Glycoconj J (2012) 29:467–479
DOI 10.1007/s10719-012-9405-2

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

Bassam Al Atalah · Pierre Rougé · David F. Smith · Paul Proost · Yi Lasanajak · Els J. M. Van Damme

Received: 3 March 2012 / Revised: 18 May 2012 / Accepted: 24 May 2012 / Published online: 10 June 2012

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 amino acid N-terminal domain followed by the EUL sequence. In this paper the original sequence of the EUL domain of OrysaEULS2 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.

Keywords Lectin · Carbohydrate-binding · Mutant · EUL protein · Glycan array

B. Al Atalah · E. J. M. Van Damme (✉)
Laboratory of Biochemistry and Glycobiology,
Department of Molecular Biotechnology, Ghent University,
Coupure links 653,
9000 Ghent, Belgium
e-mail: ElsJM.VanDamme@UGent.be

P. Rougé
Signaux et Messages Cellulaires chez les Végétaux,
UMR CNRS-UPS 5546, Pole de Biotechnologie végétale,
BP 17, 24 Chemin de Borde Rouge,
Castanet-Tolosan 31326, France

D. F. Smith · Y. Lasanajak
Department of Biochemistry,
Emory University School of Medicine,
Atlanta, GA 30322, USA

P. Proost
Laboratory of Molecular Immunology, Rega Institute for Medical Research, Katholieke Universiteit Leuven,
Minderbroedersstraat 10,
3000 Leuven, Belgium

Introduction

Lectins are heterogeneous group of proteins. They have a unique property of binding carbohydrate structures in a specific and reversible way [1]. In the last decade, a new class of inducible lectins was discovered in plants [2]. These lectins are expressed at very low concentrations in response to abiotic or biotic stresses like drought, salinity, hormone treatments and attack of pathogens and predators. Based on these observations, the hypothesis was put forward that the inducible lectins have a possible role in the stress physiology of the plant [3].

In 2008 sequencing of the cDNA clone encoding the *Euonymus europaeus* agglutinin (EEA) from spindle tree led to the discovery of a new lectin motif in plants [4]. Since then, the *Euonymus europaeus* lectin (EUL) domain is considered as the prototype for a new lectin family [5]. 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 [5]. 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) [5].

The OrysaEULS2 protein, originally named as OSR40g3, was first reported in the shoots of rice seedlings subjected to salt stress and abscisic acid treatment [6]. 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 [4] . According to the classification of plant lectins proposed by Fouquaert *et al.* [5] this protein belongs to the group of chimeric EUL proteins and consists of an EUL domain preceded by a 56 unrelated amino acid 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 [7–9]. The expression of the recombinant proteins is controlled by the alcohol oxidase 1 promoter (AOX1) and tightly induced to high levels by methanol [10, 11]. Several nucleocytoplasmic lectins from tobacco (Nictaba), maize (GNAmaize), rice (Orysata) and *Arabidopsis thaliana* (ArathEULS3) have been expressed successfully in *P. pastoris* [12–15].

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.

Materials and methods

Hydrophobic cluster analysis and molecular modeling

Hydrophobic Cluster Analysis (HCA) [16] was performed to assess the conserved secondary structural features (essentially strands of β-sheet) along the amino acid sequences of the *Euonymus europaeus* agglutinin (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 [17] 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) [18], 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 [19] 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 [20] and displayed with GRASP using the parse3 parameters [21]. 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 fixed at 4.0 and 80.0, respectively, and the calculations were performed keeping a salt concentration of 0.145 M.

Expression of EULS2 and its mutant forms in *Pichia 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 1). The mutant forms were created by using reverse primers containing several point mutations (Table 1). All primers used have the EcoRI cleavage site at the 5′ end and the XbaI cleavage site at the 3′ end, which allowed cloning all amplified PCR fragments in the vector pPICZαA after double digestion with EcoRI and XbaI. This shuttle vector contains the α–mating sequence from *Sacharomyces cerevisiae* to enable secretion of the recombinant proteins in the culture medium. The pPICZαA vectors were transformed into *Escherichia coli* Top10F cells by heat shock. After selection on LB agar plates containing 25 μg/ml zeocin (Invitrogen), the plasmid DNA was purified using the E.Z.N.A. Plasmid Mini kit I (Omega Bio-Tek, Georgia, USA) and sequenced using 5′ and 3′ AOX1 specific primers EVD 21 and EVD 22 (Table 1) (carried out by LGC Genomics GmbH, Berlin, Germany). The restriction enzyme SacI (Fermentas, St Leon-Rot, Germany) was used to linearize the pPICZαA vectors by overnight incubation at 37 °C. Approximately 10 μg linearized pPICZαA vector was used to electroporate the *Pichia* strain X-33 using the pulse settings: 25 μF, 1.5 kV and 200 Ω on the GenePulser apparatus (Bio-Rad, Hercules CA, USA). Afterwards, *Pichia* transformants were selected on YPDS plates (1 % yeast extract, 2 % peptone, 2 % dextrose, 1 M sorbitol, 2 % agar) containing 100 μg/ml zeocin.

