycobiology, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 2nd edition, Chapter 13 pp 175-198

Štefanić P, Sikić S, Cvjetko P, Balen B (2012) Cadmium and zinc induced similar changes in protein and glycoprotein patterns in tobacco (*Nicotiana tabacum* L.) seedlings and plants. Arh Hig Rada Toksikol 63: 321-335

Stefanowicz K, Lannoo N, Proost P, Van Damme EJM (2012) *Arabidopsis* F-box protein containing a Nictaba-related lectin domain interacts with N-acetyllactosamine structures. FEBS Open Bio 2: 151-158

Stillmark H (1888) Ricin, ein giftiges ferment aus den samen von Ricinus communis L. und einige anderen Euphorbiaceen. Inaugural Dissertation Dorpat, Tartu

Stoger E, Williams S, Christou P, Down RE, Gatehouse JA (1999) Expression of the insecticidal lectin from snowdrop (*Galanthus nivalis* agglutinin; GNA) in transgenic wheat plants: effects on predation by the grain aphid *Sitobion avenae*. Mol Breed 5: 65–73

Strahl-Bolsinger S, Gentzsch M, Tanner W (1999) Protein O-mannosylation. Biochim Biophys Acta 1426: 297–307

Subramanyam S, Smith DF, Clemens JC, Webb MA, Sardesai N, Williams CE (2008) Functional characterization of HFR1, a high-mannose N-glycan-specific wheat lectin induced by Hessian fly larvae. Plant Physiol 147: 1412–1426

Swanson MD, Winter HC, Goldstein IJ, Markovitz DM (2010) A lectin isolated from bananas is a potent inhibitor of HIV replication. J Biol Chem 285: 8646–8655

Tabary F, Balandreau J, Bourrillon R (1984) Purification of the rice embryo lectin and its binding to nitrogen-fixing bacteria from the rhizosphere of rice. Biochem Biophys Res Commun 119: 549-555

Tabary F, Frenoy JP (1985) Interaction of rice (*Oryza sativa*) lectin with N-acetylglucosaminides. Biochem J 229: 687-692

Takahashi H, Hotta Y, Hayashi M, Kawai-Yamada M, Komatsu S, Uchimiya H (2005) High throughput metabolome and proteome analysis of transgenic rice plants (*Oryza sativa* L.). Plant Biotechnology J 22: 47–50

Talbot NJ (2003) On the trail of a cereal killer: exploring the biology of *Magnaporthe grisea*. Annu Rev Microbiol 57: 177–202

Tanaka T, Antonio BA, Kikuchi S, Matsumoto T, Nagamura Y, Numa H, Sakai H, Wu J, Itoh T, Sasaki T, et al. (2008) The rice annotation project database (RAP-DB): 2008 update. Nucleic Acids Res 36: 1028-1033

Tang CF, Liu YG, Zeng GM, Li CF, Xu WH (2004) Effects of cadmium stress on active oxygen generation, lipid peroxidation and antioxidant enzyme activities in radish seedlings. J Plant Physiol Mol Biol 30: 469-474

Tang J, Xia H, Li D, Cao M, Tao Y, Tong W, Zhang X, Hu S, Wang J, Yu J, Yang H, Zhu L (2005) Gene expression profiling in rice young panicle and vegetative organs and identification of panicle-specific genes through known gene functions. Mol Gen Genomics 274: 467–476

Tester M, Davenport R (2003) Na$^+$ tolerance and Na$^+$ transport in higher plants. Ann Bot 91: 503–527

Tester M, Langridge P (2010) Breeding technologies to increase crop production in a changing world. Sci 327: 818–822

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tool. Nucleic Acids Res 15: 4876-4882

Tian XL, Wu XL, Li Y, Zhang SQ (2005) The effect of gamma-aminobutyric acid in superoxide dismutase, peroxiduse and catalase activity response to salt stress in maize seedling. Acta Biol Exp Sin 38: 75-79

Tiemeyer M, Selleck SB, Esko JD (2008) Arthropoda. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME. Essentials in glycobiology. Plainview, NY: Cold Spring Harbor Laboratory Press. p 347–362.

