



HAL
open science

Recombinant N-glycosylation isoforms of Legume lectins: production and purification from *Nicotiana benthamiana* leaves following RuBisCO depletion

Kevin Bellande, Alexandre Lalo, Laetitia Ligat, David Roujol, Elisabeth Jamet,
Hervé Canut

► **To cite this version:**

Kevin Bellande, Alexandre Lalo, Laetitia Ligat, David Roujol, Elisabeth Jamet, et al.. Recombinant N-glycosylation isoforms of Legume lectins: production and purification from *Nicotiana benthamiana* leaves following RuBisCO depletion. *Plant Physiology and Biochemistry*, 2020, 157, <10.1016/j.plaphy.2020.10.038>. <hal-03065006>

HAL Id: hal-03065006

<https://hal.science/hal-03065006v1>

Submitted on 14 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Recombinant *N*-glycosylation isoforms of Legume lectins: production and purification from *Nicotiana benthamiana* leaves following RuBisCO depletion

Kevin Bellande^{1†}, Alexandre Lalo¹, Lætitia Ligat^{1#}, David Roujol¹, Elisabeth Jamet¹ and Hervé Canut^{1*}

¹ Laboratoire de Recherche en Sciences Végétales, Université de Toulouse, CNRS, UPS, Auzeville-Tolosane, France.

[†]present address: UMR Diversité, Adaptation et Développement des Plantes, IRD, Université de Montpellier, 34394 Montpellier Cedex 5, France.

[#]present address: Pôle Technologique du CRCT, Centre de Recherches en Cancérologie de Toulouse, INSERM, UPS, CNRS, Université de Toulouse, Toulouse, France.

*To whom correspondence should be addressed: E-mail: canut@lrsv.ups-tlse.fr

Key words: *Arabidopsis thaliana*; *N*-glycosylation; *Nicotiana benthamiana*; Recombinant lectin; RuBisCO depletion

Abstract

An efficient purification of recombinant proteins often requires a high ratio of recombinant to host proteins. In plants, Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) is the most abundant leaf protein, thus strongly impacting purification yield. Here, we describe a simple and robust purification procedure for recombinant proteins based on a differential precipitation of RuBisCO. In this context, four Legume lectin domains of *Arabidopsis thaliana* which belong to receptor-like kinases and cell wall proteins were produced from *Nicotiana benthamiana* leaves. The recombinant proteins exhibit a unique lectin domain consisting of around 250 amino acid residues with several predicted *N*-glycosylation sites and a six His-tag at the N-terminus. After ammonium sulphate precipitation of total soluble proteins, depletion of RuBisCO was obtained using citrate and succinate buffers during the salting-in step: this depletion was pH-dependent and the presence of di- or tri-carboxylic acids was required. The depleted protein extracts were then subjected to two chromatographic steps which were used in the negative mode to submit a protein fraction enriched as much as possible in recombinant lectin domains to a third chromatographic step (immobilized metal-ion chromatography). Three of the Legume lectin domains were purified near to homogeneity and revealed multiple *N*-glycosylation isoforms, particularly those from receptor-like kinases, which were characterised using specific lectins and deglycosylation enzymes. The production and purification of recombinant lectin domains will facilitate their biochemical characterisation in the context of cell-to-cell signalling and cell wall organisation.

1. Introduction

Legume-type lectins have been identified as carbohydrate binding proteins for a long time (Van Damme et al. 2008). They are ubiquitous proteins in plants and play important roles in cell-to-cell communication, development and defence strategies. The sequencing of plant genomes has revealed that the Legume lectin domain is mainly present in chimeric proteins and associated with intracellular kinase domains to constitute receptor-like kinases (LecRKs) (Bellande et al. 2017). For example, in *Arabidopsis thaliana*, Legume-type LecRKs form a family of 43 genes while only seven genes encode cell wall proteins (Lecs) and four genes receptor-like proteins: the Legume lectin domain is therefore mainly found at the cell surfaces. Many studies suggest specific functions for Legume-type LecRKs in plant innate immunity (Singh and Zimmerli, 2013; Balagué et al. 2017). The functions of Legume-type Lecs have been rarely addressed and their involvement in cell wall organisation has been proposed (Hijazi et al. 2014). The potential ligands for Legume-type LecRKs, *i.e.* eATP, eNAD⁺, eNADP⁺ or/and RGD-containing peptides (Gouget et al. 2006; Choi et al. 2014; Wang et al. 2017; Wang et al. 2019), are far from what could be expected for a carbohydrate-binding domain. Modelling of LecRK lectin domain has confirmed the β -sandwich fold of Legume lectins, the conserved residues involved in Ca²⁺ and Mn²⁺ binding, and the potential carbohydrate binding site (Gouget et al. 2006; Nguyen et al. 2016). In most cases, the carbohydrate-binding activity of Legume-type LecRKs is not demonstrated, except in poplar, where such an activity is detected with α -L-rhamnose as a ligand (André et al. 2005). The *A. thaliana* Legume-type LecRKs are suspected to be devoid of any significant monosaccharide-binding activity. Indeed, the invariant Asp81 key residue responsible for the sugar recognition by canonical Legume lectins is replaced by a His residue (Adar and Sharon 1996; Gouget et al. 2006). However, one cannot exclude that the lectin domain of Legume-type LecRKs

accommodates complex carbohydrates or recruit another lectin domain to restore a functional protein. Finally, *A. thaliana* Legume-type lectin domains are predicted to be post-translationally modified by glycosylation and carry one to several *N*-glycans. In this context, the development of efficient methods to produce and purify recombinant Legume lectin domains in sufficient amount is necessary to decipher their recognition capability.

Nicotianeae is one of the most efficient platforms for producing heterologous recombinant proteins (Yusibov et al. 2016). In particular, agroinfiltration of *Nicotiana benthamiana* leaves together with viral vectors and coexpression of a suppressor of gene silencing (Garabagi et al. 2012) allows accumulating high levels of different heterologous proteins, ranging from reporter fluorescent proteins to complex molecules such as immunoglobulins (Yusibov et al. 2016). Agroinfiltration is an effective and inexpensive technique that can be easily set up in a laboratory. It is also beneficial because plant cells can synthesise complex proteins with post-translational modifications such as glycosylation and disulfide bonds. However, protein recovery and purification processes from plant material, especially from green leaves, remain challenging. Indeed, plants contain numerous molecules including phenolics, proteases and highly abundant proteins that must be neutralised or removed from crude samples. Also, to simplify the downstream processes of purification, recombinant proteins are often translationally fused to small affinity tags.