Expression analysis of the recombinant proteins

Selected *Pichia* colonies were grown overnight on a shaker (200 rpm) at 30 C° in a 50 ml falcon tube containing 5 ml BMGY medium (1 % yeast extract, 2 % peptone, 1.34 % yeast nitrogen base with ammonium sulphate and without amino acids, 4×10^{-5} % biotin, 100 mM potassium phosphate pH 6.0 and 1 % glycerol). After washing the *Pichia* cells with autoclaved distilled water and resuspension in 10 ml BMMY medium (BMGY medium supplemented with 1 % of methanol instead of 1 % of glycerol) the culture was grown for three days. Methanol induction (2 % final concentration) was performed twice a day. Proteins in the culture medium were analyzed after trichloroacetic acid precipitation (10 % final concentration) by SDS-PAGE and western blot analysis. Colonies that showed good expression of EULS2 were grown in 10 ml BMGY medium for 24 h at 30 °C on a rotary shaker at 200 rpm. The next day, the culture was transferred to 50 ml BMGY in 250 ml Erlenmeyer flasks and grown for 24 h, and then *Pichia* cells were washed, resuspended in 200 ml BMMY medium in a 1 l Erlenmeyer flask and grown for three days under the same conditions as mentioned above. Every 24 h 100 % methanol was added to the culture once in the morning and

Table 1 PCR scheme and primer sequences. Nucleotides that have been mutated are underlined in the primer sequence

Construct	PCR/ Template	Direction	Primer name	Primer sequence 5′ to 3′
Original EULS2	PCR1/ pFLCI vector	Forward	EVD 574	GGCGGAGAATTCACCATGGTGT ACTGCCGGGCG AACCCGAA
		Reverse	EVD 514	CCCGCTTTCTAGAATGTAGTAGG GCTGGATCTTCCAGCGC
Mutant 1	PCR2/ pFLCI vector	Forward	EVD 574	Above
		Reverse1	EVD 639	GTAGTAGGGCTGGATCTTCCAGC GCTGGTTGTCGCCCTCGCACA ACTTCCA
W changed to L	PCR3/ PCR2 product	Forward	EVD 574	Above
		Reverse 2	EVD 514	Above
Mutant 2	PCR4/ PCR2 product	Forward	EVD 574	Above
		Reverse1	EVD 640	GTAGTAGGGCTGGATCTTCCAGCG CTGGCTGTCGCCCTC
W changed into L and N into S	PCR5/ PCR4 product	Forward	EVD 574	Above
		Reverse 2	EVD 514	Above
Mutant 3	PCR6/ PCR4 product	Forward	EVD 574	Above
		Reverse	EVD 641	GTAGTAGGGCTGGATCTTCCAGC GCAGGCTGTCGCCCTC
W changed into L, N into S and Q into L	PCR7/ PCR6 product	Forward	EVD 574	Above
		Reverse 2	EVD 514	Above
AOX1 primers		Forward	EVD 21	GACTGGTTCCAATTGACAAGC
		Reverse	EVD 22	GCAAATGGCATTCTGACATCC

once in the evening (2 % final concentration). Finally, the culture was centrifuged at 3000 g for 10 min and the recombinant proteins were precipitated by adding ammonium sulphate (final concentration 80 %) to the supernatant and stored at 4 °C until use.

Purification of the recombinant proteins

Purification of the EULS2 and its mutant forms was achieved by several chromatographic steps. The protein pellet obtained after centrifugation of the ammonium sulphate precipitated *Pichia* culture medium for 25 min at 17000 g was resuspended in 90 ml 20 mM 1,3-diaminopropane and dialyzed overnight. The protein solution was loaded on a Q Fast Flow column (GE Healthcare, Uppsala, Sweden) after equilibrating with 20 mM 1,3 diaminopropane. After washing the column until optical density OD280<0.2 the bound proteins were eluted using 100 mM Tris–HCl pH 8.7 containing 0.5 M NaCl. Subsequently, imidazole (final concentration 25 mM) was added to the pooled fractions and this solution was loaded on a Nickel Sepharose column (GE Healthcare) equilibrated with start buffer (0.1 M Tris–HCl pH 7 containing 0.5 M NaCl and 25 mM imidazole) to select for the His-tagged protein. After washing the Ni-Sepharose column using the start buffer, bound proteins were eluted using the elution buffer (0.1 M Tris–HCl pH 7 containing 0.5 M NaCl) containing a series of imidazole concentrations of 75, 100, 150 and 250 mM. Afterwards, fractions eluted from Ni-Sepharose with 250 mM imidazole were diluted 5 times with 20 mM 1,3-diaminopropane and applied on a small Q Fast Flow column to concentrate the recombinant proteins and adjust them to a suitable buffer for some downstream analyses. The purity of the protein samples was verified by SDS-PAGE and/or Western blot analysis after each purification step. Starting from a 1 L culture 4 to 6 mg of recombinant EULS2 was obtained.