Treiber N, Reinert DJ, Carpusca I, Aktories K, Schulz GE (2008) Structure and mode of action of a mosquitocidal holotoxin. J Mol Biol 381: 150-159

Tsuda M (1979) Biochem J (Tokyo) 86, 1451-1461 Van Landschoot, Loontiens A, Clegg FG, R. Sharon RM, De Bruyne NCK (1977) Eur J Biochem 79: 275-283

Tucker SL, Talbot NJ (2001) Surface attachment and pre-penetration stage development by plant pathogenic fungi. Annu Rev Phytopathol 39: 385–417

Turner NC (1986) Crop water deficits: a decade of progress. Adv Agron 39: 1–51

Umezawa T, Fujita M, Fujita Y, Yamaguchi-Shinozaki K, Shinozaki K (2006) Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. Curr Opin Plant Biotech 17: 113–122

Vaid N, Pandey PK, Tuteja N (2012) Genome-wide analysis of lectin receptor-like kinase family from *Arabidopsis* and rice. Plant Mol Biol 80: 365-88

Valliyodan B, Nguyen HT (2006) Understanding regulatory networks and engineering for enhanced drought tolerance in plants. Curr Opin Plant Biotech 9: 189–195

Van Bockhaven J, De Vleesschauwer D, Höfte M (2013) Towards establishing broad-spectrum disease resistance in plants: silicon leads the way. J Exp Bot 64: 1281-1293

Van Breusegem F, Dat JF (2006) Reactive oxygen species in plant cell death. Plant Physiol 141: 384–390

Van Damme EJM, Lannoo N, Peumans WJ (2008) Plant lectins. Adv Bot Res 48: 107–209

Van Damme EJM (2008) Plant lectins as part of the plant defense system against insects. In: Schaller A (Ed.) Induced Plant Resistance to Herbivory. Springer Science pp 285–307

Van Damme EJM, Allen AK, Peumans WJ (1987) Isolation and characterization of a lectin with exclusive specificity towards mannose from snowdrop (*Galanthus nivalis*) bulbs. FEBS lett 215: 140-144

Van Damme EJM, Barre A, Mazard AM, Verhaert P, Horman A, Debray H, Rougé P, Peumans WJ (1999) Characterization and molecular cloning of the lectin from *Helianthus tuberosus*. Eur J Biochem 259: 135-142

Van Damme EJM, Barre A, Rougé P, Peumans WJ (2004b) Cytoplasmic/ nuclear plant lectins: a new story. Trends Plant Sci 9: 484–489

Van Damme EJM, Barre A, Verhaert P, Rougé P, Peumans WJ (1996) Molecular cloning of the mitogenic mannose/maltose-specific rhizome lectin from *Calystegia sepium*. FEBS Lett 397: 352-356

Van Damme EJM, De Clerq N, Claessens F, Hemschoote K, Peeters BP, Peumans WJ (1991) Molecular cloning and characterisation of multiple isoforms of the snowdrop (*Galanthus nivalis* L.) lectin. Planta 186: 35–43

Van Damme EJM, Fouquaert E, Lannoo N, Vandenborre G, Schouppe D, Peumans, WJ (2011) Novel concepts about the role of lectins in the plant cell. In: The Molecular Immunology of Complex Carbohydrates-3. Wu AM, editor. Springer New York, USA, Volume 705 pp 295–324

Van Damme EJM, Hao Q, Chen Y, Barre A, Vandenbussche F, Desmyter S, Rougé P, Peumans WJ (2001) Ribosome-inactivating proteins: a family of plant proteins that do more than inactivate ribosomes. Crit Rev Plant Sci 20: 395-465