Protein recovery and purification procedures from leaf extracts are complicated by the high abundance of ribulose biphosphate carboxylase/oxygenase (RuBisCO) which comprises about 30-60% of soluble proteins. Because it is a simple non-chromatographic step, inexpensive and easy to scale up, the selective precipitation of RuBisCO is often used in the early stages of purification processes (Zhang et al. 2005). The objective is to increase the ratio of recombinant to host proteins that is critical to ensure the overall purification efficiency. Accordingly, many attempts have been performed including isoelectric and polyelectrolyte

precipitations (Zhang et al. 2005; Holler and Zhang 2008; Kim et al. 2013; Stephan et al. 2018), polyethyleneglycol (PEG) fractionation (Xi et al. 2006; Widjaja et al. 2009; Arfi et al. 2015), or RuBisCO IgY affinity (Cellar et al. 2008; Widjaja et al. 2009). Most of these procedures have been applied to enrich protein samples in low-abundant proteins for proteomic studies and did not take into account the conservation of protein structures and activities. Isoelectric precipitation requires acidic conditions that may result in protein denaturation. Polyelectrolyte precipitations using polyethyleneimine (PEI), protamine sulphate (PS) or polyacrylic acid (PAA) are quite effective and selective in depleting RuBisCO from the extracts of various plants. However, it is still difficult to remove them from protein samples (Holler and Zhang 2008). The same difficulty occurs with PEG. Finally, RuBisCO IgY affinity efficiently removes RuBisCO from protein samples but with limited amounts of material and frequent contamination with IgY fragments have been observed (Widjaja et al. 2009).

Here we used agroinfiltration in *N. benthamiana* leaves to produce four recombinant Legume lectin domains being part of two LecRKs and two Lecs of *A. thaliana*. These recombinant lectin domains consist of around 250 amino acid residues with high sequence identity to which have been added the V5 epitope and a 6xHis tag at their C-terminus. We have developed a procedure to selectively deplete RuBisCO from total protein extracts. Based on citrate or succinate buffers at mild acidic pH, it was designed to facilitate the downstream purification process which could be achieved for three recombinant lectin domains. Then, their *N*-glycosylation status was addressed.

2. Materials and Methods

2.1. Plant expression constructs

The nucleotide sequences of *LecRK-I.5* (At3g45430), *LecRK-I.9* (At5g60300), *Lec-I.3* (At1g53070) and *Lec-I.6* (At3g16530) were retrieved from TAIR (www.arabidopsis.org). The amino acid residues 1-258 (*LecRK-I.5*), 1-287 (*LecRK-I.9*), 1-272 (*Lec-I.3*) and 1-276 (*Lec-I.6*) covered the signal peptide together with the Lectin_legB domain sequences (PF00139 – www.pfam.xfam.org – version 32.0 September 2018) (Supplementary Fig. S1 and Fig. S2). The four coding sequences (single exon coding sequences) were amplified by PCR from genomic DNA of *A. thaliana* ecotype Col-0 using the *Pfu* DNA polymerase (Promega) and specific primers. The resulting products were cloned into the vector pBAD-TOPO® (Invitrogen) allowing the fusion in-frame to the V5 epitope and the 6xHis tag sequences: these two tags were located at the C-terminus of the lectin domains (Supplementary Fig. S1 and Fig. S2).

Each of the four DNA fragments encoding the signal peptide::Lectin_legB domain::V5 epitope::6xHis tag fusion protein was subcloned by recombination into the entry vector pDONR207 (Invitrogen, CA) and the *pCaMV35S* overexpression vector pK7WG2D.1. A particular feature of the vector pK7WG2D.1 was the presence of the *rolD* promoter::enhanced green-fluorescent protein (*eGFP*) gene carried by its T-DNA.

2.2. Transient expression in *N. benthamiana* leaves

N. benthamiana transient expression was performed by agroinfiltration of leaves using a needle-less syringe. The four binary vectors were introduced into *Agrobacterium tumefaciens* strain LBA4404 using a freeze/thaw transformation method and the cultures were maintained in LB medium supplemented with 100 $\mu\text{g}\cdot\text{mL}^{-1}$ spectinomycin and 25 $\mu\text{g}\cdot\text{mL}^{-1}$ rifampicin. For infiltration, bacteria were grown to stable phase at 28°C to an OD₆₀₀ of 1.0 and collected by centrifugation. The bacterial pellets were resuspended in 10 mM MgCl₂ (pH 5.8) containing 100 μM acetosyringone by keeping a final OD₆₀₀ value of 0.5. The transformed *A. tumefaciens* suspensions were infiltrated into the leaves of 6 week-old *N.*

benthamiana plants (from the third to the fifth leaf from the apical meristem) after mixing with an equal volume of an *A. tumefaciens* suspension (final OD₆₀₀ value of 0.5) carrying a p19 suppressor of post-transcriptional gene silencing. The overall surface of the leaves was infiltrated. The plants were maintained in a growth chamber under 16 h light (24°C)/8 h dark (20°C) photoperiod, at 120 μmol.m⁻².s⁻¹ and 75% relative humidity. Infiltrated tissues were harvested three days post-infection and the material was frozen in liquid nitrogen, processed into fine powder and stored at -80°C until use.

2.3. Floral dip transformation of *A. thaliana*

A. thaliana Col 0 ecotype transformation was performed by floral dipping. The previous binary vector containing the DNA fragment encoding the lectin domain of LecRK-I.9 described above was introduced into the *A. tumefaciens* strain C58 containing the pMP90 (GV3101) plasmid using a freeze/thaw transformation method. The culture was maintained in LB medium supplemented with 100 μg.mL⁻¹ gentamycin and 25 μg.mL⁻¹ rifampicin. One colony of the *A. tumefaciens* strain was grown overnight in LB medium. Cells were harvested by centrifugation and resuspended into the infiltration medium (see below) to a final OD₆₀₀ of approximately 1.0. Inoculations were performed by dipping aerial parts of the plants for a few seconds in 500 mL of a solution containing 5% (w/v) sucrose, 0.03% (v/v) Silwett L-77, 10 mM MgCl₂ and the *A. tumefaciens* cells. The plants were kept into the dark with high humidity for 24 h and then transferred to the growth chamber. Seeds were sown on medium supplemented with 100 μg.mL⁻¹ kanamycin to select transformants.

2.4. Extraction and fractionation of total soluble proteins from *N. benthamiana* leaves

Ground leaf material (1 g) was homogenised with 5 mL ice-cold protein extraction buffer consisting of 20 mM NaH₂PO₄/Na₂HPO₄ buffer, pH 7.4, 150 mM NaCl, 10 mM β-mercaptoethanol, 100 mg polyvinylpyrrolidone. The buffer also contained 100 μL of a protease inhibitor cocktail (P8849, Sigma-Aldrich), along with 100 μM L-benzylsuccinic acid

as a carboxypeptidase inhibitor. After 15 min vortexing vigorously at 4°C, the slurry was clarified by centrifugation at 40,000 x g for 10 min at 4°C. The resulting pellet was resuspended with 5 mL ice-cold protein extraction buffer for a second extraction following the same procedure. The two supernatants were combined and filtered through glass wool. The final filtrate was referred to as total soluble proteins (TSP). The protein fractionation started with protein precipitation using either ammonium sulphate, protamine sulphate (PS), polyethyleneimine (PEI) or polyethylene glycol (PEG). The four different procedures were compared to evaluate their efficiency with regard to the purification process of the lectin domains.