N-terminal sequence analysis

A sample from the affinity purified EULS2 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).

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). In a glass tube, 10 µl of the purified protein (54 µg/ml), 10 µl of 1 M ammonium sulphate and 30 µl of trypsinized erythrocytes were mixed. The negative control contained 20 µl 1 M ammonium sulphate 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. To determine the agglutination titer, a dilution series of the recombinant proteins was analyzed.

Carbohydrate inhibition test

To test the carbohydrate specificity of the recombinant proteins, several carbohydrates (mannose, trehalose, glucose, galactose, GlcNAc or methyl mannopyranoside, at 0.5 M) and glycoproteins (thryoglubulin, ovomucoid and asialomucin at 10 mg/ml) were used. Therefore 10 µl aliquots of a serially twofold diluted purified lectin were mixed with 10 µl of 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 1 h at room temperature.

Glycan array screening

The microarrays are printed as described previously [22] 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 *Pichia 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 s 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.

Analytical methods

The Coomassie (Bradford) Protein Assay Kit (Thermo Fischer Scientific, Rockford, IL USA) was used to measure the protein concentration based on the Bradford dye-binding procedure [23]. SDS-PAGE was performed using 15 % polyacrylamide gels under reducing conditions as described by Laemmli [24]. Proteins were visualized 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 (Biotrace™ PVDF, PALL, Gelman Laboratory, Ann Arbor, MI USA). Blots were blocked for 1 h 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. Afterwards, 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. All washes and incubations were conducted at room temperature with gentle shaking.

Results

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). 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. 1a) and are nicely superposable (Fig. 1b). This β-trefoil structure is reminiscent to that found in ricin-B and other lectins from the ricin-B family [25]. 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.

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) [26], 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 EULS2 from rice) was predicted to occur at the C-terminal end of each of the EUL domains (Fig. 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. 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 ArathEULS3 and OrysaEULS2, respectively, could also participate in stacking interactions that often reinforce the binding of a simple sugar to the carbohydrate-binding site of plant lectins [27]. 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