Van Damme EJM, Hause B, Hu J, Barre A, Rougé P, Proost P, Peumans WJ (2002) Two distinct jacalin-related lectins with a different specificity and subcellular location are major vegetative storage proteins in the bark of the black mulberry tree. Plant Physiol 130: 757-769

Van Damme EJM, Lannoo N, Fouquaert E, Peumans WJ (2004a) The identification of inducible cytoplasmic/nuclear carbohydrate binding proteins urges to develop novel concepts about the role of plant lectins. Glycoconj J 20: 449–460

Van Damme EJM, Nakamura-Tsuruta S, Smith DF, Ongenaert M, Winter HC, Rougé P, Goldstein IJ, Hanqing MO, Kominami J, Culerrier R, et al. (2007b) Phylogenetic and specificity studies of two-domain GNA-related lectins: generation of multispecificity through domain duplication and divergent evolution. Biochem J 404: 51–61

Van Damme EJM, Peumans WJ, Barre A, Rougé P (1998) Plant lectins: a composite of several distinct families of structurally and revolutionary related proteins and diverse biological roles. Crit Rev Plant Sci 17: 575-692

Van Damme EJM, Rougé P, Peumans WJ (2007a) Carbohydrate-protein interactions: plant lectins. In Comprehensive Glycoscience - From Chemistry to Systems Biology (Kamerling JP, Boons GJ, Lee YC, Suzuki A, Taniguchi N, Voragen AGJ, (Ed.) Elsevier, NY pp 563-599

Van Dammes EJM, Fouquaert E, Lannoo N, Vandenborre G, Schouppe D, Peumans WJ (2010) Novel concepts about the role of lectins in the plant cell. In: the Molecular Immunology of Complex Carbohydrate (Ed.) Wu AM, Springer pp 295-324

Van Emden H and Harrington R (2007) Aphids as crop pests. CAB international pp 699

Van Guilder HD, Vrana KE, Freeman WM (2008) Twenty-five years of quantitative PCR for gene expression analysis. Biotechniques 44: 619–626

Van Hove J, Fouquaert E, Smith DF, Proost P, Van Damme EJM (2011) Lectin activity of the nucleocytoplasmic EUL protein from *Arabidopsis thaliana*. Biochem Biophys Res Commun 414: 101–105

Vandenborre G, Miersch O, Hause B, Smagghe G, Wasternack C, Van Damme EJM (2009a) *Spodoptera littoralis*-induced lectin expression in tobacco. Plant Cell Physiol 50: 1142–1155

Vandenborre G, Smagghe G, Gesquière B, Menschaert G, Rao RN, Gevaert K, Van Damme EJM (2011a) Diversity in protein glycosylation among insect species. PloS ONE 6: e16682. doi:10.1371/journal.pone.0016682

Vandenborre G, Smagghe G, Van Damme EJM (2011b) Plant lectins as defense proteins against phytophagous insects. Phytochem 72: 1538–1550

Vandenborre G, Van Damme EJM, Smagghe G (2009b) *Nicotiana tabacum* agglutinin expression in response to different biotic challengers. Arthropod-Plant Interact 3: 193–202

Vandenborre G, Groten K, Smagghe G, Lannoo N, Baldwin IT, Van Damme EJM (2010a) *Nicotiana tabacum* agglutinin is active against Lepidopteran pest insects. J Exp Bot 61: 1003-1014

Vandenborre G, Van Damme EJM, Ghesquière B, Menschaert G, Hamshou M, Rao RN, Gevaert K, Smagghe G (2010b) Glycosylation signatures in *Drosophila*: fishing with lectins. J Proteome Res 9: 3235–3242

Vandoorne B, Mathieu AS, Van den Ende W, Vergauwen R, Périlleux C, Javaux M, Lutts S (2012) Water stress drastically reduces root growth and inulin yield in *Cichorium* (var.*sativum*) independently of photosynthesis. J Exp Bot 63: 4359-4373

Vinocur B, Altman A (2005) Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. Curr Opin Biotech 16: 123–132