Solid $(\text{NH}_4)_2\text{SO}_4$ was added to the TSP to a final concentration of 80 % and allowed to fully dissolve for 30 min at 4°C under continuous vortexing. Following centrifugation at 40,000 x g for 10 min, the resulting 80 % pellet was resuspended in an ice-cold 10 mM citrate solution containing 150 mM NaCl buffered to pH 5.0 (NaOH), unless otherwise indicated. The solubilisation step was carried out at a ratio of 1.5 mL citrate solution to a pellet corresponding to 1 g of ground leaf material. The mixture was kept on ice for 10 min to ensure protein solubilisation and centrifuged at 12,000 x g for 5 min at 4°C. The supernatant was collected and the precipitate was solubilised in 20 mM $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ buffer, pH 7.4, 150 mM NaCl. Alternatively, the precipitate was solubilised in 10 mM citrate or succinate buffer at different pH ranging from 3.0 to 6.0, 150 mM NaCl.

For the three other procedures, TSP was divided into equal portions and mixed with (i) the desired PS concentration (1% stock solution in distilled water), from 0.02% to 0.20% (PS salt from salmon, P4020, Sigma-Aldrich), (ii) the desired PEI concentration (50% aqueous stock solution), from 0.1% to 0.5% (PEI long chain average molecular mass 750 kDa, 181978, Sigma-Aldrich), or (iii) 5, 10, 15, and 20% w/v polyethylene glycol (PEG) (PEG approx. molecular mass 3350, P3640, Sigma-Aldrich). The mixtures were kept on ice for 10

min to ensure protein precipitation and centrifuged at 12,000 x g for 5 min at 4°C. The supernatants were collected and the precipitates were solubilised in 20 mM NaH₂PO₄/Na₂HPO₄ buffer, pH 7.4, 150 mM NaCl.

All the protein fractions were stored at -20°C before protein amount determination, SDS-PAGE and immune- or lectin-blot analysis. Protein concentration was determined by the Bradford dye binding assay (Pierce™).

2.5. Extraction and fractionation of total soluble proteins from *A. thaliana* and *Brachypodium distachyon* leaves

To know whether the procedures described above could be applied to other plants, TSP from *A. thaliana* and *B. distachyon* leaves were prepared as for *N. benthamiana*. Young leaves from four week-old rosettes of *A. thaliana* and young leaves (3 to 6 cm-long) from two-month-old plants of *B. distachyon* were collected.

2.6. Purification of recombinant proteins

Recombinant lectin domains were isolated according to the same purification scheme involving three consecutive chromatographic steps. The first two steps (anion-exchange and hydroxyapatite chromatographies) were used in the negative mode in order to submit a protein fraction enriched as much as possible in lectin domains to the third chromatographic step (immobilized metal-ion chromatography) that relies on the presence of a 6xHis tag at the C-terminus of the proteins. All the purification processes started after protein depletion of TSP using the citrate procedure (see above).

The depleted protein fractions were desalted using Econo-Pac® 10DG columns (Bio-Rad) equilibrated with 20 mM Tris pH 7.5 (HCl), 50 mM NaCl. In the case of the protein fractions containing the lectin domain of LecRK-I.5, the equilibration buffer was 20 mM Tris pH 7.5 (HCl), 100 mM NaCl. DEAE Sepharose FF (GE Healthcare) equilibrated in their respective buffers was added to the desalted fractions in the following ratio: 1 mL of

chromatographic support to a protein fraction obtained from 1 g of ground leaf material. The mixture was then introduced into a column and the flow-through fraction was desalted as before with columns equilibrated with 10 mM NaH₂PO₄/Na₂HPO₄ buffer, pH 6.5, 500 mM NaCl. Hydroxyapatite (Bio-Gel HTP, Bio-Rad) equilibrated in the same buffer was added to the new desalted fractions (500 µL of chromatographic support /1 g of ground leaf material). The mixture was loaded onto a column and the flow-through fraction was adjusted to pH 8.5 (NaOH 1 M). Ni-NTA Superflow (Qiagen) equilibrated with 50 mM NaH₂PO₄/Na₂HPO₄ buffer, pH 8.5, 500 mM NaCl was added to the hydroxyapatite flow-through fraction (500 µL of chromatographic support /1 g of ground leaf material) and a protein fraction was eluted in the same buffer plus 20 mM imidazole. All steps were carried out at 4°C. All protein fractions were stored at -20°C before protein amount determination, SDS-PAGE, immuno- or lectin-blot analysis, and enzymatic deglycosylation. During the purification steps, protein concentration was determined by the Bradford dye binding assay.

2.7. SDS-PAGE, immunoblotting and lectin blot analysis

Electrophoretic separation of the proteins was performed by 11% (w/v) SDS-PAGE in reducing conditions. The resolved proteins were stained with Coomassie colloidal blue to visualise protein patterns, or electrotransferred onto nitrocellulose membranes for immunodetection with mouse anti-V5 epitope antibody (Invitrogen) and alkaline phosphatase-conjugated goat anti-mouse IgG as secondary antibody (Promega) followed by a staining reaction using nitroblue tetrazolium (NBT) and 5-bromo-4-chloro-3-indoyl phosphate (BCIP).

Presence of *N*-glycans on purified lectin fractions was analysed by reducing SDS-PAGE and lectin-blotting, followed by detection with DIG-labelled lectins according to the instructions supplied by the manufacturer (DIG Glycan differentiation Kit, Roche Diagnostics). Two DIG conjugated lectins were used in this analysis: *Datura stramonium* agglutinin (DSA) recognises unsubstituted galactose-(1-4)-*N*-acetylglucosamine; *Galanthus*

nivalis agglutinin (GNA) recognises terminal mannose, (1-3), (1-6) or (1-2) linked to mannose. Lectin binding was detected by alkaline phosphatase-conjugated anti-digoxigenin followed by a staining reaction using NBT/BCIP.