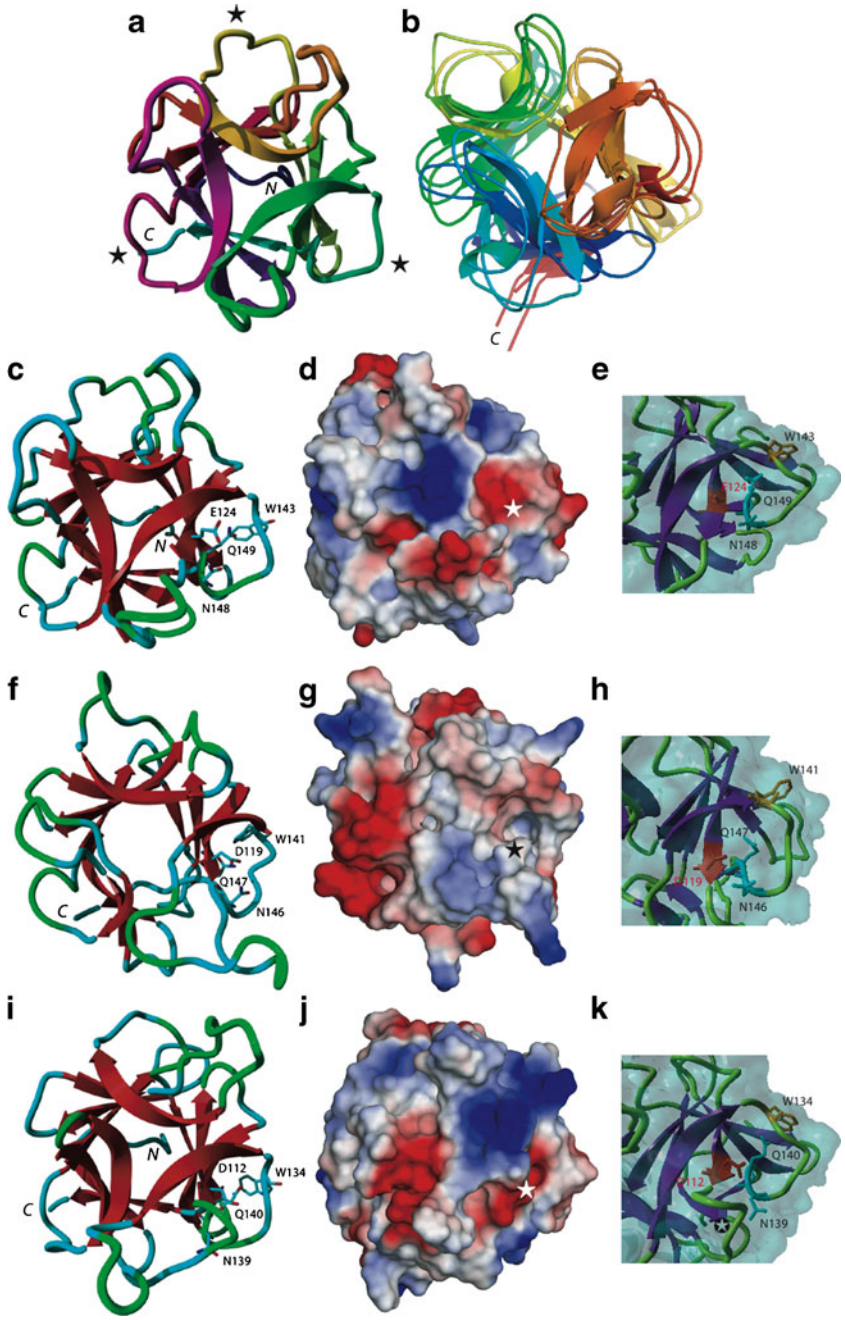

Fig. 1 **a**. Ribbon diagram of EEA showing the three bundles of β-sheet in different colors (orange, green, magenta) associated to the extended loops (indicated by stars) forming the three lobes of the β-trefoil. **b**. Superposition of the three models for EEA and the EUL domains of ArathEUL3 and OrysaEULS2 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 domains 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

buried (Fig. 1k). In this respect, OrysaEULS2 readily differs from EEA and ArathEULS3, since in the latter lectins the corresponding Asn residue remains fully exposed (Fig. 1e,h).

Purification and characterization of the recombinant proteins from *Pichia 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). 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, Gln140 (Gln196 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 *Pichia 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 l 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.

a

M[1]RFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPFSNSTNNG
LLFINTTIASIAAKEEGVSLEKRE[86]A[87]EA[89]EFTM[93]VYCRANPNYAMTARNGAVVLAPANPKD
EYQHWIKDMRWSTSIKDEEGYPAFALVNKATGQAIKHSLGQSHPVRLVPYNPEVMDESVLWTE
SRDVGNGFRCIRMVNNIYLNFDAFHGDKYHGGVRDGTDIVLWKWCEGDNQRWKIQPYYILEQK
LISEEDLNSAVDH[259]HHHHH[264]*

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. 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 bold. Mutated residues are indicated by arrows. The unrelated N-terminal sequences in OrysaEULS2 and ArathEULS3 are shown in red and italics. Residues being the first amino acid of the EUL domain in the three proteins are underlined. *: identical residues and .: similar residues

As shown by SDS-PAGE analysis (Fig. 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 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. 2a), resulting in a polypeptide with a calculated molecular mass of 20.7 kDa.

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 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 EULS2 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. 4. The interaction of the protein with the array was relatively weak compared to other related lectins such as EEA [4] and ArathEULS3 [15]. The data in Fig. 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. 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 [28]. The strongest binding glycans are summarized in Table 2 and indicate that EULS2 exhibits selectivity toward the high mannose N-linked glycans, especially $Man_3GlcNAc_2$ (glycans #51 and 52) with somewhat lower binding to $Man_5GlcNAc_2$ (#217), $Man_6GlcNAc_2$ (#216), $Man_7GlcNAc_2$ (#211), $Man_8GlcNAc_2$ (#212), and $Man_9GlcNAc_2$ (#213). Interestingly, no binding was observed by the lectin to the high mannose structures without the chitobiose core including Man_3 (#214), Man_5 (#215 and 315), Man_8, (#316), and Man_8, (#317) (data not shown in Table 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. 4b–d).

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 [6]. Since inducible lectins are expressed in response to (a)biotic stress factors, it is believed that these lectins are involved in stress signaling [29]. In this paper we describe the expression of the EUL domain from OrysaEULS2 in *Pichia 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.