Waljuno K, Scholma RA, Beintema J, Mariono A, Hahn AM (1975) Amino acid sequence of hevein. In: Proceedings of the International Rubber Conference, Kuala Lumpur, Vol 2. Rubber Research Institute of Malaysia. Kuala Lumpur pp 518–531

Wang G, Ge Z, Rasko DA, Taylor DE (2000) Lewis antigens in Helicobacter pylori: biosynthesis and phase variation. Mol Microbiol 36: 1187-1196

Wang W, Mo J, Cheng J, Zhuang P, Tang Z (2006) Selection and characterization of spinosad resistance in *Spodoptera exigua* (Hübner) (Lepidoptera: Noctuidae). Pest Biochem Physiol 80: 180-187

Wang W, Vinocur B, Altman A (2003) Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. Planta 218: 1–14

Wang XM, Ma QH (2005) Characterization of a jasmonate-regulated wheat protein related to a beta-glucosidase-aggregating factor. Plant Physiol Biochem 43: 185–92

Wang Y, Barbacioru C, Hyland F, Xiao W, Hunkapiller KL, Blake J, Chan F, Gonzalez C, Zhang L, Samaha RR (2006) Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays. BMC Genomics 7: 59

Wang YQ, Zhu SY, Wang Y, Zhang MY (2007) Tissue and inducible expression of a rice glutathione transporter gene promoter in transgenic *Arabidopsis*. Botanical Studies 48: 35-41

Wang Z, Zhang K, Sun X, Tang K and Zhang J (2005) Enhancement of resistance to aphids by introducing the snowdrop lectin gene gna into maize plants. J Biosci 30: 627–638

Whalen MC, Innes RW, Bent AF, Staskawicz BJ (1991) Identification of *Pseudomonas syringae* pathogens of *Arabidopsis* and a bacterial locus determining avirulence on both *Arabidopsis* and soybean. Plant Cell 3: 49–59

Wilkins TA, Raikhel NV (1989 Expression of rice lectin is governed by two temporally and spatially regulated mRNAs in developing embryos. Plant Cell 1: 541–549

Wilkinson S, Davies WJ (2002) ABA-based chemical signalling: the coordination of responses to stress in plants. Plant Cell Environ 25: 195–210

Williams CE, Collier CC, Nemacheck JA, Liang C, Cambron SE (2002) A lectin-like gene responds systemically to attempted feeding by avirulent firstinstar Hessian fly larvae. J Chem Ecol 28: 1411–28

Wilson RA, Talbot NJ (2009) Under pressure: investigating the biology of plant infection by *Magnaporthe oryzae*. Nat Rev Microbiol 7: 185–195

Witcombe JR, Hollington PA, Howarth CJ, Reader S, Steele KA (2008) Breeding for abiotic stresses for sustainable agriculture. Philos Trans R Soc Lond B Biol Sci 363: 703–716

Xiang Y, Tang N, Du H, Ye H, Xiong L (2008) Characterization of OsbZIP23 as a key player of bZIP transcription factor family for conferring ABA sensitivity and salinity and drought tolerance in rice. Plant Physiol 148: 1938–1952

Xu GY, Rocha PSCF, Wang ML, Xu ML, Cui YC, Li LY, Zhu YX, Xia, X (2011) A novel rice calmodulin-like gene, OsMSR2, enhances drought and salt tolerance and increases ABA sensitivity in *Arabidopsis*. Planta 234: 47–59

Xu J, Audenaert K, Hofte M, De Vleesschauwer D (2013) Abscisic acid promotes susceptibility to the rice leaf blight pathogen *Xanthomonas oryzae* pv *oryzae* by suppressing salicylic acid-mediated defenses. PLoS ONE 8: e67413. doi:10.1371/journal.pone.0067413

Xu K, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Baile-Serres J, Ronald PC, Mackill DJ (2006) Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nat 422: 705–708