2.8. Enzymatic deglycosylation

N-glycans were released from proteins with Endoglycosidase H (EndoH) and Peptide *N*-Glycosidase F (PNGaseF) according to the instructions supplied by the manufacturer (EndoH, V4871; PNGaseF, V483A; Promega). The purified protein samples (0.5 – 2 µg) were first denaturated for 5 min at 95°C by adding 100 mM DTT and 0.5% (w/v) SDS final concentrations. EndoH digestions were carried out at 37°C for 24 h in 50 mM sodium citrate buffer (pH 5.5) containing 500 units EndoH and 1 µL of a protease inhibitor cocktail (P9599, Sigma-Aldrich). PNGaseF digestions were carried out at 37°C for 3 h in 50 mM sodium phosphate buffer (pH 7.5) containing 20 units PNGaseF, 2% (v/v) TritonX-100 and 1 µL of a protease inhibitor cocktail (P9599, Sigma-Aldrich). The final volume of the reaction mixtures was 25 µL. The enzymatic reactions were stopped by adding Laemmli buffer and immediately stored at -20°C until SDS-PAGE and lectin-blot analysis.

3. Results

3.1. Multiple isoforms of the recombinant lectin domains are produced in *N. benthamiana* leaves

Recombinant lectin domains originating from either lectin receptor kinases (LecRK-I.5 and LecRK-I.9) or cell wall lectins (Lec-I.3 and Lec-I.6) were produced in transiently transformed *N. benthamiana* leaves. The four constructs contained the coding sequences of lectin domains including their original signal peptides targeting the recombinant proteins to the secretory pathway as well as sequences encoding V5 epitope and 6xHis tags. Total soluble proteins (TSP) were analysed by SDS-PAGE and immunoblotting (Fig. 1). Recombinant lectin

domains exhibited significant molecular mass heterogeneity with multiple closely spaced bands separated from each other by about 2 kDa. The apparent molecular masses were between 40 and 30 kDa for LecRK-I.5 (4 bands), 55 and 35 kDa for LecRK-I.9 (6 bands), 40 and 35 kDa for Lec-I.3 (1 major plus 2 minor bands) and around 35 kDa for Lec-I.6 (1 thick plus 2 minor bands). The expected molecular masses were lower than those observed: 28.7, 32.4, 30.9 and 31.5 kDa for LecRK-I.5, LecRK-I.9, Lec-I.3 and Lec-I.6 respectively (Supplementary Fig. S1 and Fig. S2). This result suggests that the observed bands correspond to *N*-glycosylated forms of the lectin domains. Indeed, bioinformatics predictions indicate the presence of 4, 7, 3 and 3 *N*-glycosylation sites for LecRK-I.5, LecRK-I.9, Lec-I.3 and Lec-I.6 respectively (Supplementary Fig. S1 and Fig. S2). The Legume lectin of LecRKs thus appeared to be highly glycosylated with most of the *N*-glycosylation sites occupied by carbohydrates while cell wall lectins showed one prevailing glycoform.

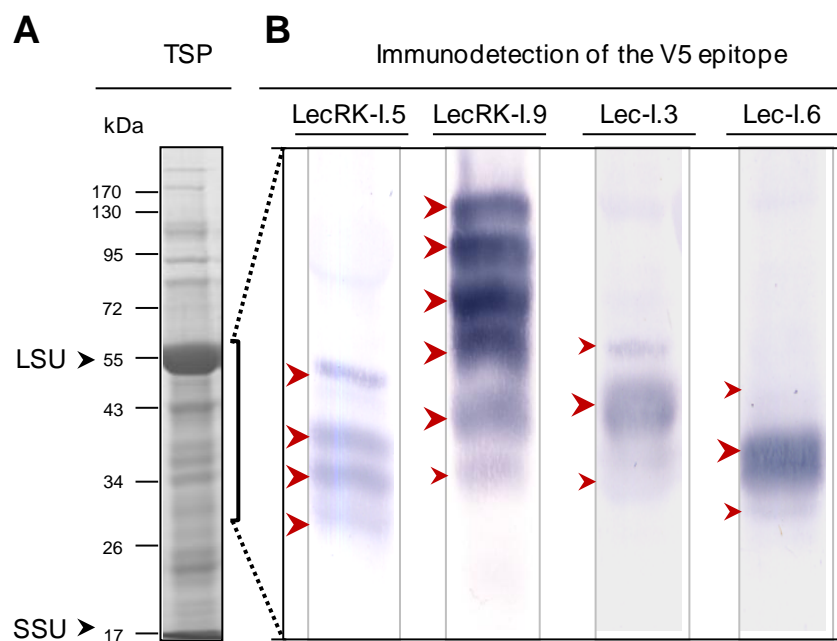


Fig. 1. Immunoblotting analysis of recombinant lectin domains from total protein extracts of *N. benthamiana* leaves after transient expression. (A) SDS-PAGE analysis of a total soluble protein extract (TSP) from leaves transformed with the LecRK-I.9::V5::6xHis construct as an example, followed by staining with Coomassie colloidal blue: the abundant RuBisCO large (LSU) and small (SSU) subunits are marked by black arrowheads; (B) Immunoblots to reveal the recombinant lectin domains (red arrowheads) of LecRK-I.5, LecRK-I.9, Lec-I.3 and Lec-I.6 using anti-V5 antibody.

3.2. Citrate buffers deplete RuBisCO from leaf protein extracts during a salting-in step

In order to clarify protein extracts by removing residual particulates while maintaining a high extraction efficiency of the recombinant lectins, we have developed a method using citrate buffers during a salting-in step, the flow chart of which is presented in Fig. 2. Once total soluble proteins (TSP) were precipitated by ammonium sulphate to 80% saturation (salting-out step), citrate buffers at different pH ranging from 3.0 to 6.0 and containing 0.15 M NaCl were added in order to solubilise proteins (salting-in step) from pellets. After centrifugation, proteins from supernatants and pellets were analysed by SDS-PAGE and immunoblotting (Fig. 3). Protein patterns revealed that RuBisCO was fully depleted from the supernatants obtained with citrate buffer at pH 5.0 or below (Fig. 3A) while the recombinant lectin domains of LecRK-I.9 and Lec-I.6 remained present (Fig. 3B and C). Similar patterns were obtained for the lectin domains of LecRK-I.5 and Lec-I.3 (not shown). Notably, all the isoforms of each lectin domain present in TSP were also detected in depleted protein extracts.