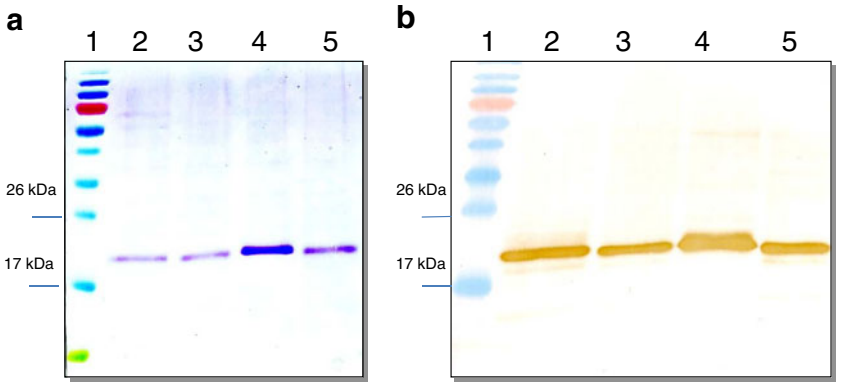

Fig. 3 Purified EUL domain from OrysaEULS2 and its mutant forms were analyzed by SDS-PAGE (**a**) and Western blot analysis with a monoclonal anti-His antibody (**b**). Samples are loaded as follows: lane 1: protein ladder (increasing molecular mass: 10, 17, 26, 34, 43, 55, 72, 95, 130, 170 kDa) (Fermentas, 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)

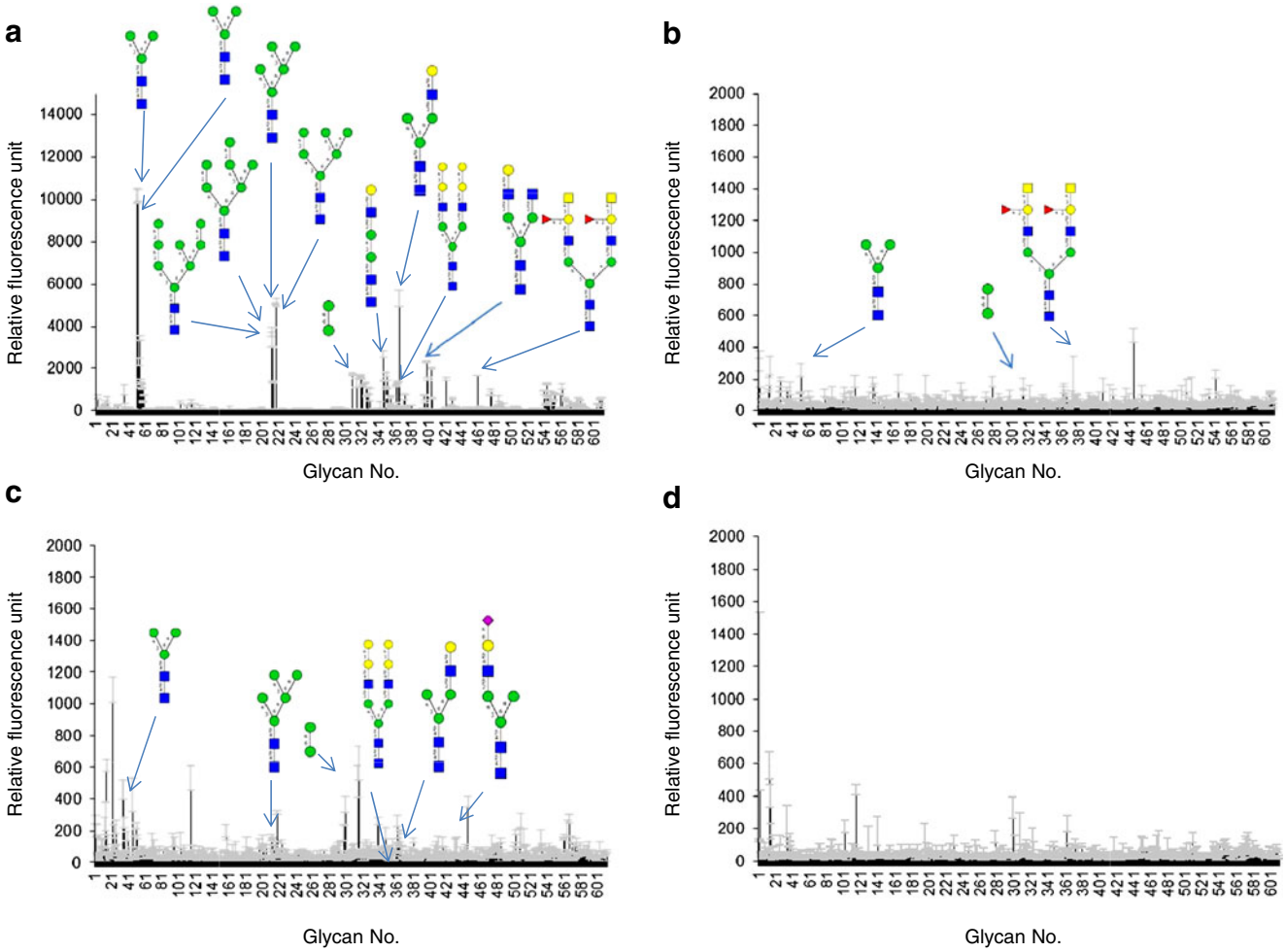

Fig. 4 Comparative analysis of the glycan array binding of the recombinant EUL domain from OrysaEULS2 (**a**) and mutant forms 1–3 (**b**–**d**) tested at 200 μg/ml. The complete primary data set for each protein is available on the website of the Consortium for Functional Glycomics (www.functionalglycomics.org)