Xu X, Song Y, Li Y, Chang J, zhang H, An L (2010) The tandem affinity purification method: An efficient system for protein complex purification and protein interaction identification. Protein Exp Purif 72: 149–156

Yamaguchi-Shinozaki K, Shinozaki K (2005) Organization of cis-acting regulatory elements in osmotic-and cold-stress-responsive promoters. Trends Plant Sci 10: 88–94

Yan S, Tang Z, Su W, Sun W (2005) Proteomic analysis of salt stress-responsive proteins in rice roots. Proteomics 5: 235-244

Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN (1982) Living with water stress: evolution of osmolyte systems. Sci 217: 1214–1222

Yang J, Zhang J, Liu L, Wang Z and Zhu Q (2002) Carbon remobilization and grain filling in japonica/indica hybrid rice subjected to postanthesis water deficits. Agron J 94: 102-109

Yang J, Zhang J, Wang Z, Zhu Q and Wang W (2001a) Remobilization of carbon reserves in response to water deficit during grain filling of rice. Field crops Res 71: 47-55

Yang J, Zhang J, Wang Z, Zhu Q and Wang W (2001b) Hormonal changes in the grains of rice subjected to water stress during grain filling. Plant Physiol 127: 315-323

Yang Y, Xu S, An L, Chen N (2007) NADPH oxidase dependent hydrogen peroxide production, induced by salinity stress, may be involved in the regulation of total calcium in roots of wheat. J Plant Physiol 164: 1429–1435

Yang Z, Lu Q, Wen X, Chen F, Lu (2012) Functional analysis of the rice rubisco activase promoter in transgenic *Arabidopsis*. Biochem Biophys Res Commun 418: 565-570

Yazawa K, Jiang CJ, Kojima M, Sakakibara H, Takatsuji H (2012) Reduction of abscisic acid levels or inhibition of abscisic acid signaling in rice during the early phase of *Magnaporthe oryzae* infection decreases its susceptibility to the fungus. Physiol Mol Plant Pathol 78: 1–7

Yeo AR, Flowers TJ (1983) Varietal differences in the toxicity of sodium-ions in rice leaves. Physiol Plantarum 59: 189–195

Yeo AR, Lee KS, Izard P, Boursier PJ, Flowers TJ (1991) Short and long-term effects of salinity on leaf growth in rice (*Oryza sativa* L.). J Exp Bot 42: 881–889

Yeo AR, Yeo ME, Flowers SA, Flowers TJ (1990) Screening of rice (*Oryza sativa* L) genotypes for physiological characters contributing to salinity resistance, and their relationship to overall performance. Theor App Gen 79: 377–384

Yi K, Wu Z, Zhou J, Du L, Guo L, Wu Y, Wu P (2005) OsPTF1, a novel transcription factor involved in tolerance to phosphate starvation in rice. Plant Physiol 138: 2087–2096

Yokotani N, Ichikawa T, Kondou Y, Maeda S, Iwabuchi M, Mori M, Hirochika H, Matsui M, Oda K (2009) Overexpression of a rice gene encoding a small C2 domain

protein OsSMCP1 increases tolerance to abiotic and biotic stresses in transgenic *Arabidopsis*. Plant Mol Biol 71: 391–402

Yong WD, Xu YY, Xu WZ, Wang X, Li N, Wu JS, Liang TB, Chong K, Xu ZH, Tan KH, Zhu ZQ (2003) Vernalization-induced flowering in wheat is mediated by a lectin-like gene VER2. Planta 217: 261–270

Zeng L, Shannon MC, Grieve CM (2002) Evaluation of salt tolerance in rice genotypes by multiple agronomic parameters. Euphytica 127: 235–245

Zhang F, Zhu L, He G (2004) Differential gene expression in response to brown planthopper feeding in rice. J Plant Physiol 161: 53–62

Zhang HY, Liu XZ, Wei L, Zhou LY, Yang YM (2007) Transgenic tobacco plants containing Bt and GNA genes. Biologia Plantarum 51: 746-748