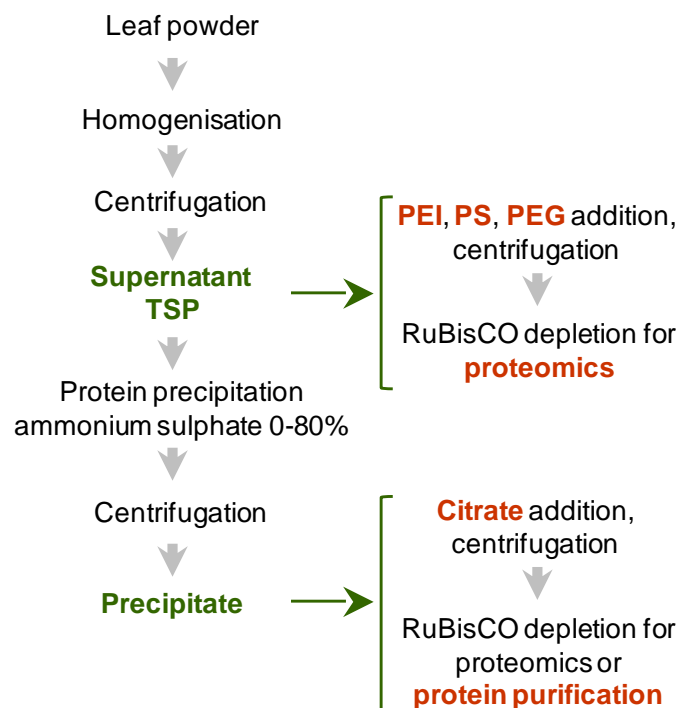


Fig. 2. Flow chart describing the procedures for the depletion of RuBisCO from plant leaf protein extracts. TSP: total soluble proteins. PEI: polyethyleneimine. PS: protamine sulphate. PEG: polyethylene glycol.

Consistently with the protein patterns observed for supernatants, those of pellets showed huge amount of RuBisCO and only traces of the recombinant lectin domains (Fig. 3, compare panels A and B-C). However, the protein depletion was a non-specific process since

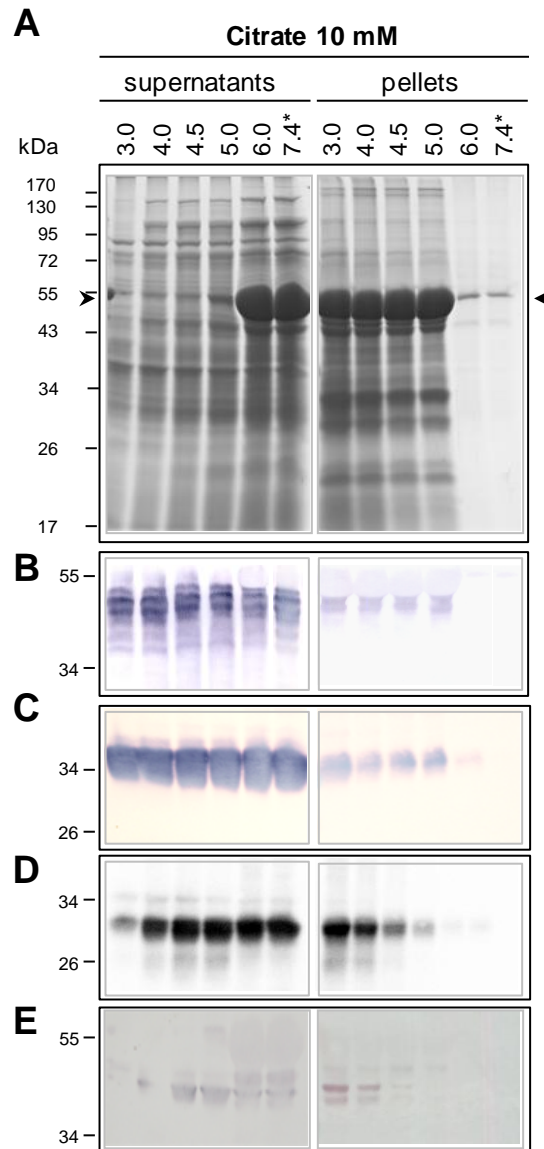


Fig. 3. Effect of citrate buffers on RuBisCO precipitation, recombinant lectin domains recovery, GFP and actin recovery. (A) After ammonium sulphate precipitation, protein precipitates were re-suspended in citrate buffers of different pH ranging from 3.0 to 6.0, or in phosphate buffer pH 7.4 as a reference (*). After centrifugation, proteins from supernatants and pellets were analysed by SDS-PAGE and stained by Coomassie colloidal blue. Total soluble protein extracts from leaves transformed with the LecRK-I.9::V5::6xHis construct are shown as an example. Abundant RuBisCO large subunit (LSU) is marked by black arrowheads. (B) Immunoblots corresponding to (A) were revealed to detect recombinant lectin domains of LecRK-I.9 with anti-V5 antibody. (C) Immunoblots corresponding to (A) were revealed to detect recombinant lectin domains of Lec-I.6 with anti-V5 antibody. (D) Immunoblots corresponding to (A) were revealed to detect a recombinant GFP (Green Fluorescent Protein) with anti-GFP antibody. (E) Immunoblots corresponding to (A) were revealed to detect native actin with anti-actin antibody.

many proteins other than RuBisCO were present in the pellets. Besides, two other proteins, namely a recombinant GFP (Green Fluorescent Protein) and native actin, were assessed and showed a distinct behaviour during the salting-in step (Fig. 3D and E): while soluble in citrate buffers at pH 4.5 to 6.0, GFP partly precipitated while actin completely precipitated at more acidic pH (3.0 and 4.0). From these results, it appears that the citrate buffer at pH 5.0 is the most effective in depleting RuBisCO and enriching the abundance of other proteins within the supernatants with little or no effect on their depletion. In fact, any citrate buffer below pH 5.0 might be selected if we consider only the selective extraction of lectin domains. However, a citrate buffer at pH 5.0 was chosen because the potential deleterious effects of a more acidic pH on protein structures and activities. In addition, this step significantly contributed in clarifying the protein samples as judged by the chlorophyll pigments present in the pellets.

3.3. RuBisCO depletion by citrate buffers is beyond a simple acidic precipitation

The selective extraction observed during the salting-in step apparently relied on the differential behaviour of proteins towards the pH of the added solutions. Citrate buffers depleted RuBisCO in the acidic range starting at pH 5.0 while lectin domains remained soluble even under highly acidic conditions (pH 3.0 – Fig. 3B and C). Citrate buffers were ineffective in precipitating RuBisCO at pH 6.0. To further examine this step, succinate and MES buffers at different pH ranging from 3.0 to 6.0 and containing 0.15 M NaCl were tested in a procedure identical to that used for the citrate buffers (Fig. 2). The resulting protein supernatants and pellets were analysed by SDS-PAGE and immunoblotting (Fig. 4). The succinate buffers precipitated RuBisCO at pH 5.0 or below (Fig. 4A) while the recombinant lectin domains of LecRK-I.9 remained present (Fig. 4B): the overall protein profiles were identical to those obtained with citrate buffers. In contrast, MES buffers failed to precipitate RuBisCO at pH 5.0 (Fig. 4C): RuBisCO was only depleted from supernatant of pH 3.0 and recovered as a precipitate in the corresponding pellet. MES buffers let lectin domains mainly