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 [12, 14, 30]. 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 [31]. 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 [27]. 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 domains 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,

Table 2 Overview of the top 30 glycans interacting with the EUL domain from OrysaEULS2. Lactosamine structures are underlined. Terminal GlcNAc is shown in italic. The glycan with the highest relative 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	
51	Manα1-6(Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp13	100.0
50	Manα1-6(Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	99.8
217	Manα1-6(Manα1-3)Manα1-6(Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	50.5
216	Manα1-6(Manα1-3)Manα1-6(Manα1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	49.9
211	Manα1-6(Manα1-2Manα1-3)Manα1-6(Manα1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	35.6
212	Manα1-2Manα1-6(Manα1-3)Manα1-6(Manα1-2Manα1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	34.1
213	Manα1-2Manα1-6(Manα1-2Manα1-3)Manα1-6(Manα1-2Manα1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	13.6
314	Manα1-6Manβ-Sp10	13.1
	Lactosamine related structures	
404	Galα1-4Galβ1-3GlcNAcβ1-2Manα1-6(Galα1-4Galβ1-3GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp19	19.7
319	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-6(Neu5Acα2-3Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	16.6
459	Galβ1-4GlcNAcβ1-6(Galβ1-4GlcNAcβ1-2)Manα1-6(Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp19	16.5
320	Galβ1-4GlcNAcβ1-2Manα1-6(Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	15.1
54	Galβ1-4GlcNAcβ1-2Manα1-6(Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	14.2
364	Galα1-3Galβ1-4GlcNAcβ1-2Manα1-6(Galα1-3Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp20	13.6
325	Galβ1-3GlcNAcβ1-2Manα1-6(Galβ1-3GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp19	13.1
543	Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-2Manα1-6(Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp24	12.7
362	Fucα1-2Galβ1-4GlcNAcβ1-2Manα1-6(Fucα1-2Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp20	11.7
55	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-6(Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	11.5
	Mannose /Lactosamine hybrid structures	
365	Galβ1-4GlcNAcβ1-2Manα1-6(Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	52.1
346	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-6(Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAc-Sp12	26.4
398	GlcNAcβ1-2Manα1-6(Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAc-Sp12	22.8
352	Manα1-6(Galb1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	12.3
309	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-6(GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	17.1
350	Galβ1-4GlcNAcβ1-2Manα1-3Manβ1-4GlcNAcβ1-4GlcNAc-Sp12	17.0
399	Galb1-4GlcNAcβ1-2Manα1-6(GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAc-Sp12	14.9
347	Manα1-6(Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAc-Sp12	14.2
321	GlcNAcβ1-2Manα1-6(Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	11.7
	GlcNAc terminated structures	
53	*GlcNAc*β1-2Manα1-6(*GlcNAc*β1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp13	33.7
52	*GlcNAc*β1-2Manα1-6(*GlcNAc*β1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp12	22.6
422	*GlcNAc*β1-2(*GlcNAc*β1-6)Manα1-6(*GlcNAc*β1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp19	15.0

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) [32]. An Asn residue also occurs in the carbohydrate binding sites of the ricin-B domain of *Streptomyces olivaceoviridis* xylanase (PDB code 1ISZ) [33] and the C-terminal domain of the HA33/A protein from *Clostridium botulinum* (PDB code 1YBI) [34], 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 [35]. 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 [14]. The protein 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 lectins 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 [4]. 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 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 [15]. 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 [36], the jacalin-related lectins [37] and the GNA-related lectins [13] has resulted in changes in carbohydrate-binding specificity. Furthermore, evidence shows that the ricin-B domain also displays considerable plasticity in its 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 [38]. 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 [39]. 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 [27]. 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) [40], bacteria and viruses [41–43]. However, in plants only the Lewis A (Galβ1-3(Fucα1-4) GlcNAc) motif has been identified [44–46]. 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.

Acknowledgments This work was funded primarily by the Fund for Scientific Research – Flanders (FWO grants G.0022.08 and KAN 1.5.069.09.N.), the Research Council of Ghent University (projects BOF2005/GOA/008 and BOF2007/GOA/0017). Bassam Al Atalah is recipient of a doctoral grant from the Special Research Council of Ghent University. The authors want to thank the Consortium for Functional Glycomics funded by the NIGMS GM62116 for the glycan array analysis. The authors also want to thank the Rice Genome Resource Center, National Institute of Agrobiological Sciences, Japan for providing the cDNA clone encoding OrysaEULS2.