Zhang J, Nguyen H and Blum A (1999) Genetic analysis of osmotic adjustment in crop plants. J Exp Bot 50: 291-302

Zhang Q (2007) Strategies for developing green super rice. Proc Natl Acad Sci 104: 16402–16409

Zhang W, Peumans WJ, Barre A, Astoul CH, Rovira P, Rougé P, Proost P, Truffa-Bachi P, Jalali AA, Van Damme EJM (2000) Isolation and characterization af a jacalin-related mannose-binding lectin from salt-stressed rice (*Oryza sativa*) plants. Planta 210: 970-978

Zhu GY, Kinet JM, Lutt S (2004) Characterisation of rice (*Oryza sativa*) F3 populations selected for salt resitance. 2. Relationship between yield-related parameters and physiological properties. Aust J Exp Agric 44: 333–342

Zhu JK (2001) Plant salt tolerance. Trends Plant Sci 6: 66–71

Zhu JK (2002) Salt and drought stress signal transduction in plants. Annu Rev Plant Biol 53: 247–273

Zhu JK (2003) Regulation of ion homeostasis under salt stress. Curr Opin Plant Biol 6: 441–445

Acknowledgements

I had the first lesson on DNA when I was at high school. Since then, I have been dreaming that one day I will be studying something related to DNA (I did not know what) and will acquire a PhD degree in this field. Finally, my dream has become true. Definitely, without the support and help of a lot of people around me, this dream would never have been accomplished.

First of all I would like to express my deep appreciation to my promoter **Prof. Dr. Els Van Damme**. Thank you very much for your patience and your scientific guidance. Your office door was always open for me. I am really grateful for your help and support. Without your fundamental contribution, this work could not be performed. Thank you very much **Els**!

I would also like to thank the members of the examination committee: **Prof. Dr. Kris Verheyen**, **Prof. Dr. Wim Van den Ende**, **Prof. Dr. Guy Smagghe**, **Prof. Dr. Monica Höfte** and **Dr. Tina Kyndt**.

Very special thanks are directed to the "Special Research Fund" (**BOF**)/ Ghent University for the financial support of this project. My sincere thanks also go to the "General Commission for Scientific Agricultural Research (**GCSAR**)"/ Syria.

I would like to take this opportunity to express my deepest gratitude to the **Glyco** group: **Dr. Nausicaä Lannoo**, **Jonas Van Hove**, **Dr. Annelies Delporte**, **Karolina Stefanowicz**, **Tomasz Walski**, **Chenjing Shang**, **Ying Chen, Lore Eggermont**, **Sofie Van Holle**, **Liuyi Dang, Isabel Verbeke** and **Jeroen de Zaeytijd**. I will miss you all guys! Thank you very much for your help and support. Thank you for the non-forgettable moments and the lovely atmosphere. You are all very kind and very nice. I like you very much and you are all very close to my heart. I wish you all the best and for those who still have to finalize their PhD I wish you a lot of success.

I will not forget to thank all the people who already left the group: **Dr. Elke Fouquaert**, **Dr. Gianni Vandenborre**, **Dr. Dieter Schouppe**, **Romina Termote-Verhalle**, **Kirsten Plas**, **Isabelle Tilmant** and **Dr. Megumi Maeda**. I would also like to thank all people from the department of **Molecular Biotechnology**. My thanks are also extended to the people from Agrozology and Phytopathology labs, especially to **Dr. David De Vleesschauwer**.

To my **father**, **Fahed Al Atalah**, and my **mothe**r, **Nashmya Abo Aasy**, thank you very much for your love, prayers and encouragements throughout all my life. I still remember your words when you told me you can do it! We very much trust you. Finally I did it because of your support and your deep emotions that I always feel them around me to protect and save me. I am so grateful and I am proudly presenting this PhD work to you.