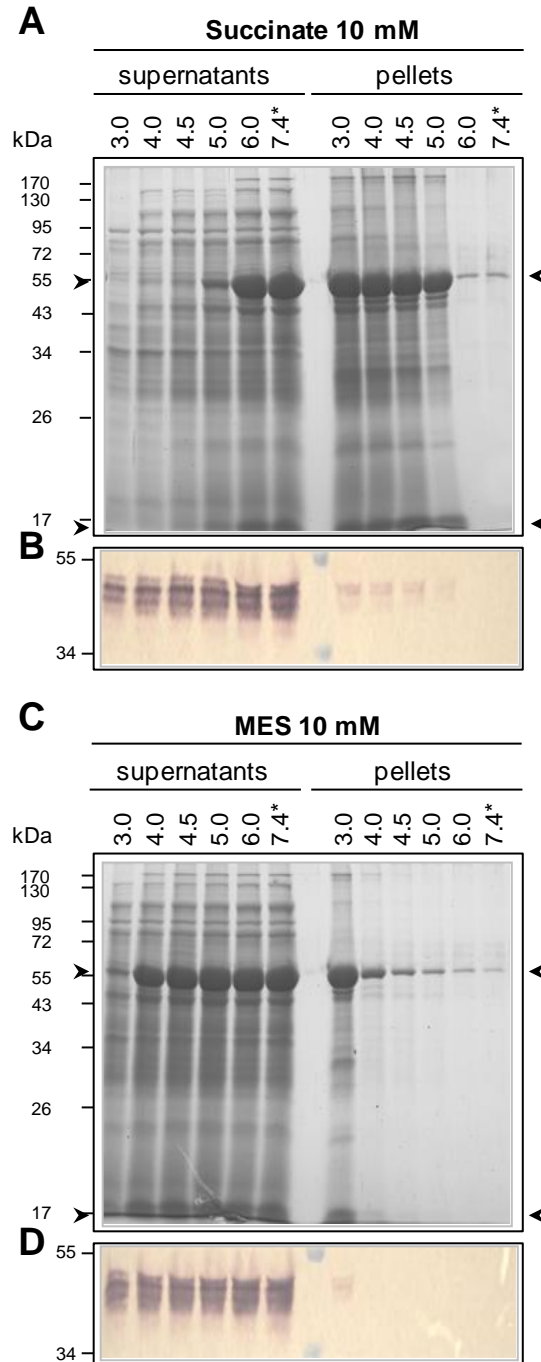


Fig. 4. Effect of succinate and MES buffers on RuBisCO precipitation and recombinant lectin domain recovery. Total soluble proteins of *N. benthamiana* leaves after transient transformation for the production of recombinant lectin domains of LecRK-I.9: after ammonium sulphate precipitation, protein precipitates were re-suspended in succinate (A) and MES buffers (C) of different pH ranging from 3.0 to 6.0, or in phosphate buffer pH 7.4 as a reference (*). After centrifugation, proteins from supernatants and pellets were analysed by SDS-PAGE and stained with Coomassie colloidal blue. Abundant RuBisCO large (LSU) and small (SSU) subunits are marked by black arrowheads. Immunoblots (B) and (D) corresponding to (A) and (C) were revealed with anti-V5 antibody to detect recombinant lectin domains of LecRK-I.9. Details are given in the experimental section.

soluble whatever the pH (Fig. 4D). These results suggest an acidic precipitation of RuBisCO under highly acidic conditions (pH 3.0) whatever the buffer used, and a distinctive effect of citrate and succinate in precipitating RuBisCO in mild acidic conditions.

RuBisCO was equally depleted from supernatants and precipitated in the pellets obtained from *A. thaliana* and *B. distachyon* leaves using citrate buffers during the salting-in step (Fig. 5). The most effective citrate buffer appeared to be at pH 4.5: the protein depletion

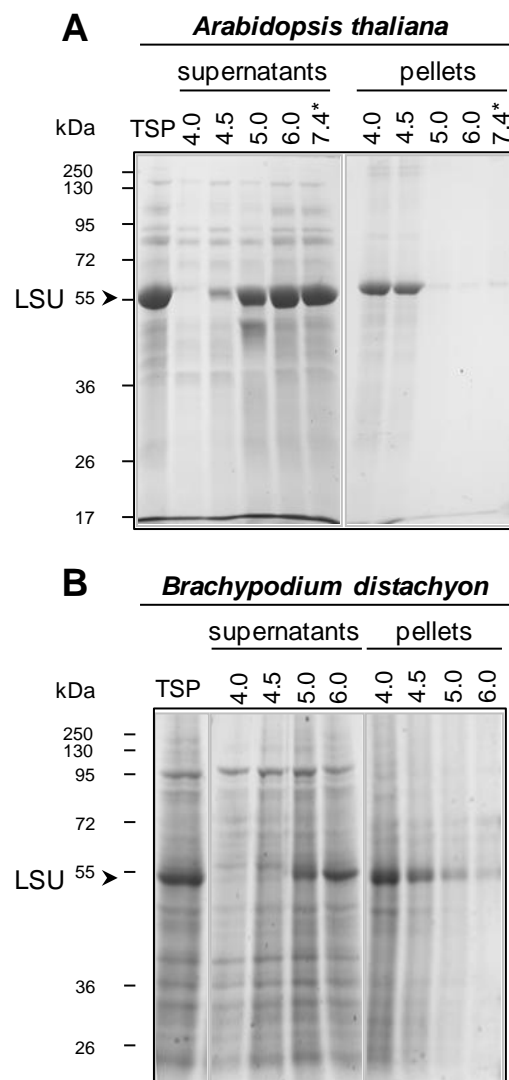


Fig. 5. Effect of citrate buffers on RuBisCO precipitation in leaf extracts of *A. thaliana* and *B. distachyon*. After ammonium sulphate precipitation, protein precipitates were re-suspended in citrate buffers of different pH ranging from 4.0 to 6.0, or in phosphate buffer pH 7.4 as reference (*). After centrifugation, proteins from supernatants and pellets were analysed by SDS-PAGE and stained with Coomassie colloidal blue. TSP, total soluble proteins. LSU, RuBisCO large subunit.

was still a non-specific process with other proteins being precipitated together with RuBisCO. The procedure thus seems reliable for other plants with a condition of pH adjustment.

3.4. Comparison of the procedure using citrate with those using polyelectrolyte precipitation and PEG fractionation

Several methods to enrich plant extracts in recombinant proteins by depleting RuBisCO have been previously established. The most common methods are based on either polyethyleneglycol (PEG) fractionation or polyelectrolyte precipitation using polyethyleneimine (PEI) and protamine sulphate (PS). We therefore applied them for comparison with the citrate procedure following the flow chart presented in Fig. 2.