References

1. Peumans, W.J., Van Damme, E.J.M.: Lectins as plant defense proteins. Plant Physiol. **109**, 347–352 (1995)
2. Van Damme, E.J.M., Lannoo, N., Fouquaert, E., Peumans, W.J.: 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 (2004)
3. Van Damme, E.J.M., Lannoo, N., Peumans, W.J.: Plant lectins. Adv. Bot. Res. **48**, 107–209 (2008)
4. Fouquaert, E., Peumans, W.J., Smith, D.F., Proost, P., Savvides, S., Van Damme, E.J.M.: The old *Euonymus europaeus* agglutinin represents a novel family of ubiquitous plant proteins. Plant Physiol. **147**, 1316–1324 (2008)
5. Fouquaert, E., Peumans, W.J., Vandekerckhove, T.T.M., Ongenaert, M., Van Damme, E.J.M.: Proteins with an Euonymus lectin-like domain are ubiquitous in Embryophyta. BMC Plant Biol. **9**, 136 (2009)
6. Moons, A., Gielen, J., Vandekerckhove, J., Van Der Straeten, D., Gheysen, G., Van Montagu, M.: An abscisic-acid and salt-stress-responsive rice cDNA from a novel plant gene family. Planta **202**, 443–454 (1997)
7. Cereghino, J.L., Cregg, J.M.: Heterologous protein, expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol. Rev. **24**, 45–66 (2000)
8. Macauley-Patrick, S., Fazenda, M.L., McNeil, B., Harvey, L.M.: Heterologous protein production using the *Pichia pastoris* expression system. Yeast **22**, 249–270 (2005)
9. Daly, R., Hearn, M.T.W.: Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. J. Mol. Recognit. **18**, 119–138 (2005)
10. Elias, S.B., Brust, P.F., Koutz, P.J., Waters, A.F., Harpold, M.M., Gingeras, T.R.: Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast *Pichia pastoris*. Mol. Cell. Biol. **5**, 1111–1121 (1985)
11. Hartner, F.S., Glieder, A.: Regulation of methanol utilisation pathway genes in yeasts. Microb. Cell Fact. **5**, 39 (2006)
12. Lannoo, N., Vervecken, W., Proost, P., Rougé, P., Van Damme, E.J.M.: Expression of the nucleocytoplasmic tobacco lectin in the yeast *Pichia pastoris*. Protein Expr. Purif. **53**, 275–282 (2007)
13. Fouquaert, E., Smith, D.F., Peumans, W.J., Proost, P., Balzarini, J., Savvides, S.N., Van Damme, E.J.M.: Related lectins from snowdrop and maize differ in their carbohydrate- binding specificity. Biochem. Biophys. Res. Commun. **380**, 260–265 (2009)
14. Al Atalah, B., Fouquaert, E., Vanderschaeghe, D., Proost, P., Balzarini, J., Smith, D.F., Rougé, P., Lasanajak, Y., Callewaert, N., Van Damme, E.J.M.: Expression analysis of the nucleocytoplasmic lectin "Orysata" from rice in *Pichia pastoris*. FEBS J. **278**, 2064–2079 (2011)
15. Van Hove, J., Fouquaert, E., Smith, D.F., Proost, P., Van Damme, E.J.M.: Lectin activity of the nucleocytoplasmic EUL protein from *Arabidopsis thaliana*. Biochem. Biophys. Res. Commun. **414**, 101–105 (2011)
16. Treiber, N., Reinert, D.J., Carpusca, I., Aktories, K., Schulz, G.E.: Structure and mode of action of a mosquitocidal holotoxin. J. Mol. Biol. **381**, 150–159 (2008)
17. Krieger, E., Koraimann, G., Vriend, G.: Increasing the precision of comparative models with YASARA NOVA - a self-parameterizing force field. Proteins **47**, 393–402 (2002)
18. Gaboriaud, C., Bissery, V., Benchetrit, T., Mornon, J.P.: Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. FEBS Lett. **224**, 149–155 (1987)
19. Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M.: PROCHECK: a program to check the stereochemistry of protein structures. J. Appl. Cryst. **26**, 283–291 (1993)
20. Nicholls, A., Sharp, K.A., Honig, B.: Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins Struc. Func. Genet. **11**, 281–296 (1991)
21. Gilson, M.K., Honig, B.H.: Calculation of electrostatic potential in an enzyme active site. Nature **330**, 84–86 (1987)
22. Blixt, O., Head, S., Mondala, T., Scanlan, C., Huflejt, M.E., Alvarez, R., Bryan, M.C., Fazio, F., Calarese, D., Stevens, J., Skehel, J., van Die, I., Burton, R., Wilson, A., Cummings, R., Bovin, N., Wong, C.-H., Paulson, C.: Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. Proc. Natl. Acad. Sci. U.S.A. **101**, 17033–17038 (2004)
23. Bradford, M.M.: 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 (1976)
24. Laemmli, U.K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**, 680–685 (1970)
25. Van Damme, E.J.M., Hao, Q., Chen, Y., Barre, A., Vandenbussche, F., Desmyter, S., Rougé, P., Peumans, W.J.: Ribosome-inactivating proteins: a family of plant proteins that do more than inactivate ribosomes. Crit. Rev. Plant Sci. **20**, 395–465 (2001)
26. Arndt, J.W., Gu, J., Jaroszewski, L., Schwarzenbacher, R., Hanson, M.A., Lebeda, F.J., Stevens, R.C.: 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 (2005)
27. Van Damme, E.J.M., Rougé, P., Peumans, W.J.: Carbohydrate-protein interactions: plant lectins. In: Kamerling, J.P., Boons, G.J., Lee, Y.C., Suzuki, A., Taniguchi, N., Voragen, A.G.J. (eds.) Comprehensive glycoscience - from chemistry to systems biology, pp. 563–599. Elsevier, New York (2007)
28. Smith, D.F., Song, X., Cummings, R.D.: Use of glycan microarrays to explore specificity of glycan-binding proteins. In: Fukuda, M. (ed.) Meth. Enzymol, vol. 480, pp. 417–444. Academic, Burlington (2010)
29. Van Damme, E.J.M., Barre, A., Rougé, P., Peumans, W.J.: Cytoplasmic/ nuclear plant lectins: a new story. Trends Plant Sci. **9**, 484–489 (2004)
30. Oliveira, C., Felix, W., Moreira, R.A., Teixeira, J.A., Domingues, L.: Expression of frutalin, an α-D-galactose-binding jacalin-related