To my **brothers** and **sisters** and their families: **Hossam**, his wife **Johina** and their lovely child **Layth**, **Basema**, her husband **Yehya** and their beautiful children **Hadeel** and **Hasan**, **Hisham**, his wife **Shireen** and their beautiful children **Shada** and **Shadi**, **Wasema**, her husband **Khaiser** and their beautiful children **Zaenah** and **Omran**, **Haifa'a**, her husband **Feda'a** and their lovely child **Loai**, my little brother **Wisam** and my little sister **Reham**. To all of you, thank you very much for your love and your support. You mean a lot to me, I wish you joyful and happy lives.

To my **father in law**, **Ammar Abo Mogdeb**, and my **mother in law**, **Aml Al Nasar**, thank you very much for your love and support. I would like to express my deepest gratitude for the beautiful flower that you gave to me! My sincere thanks are also extended to **my brothers in law**, **Oudai** and **Qutiba**. I wish you the best for your lives.

To my closest and dearest friends, **Bassam Al Barbour**, his wife **Ashwaq** and their lovely child **Al Hakam**, **Bashar Ghazali**, his wife **Siita** and their lovely children **Lein**, **Rwan** and **Karum** , **Nawras Abo Shaheen**, his wife **Nuha** and their lovely child **Maisam**, **Radi Abo Trabi** and his family and **Tahseen Al Shaer**. Thank you very much for your love and support. Do you remember, **Nawras**, I told you once that I would like to get a PhD degree, I finally did it! I wish to all of you joyful and happy lives. My thanks are also extended to all my University (Faculty of Agriculture/ Syria) colleagues and friends, especially to **Jalal Ghazali**. It is impossible to list all your names here but I wish to all of you happy and successful lives.

I would like to thank all the teachers that taught me since I was at the preliminary school until I got my PhD degree, and specifically teacher **Aqbal Subh**. Thank you all very much for your encouragements and support.

I would also like to thank all my friends, all people in my village, **Alhwaia**, all my neighbors and all my relatives. Thank you very much for your support and interest.

To my lovely, wonderful, kind and dear wife, **Rema Abo Mogdeb**, You are the rose of my life; you are a gift from God; you are my angel and you have become my spirit and my whole life. I am really in debt for your continuous love, your warm emotions, your deep interest, your complete support and your kind patience. You mean everything to me, my sweetheart. I love you very very very much!

Bassam Al Atalah

December 2013

Curriculum Vitae

1. Personal information

Name: Bassam

Family name: Al Atalah

Birth date: 15th October 1977

Nationality: Syrian

Civil status: Married

E-mail: bassamatalah@hotmail.com

2. Education

2010-2013   PhD in Applied Biological Sciences, Faculty of Bioscience Engineering, Ghent University, Belgium

PhD research entitled **'Use of carbohydrate-binding proteins to increase stress tolerance of crops'**

2004 – 2006   M.Sc. in Biotechnology 'Applied Biotechnology' under the European project (TEMPUS, JEP-CD-30018-2002), Faculty of Agriculture, Damascus University, Syria

Master research project entitled **'Characterization and expression analysis of new carbohydrate binding proteins'**, Faculty of Bioscience Engineering, Ghent University, Belgium

2003 – 2004   High studies Diploma, Faculty of Agriculture, Damascus University, Syria

Diploma Research project entitled **'The physiological and breeding criteria for increasing the wheat grain yield under irrigated and non-irrigated conditions'**

1996 – 2003   B.Sc. in Agronomy, Spec. in Field Crops, Faculty of Agriculture, Damascus University, Damascus, Syria

Research report entitled **'Improving a screening technique for some wheat cultivars for water stress tolerance'**

3. Career

Job in Syria

2004 onwards   Research Assistant (permanent post) at the General Commission for Scientific Agricultural Research (GCSAR), Syria

Supervision of undergraduate students

Vroman Jolien (2011-2012) 'Expressie-analyse van enkele lectinen uit rijst'

Sylwia Struk (2011-2012) 'Localization and Expression Analyses of Some Rice Lectins'