PEI precipitation was conducted on TSP from *N. benthamiana* leaves obtained in phosphate buffer with a protein concentration of 1 mg.mL⁻¹. A nearly complete precipitation of RuBisCO and a maximum recovery in the pellet fractions were observed at concentrations of 0.1% and 0.2% (Fig. 6A). Increasing the PEI concentration actually decreased the amount of precipitated RuBisCO. Similarly to the citrate procedure, the process was non-specific since many other proteins were recovered in the pellets. All the isoforms of the lectin domain of LecRK-I.9 were found in equal amounts in the different supernatants (Fig. 6B). However, they were also present in the pellets thus leading to the loss of the protein of interest.

Using PS, a very efficient RuBisCO precipitation started at the lowest concentration of 0.02% (Fig. 6C). Although still efficient, higher concentrations of PS left some RuBisCO in the supernatants in the same way as PEI: a protein-polyelectrolyte complex with a stoichiometry below 1 is usually insoluble, while overdosed system might be soluble (Romanini et al. 2013). Once again, protein removal using PS was non-specific. All the isoforms of the lectin domain of LecRK-I.9 were recovered at high yield in all supernatants in a way similar to the citrate procedure (Fig. 6D).

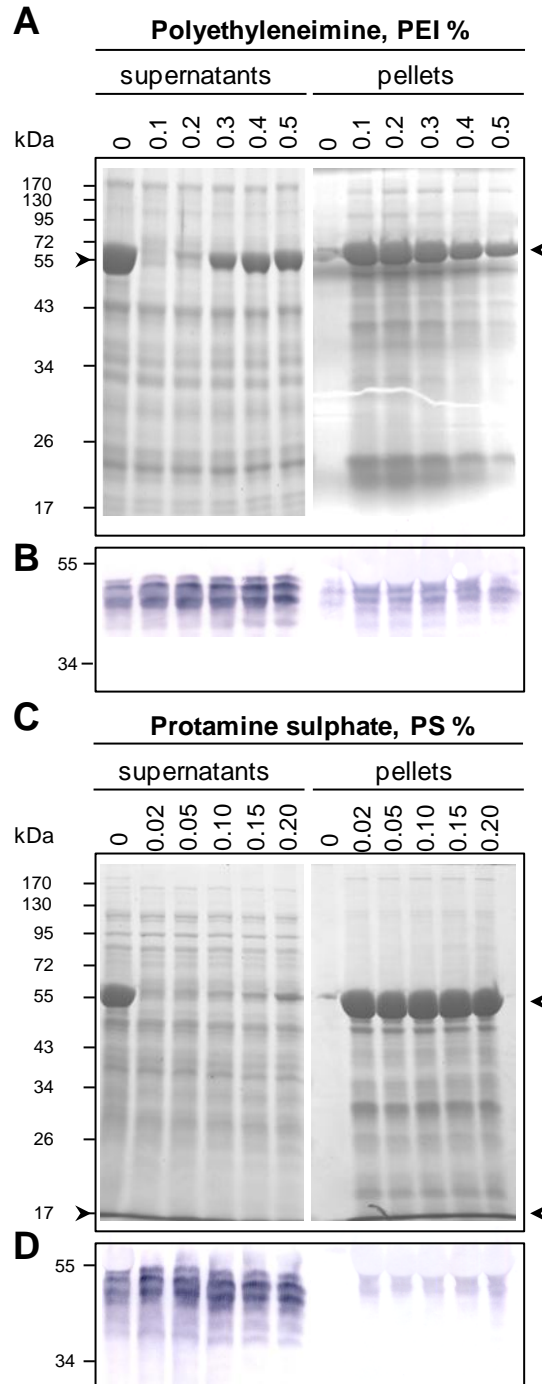


Fig. 6. Effect of polyethyleneimine (PEI) and protamine sulphate (PS) on RuBisCO precipitation and recombinant lectin domain recovery. (A) Total soluble extracts of *N. benthamiana* leaves after transient transformation for the production of recombinant lectin domains of lecRK-I.9 were subjected to different concentrations of PEI ranging from 0.1% to 0.5%. After centrifugation, proteins from supernatants and pellets were analysed by SDS-PAGE and stained by Coomassie colloidal blue. Abundant RuBisCO large (LSU) and small (SSU) subunits are marked by black arrowheads. (B) Immunoblots corresponding to (A) were revealed to detect recombinant lectin domains of LecRK-I.9 with anti-V5 antibody. (C) Total soluble extracts were subjected to different concentrations of PS ranging from 0.02% to 0.2%. (D) Immunoblots corresponding to (C) were revealed with anti-V5 antibody to detect recombinant lectin domains of LecRK-I.9.

PEG precipitated RuBisCO in a dose-dependent manner (Supplementary Fig. S3A). However, as PEG concentration increased, numerous proteins pulled down together with RuBisCO in the pellets, and particularly all the isoforms of the lectin domain of LecRK-I.9 (Supplementary Fig. S3B).

Protein determination (Fig. 7) confirmed the protein profiles and indicated that 60 to 70% of TSP were removed using succinate and citrate buffers. Such a protein depletion was comparable to that obtained with the most efficient concentrations of PEI and PS.