lectin, in the yeast *Pichia pastoris*. Protein Expr. Purif. **60**, 188–193 (2008)
31. Sreekrishna, K., Brankamp, R.G., Kropp, K.E., Blankenship, D.T., Tsay, J.T., Smith, P.L., Wierschke, J.D., Subramaniam, A., Birkenberger, L.A.: Strategies for the optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. Gene **190**, 55–62 (1997)
32. Rutenber, E., Katzin, B.J., Ernst, S., Collins, E.J., Mlsna, D., Ready, M.P., Robertus, J.D.: Crystallographic refinement of ricin to 2.5-Å. Proteins **10**, 240–250 (1991)
33. Fujimoto, Z., Kuno, A., Kaneko, S., Kobayashi, H., Kusakabe, I., Mizuno, H.: 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 (2002)
34. Arndt, J.W., Gu, J., Jaroszewski, L., Schwarzenbacher, R., Hanson, M.A., Lebeda, F.J., Stevens, R.C.: 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 (2005)
35. Schouppe, D., Rougé, P., Lasanajak, Y., Barre, A., Smith, D.F., Proost, P., Van Damme, E.J.M.: Mutational analysis of the carbohydrate binding activity of the tobacco lectin. Glycoconj. J. **27**, 613–623 (2010)
36. Loris, R., Hamelryck, T., Bouckaert, J., Wyns, L.: Legume lectin structure. Biochim. Biophys. Acta **1383**, 9–36 (1998)
37. Rougé, P., Peumans, W.J., Barre, A., Van Damme, E.J.M.: 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 (2003)
38. Hao, Q., Van Damme, E.J.M., Hause, B., Barre, A., Chen, Y., Rougé, P., Peumans, W.J.: Iris bulbs express type 1 and type 2 ribosome inactivating proteins with unusual properties. Plant Physiol. **125**, 866–876 (2001)
39. Notenboom, V., Boraston, A.B., Williams, S.J., Kilburn, D.G., Rose, D.R.: 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. Biochemistry **41**, 4246–4254 (2002)
40. Stanley, P., Cummings, R.D.: Structures common to different glycans. In: Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W., Etzler, M.E. (eds.) Essentials of glycobiology, 2nd edition, Chapter 13, pp. 175–198. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (2009)
41. Preston, A., Mandrell, R.E., Gibson, B.W., Apicella, M.A.: The lipooligosaccharides of pathogenic gram-negative bacteria. Crit. Rev. Microbiol. **22**, 139–180 (1996)
42. Wang, G., Ge, Z., Rasko, D.A., Taylor, D.E.: Lewis antigens in *Helicobacter pylori*: biosynthesis and phase variation. Mol. Microbiol. **36**, 1187–1196 (2000)
43. Monzavi-Karbassi, B., Luo, P., Cunto-Amesty, G., Jousheghany, F., Pashov, A., Weissman, D., Kieber-Emmons, T.: Fucosylated lactosamines participate in adhesion of HIV-1 envelope glycoprotein to dendritic cells. Arch. Virol. **149**, 75–91 (2004)
44. Leonard, R.: The presence of Lewis A epitopes in *Arabidopsis thaliana* glycoconjugates depends on an active α4-fucosyltransferase gene. Glycobiology **12**, 299–306 (2002)
45. Melo, N.: 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 (1997)
46. Dam, T.K.: 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. Glycobiology **21**, 925–933 (2011)