Maxim Braeckevelt (2012-2013) 'Expressie-analyse van rijstlectinen in transgene *Arabidopsis thaliana* planten die onderhevig zijn aan stress'

Pieter De Beule (2012-2013) 'Expressie van *Oryza sativa* agglutinine uit rijst in transgene tabaksplanten en het effect op insecten'

4. Languages

Arabic: mother tongue

English: fluent

5. Publications

Peer reviewed manuscripts

Al Atalah B, Fouquaert E, Vanderschaeghe D, Proost P, Balzarini J, Smith D.F, Rougé P, Lasanajak Y, Callewaert N, Van Damme EJM (2011) Expression analysis of the nucleocytoplasmic lectin 'Orysata' from rice in *Pichia pastoris*. FEBS J 278: 2064–2079

Al Atalah B, Rougé P, Smith DF, Proost P, Lasanajak Y, Van Damme EJM (2012) Expression analysis of a type S2 EUL-related lectin from rice in *Pichia pastoris*. Glycoconj J 29: 467–479

Al Atalah B, Fouquaert E, Van Damme EJM (2013) Promoter analysis for three types of EUL-related rice lectins in transgenic *Arabidopsis*. Plant Mol Biol Rep 31: 1315–1324

Al Atalah B, Vanderschaeghe D, Bloch Y, Proost P, Plas K, Callewaert N, Savvides SN, Van Damme EJM (2013) Characterization of a type D1A EUL-related lectin from rice expressed in *Pichia pastoris*. Biol Chem, in press

Manuscripts submitted

Al Atalah B, Smagghe G, Van Damme EJM (2013) Insecticidal activity of Orysata, a jacalin-related lectin from rice, against biting-chewing and piercing-sucking insects.

Al Atalah B, De Vleesschauwer D, Xu J, Fouquaert E, Höfte M, Van Damme EJM (2013) Responsiveness of EUL-related rice lectins towards important abiotic and biotic stresses

6. Conferences

With oral presentation

Al Atalah B, Van Damme EJM (2012) Overexpression analysis of some rice lectins. Plant Research FBW Meeting, 4th of May, Ghent, Belgium.

Al Atalah B, Fouquaert E, Van Damme EJM (2013) Expression analysis for EUL-related rice lectins under different stress conditions. The 65th on International Symbosium on Crop Protection, May 21st, Ghent, Belgium

Al Atalah B, Van Damme EJM (2013) Characterization and expression analysis for some rice lectins under abiotic stresses. The 5th European Plant Science Retreat, 23rd -26 of July, Ghent, Belgium

Al Atalah B, Van Damme EJM (2013) *Euonymus* related lectins from rice are stress-related proteins. 24th Joint Glycobiology Meeting, 24th -26th of November, Wittenberg), Germany

With poster presentation

Al Atalah B, Fouquaert E, Van Damme EJM (2010) Expression of the nucleocytoplasmic lectin from rice in *Pichia pastoris*. The 21st Joint Glycobiology Meeting, 7th -9th of November, Ghent, Belgium

Van Hove J, **Al Atalah B**, Fouquaert F, Plas K, Van Damme EJM (2010) Cytoplasmic/nuclear proteins with conserved *Euonymus* lectin-like domains: indications for a stress related role in Embryophyta. Belgian Plant Biotechnology Association, the 3rd of December, Leuven, Belgium

Al Atalah B, Van Damme EJM (2011) Expression of a stress-inducible EUL lectin from rice in *Pichia pastoris*. The 22nd Joint Glycobiology Meeting, November, Lille, France

Al Atalah B, Van Damme EJM (2012) Nucleocytoplasmic lectins from rice are expressed in response to (a)biotic stresses. Pant Abiotic Stress Tolerance II, 22nd -25 of February, Vienna, Austria

Al Atalah B, Plas K, Damme EJM (2013) Expression of nucleocytoplasmic lectins from rice in *Pichia pastoris*. Yeasterday, May 24, Ghent, Belgium