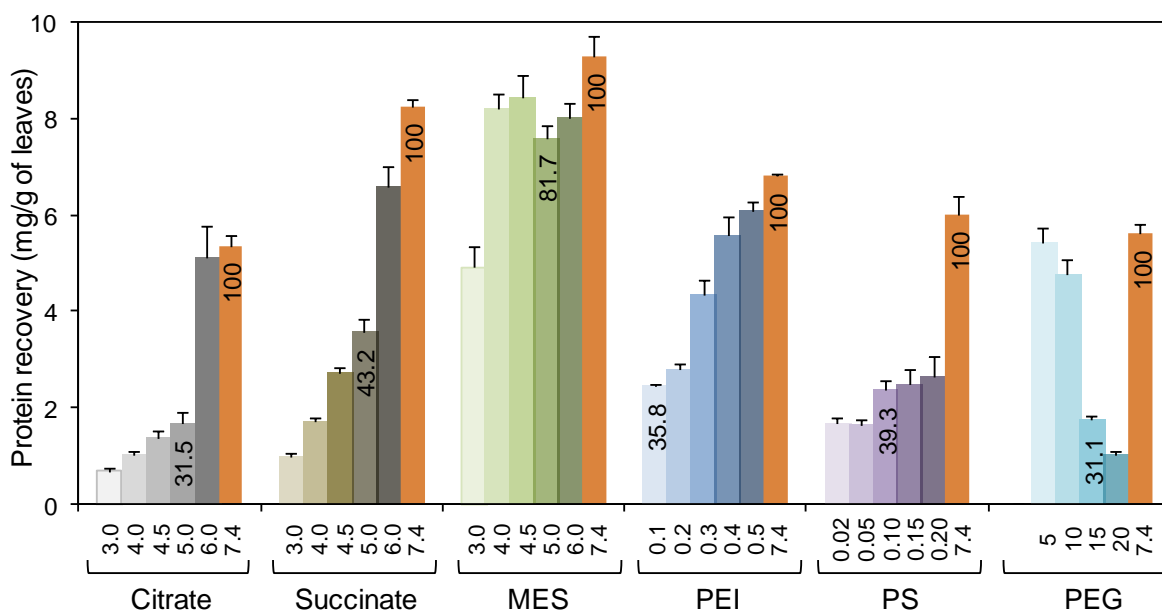


Fig. 7. Protein depletion from *N. benthamiana* extracts. Several methods were tested for protein removal: 10 mM citrate, succinate or MES solutions at different pHs; polyethyleneimine (PEI), protamine sulphate (PS) or polyethyleneglycol (PEG) solutions in phosphate buffer (pH 7.4) ranging from concentrations of 0.1% to 0.5%, 0.02% to 0.2% and 5% to 20% respectively. Protein contents were determined on supernatants obtained after treatments. Extracts obtained in phosphate buffer pH 7.4 without additives (orange columns) were used as reference (100) to indicate yields in selected columns. Errors bars indicate standard deviation (n=4).

3.5. Purified recombinant lectin domains from *N. benthamiana* are N-glycosylated

Purification of the recombinant lectin domains started with 2 g of agro-infiltrated *N. benthamiana* leaves. After the selective extraction of the lectin domains, using citrate buffer pH 5.0 during a salting-in step, three consecutive chromatographic steps were performed for their purification (see Materials and Methods). In the course of establishing the purification

process, cation exchange and hydrophobic interaction chromatographies were excluded as well as Concanavalin A (ConA) affinity chromatography. In the former cases, the recombinant lectin domains did not bind efficiently to the columns. In the latter case, leaking of ConA, a member of the Legume lectin family, heavily contaminated the eluted fractions. This was a major drawback for the biochemical characterisation of the recombinant Legume lectin domains. Of the four lectin domains to be purified, three of them bound to the Ni-NTA support consistently with the presence of an efficient 6xHis tag at the C-terminus of the proteins (LecRK-I.9, LecRK-I.5, Lec-I.3). However, the recombinant lectin domain of Lec-I.6 did not bind to the Ni-NTA column. The eluted fractions were analysed by SDS-PAGE.

After immunoblotting using an antibody against the V5 epitope, the purified recombinant lectin domains appeared as groups of bands (Fig. 8A) with patterns very similar to those obtained from TSP (Fig. 1) indicating an analogous behaviour of the different isoforms during the purification process. After Coomassie colloidal blue staining, the purified lectin domains appeared as the major bands of the eluted fractions with no visible or only a few protein contaminants. Only the fainter bands detected with the anti-V5 antibody were not stained by Coomassie colloidal blue (Fig. 8A). The recombinant lectin content in the eluted fraction related to LecRK-I.5 amounted to $14 \mu\text{g}\cdot\text{g}^{-1}$ fresh mass, corresponding to 0.26% of TSP. For the recombinant lectin domains of LecRK-I.9 and Lec-I.3 similar purification yields were achieved (17 and $21 \mu\text{g}\cdot\text{g}^{-1}$ fresh mass, corresponding to 0.28 and 0.37% of TSP, respectively).

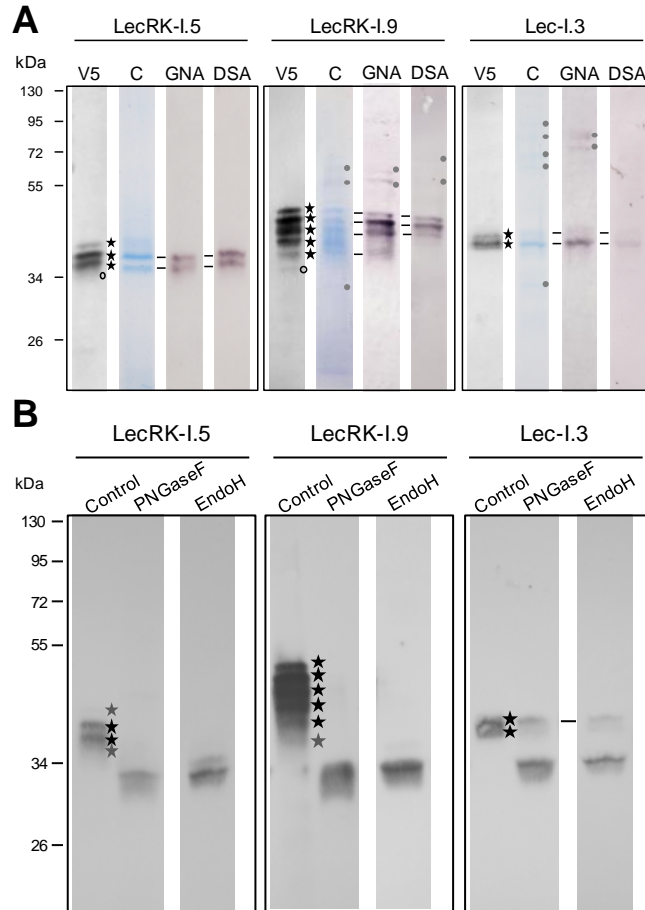


Fig. 8. N-glycosylation analysis of the purified recombinant lectin domains. After RuBisCO depletion of protein extracts of transformed *N. benthamiana* leaves, the recombinant lectin domains of LecRK-I.9, LecRK-I.5 and Lec-I.3 were purified by immobilised metal affinity chromatography. **(A)** The three purified recombinant lectin domains were detected by (i) immunoblotting using anti-V5 antibody (lanes V5), (ii) SDS-PAGE followed by Coomassie colloidal blue staining (lanes C), (iii) lectin-blotting using *Galanthus nivalis* agglutinin (lanes GNA), (iv) lectin-blotting using *Datura stramonium* agglutinin (lanes DSA). Stars indicate isoforms of the different recombinant lectin domains, open circles isoforms that were not detected by Coomassie colloidal blue staining, dots contaminating proteins and dashes isoforms revealed by agglutinins. **(B)** Enzymatic deglycosylation of the three recombinant lectin domains by Peptide-N-Glycosidase F (lane PNGaseF) or Endoglycosidase H (lane EndoH). Stars indicate isoforms of the different lectin domains in the control lanes (without enzyme). Dashes indicate the remaining high molecular mass bands for Lec-I.3 after enzymatic digestions.

The glycosylation status of the recombinant lectin domains was first examined using the EndoH and PNGaseF glycosidases. EndoH cleaves the chitobiose core of *N*-linked oligosaccharides belonging to the high-mannose-class or the hybrid-class of *N*-glycans. It does not cleave the complex-class of *N*-glycans. PNGaseF cleaves all *N*-linked oligosaccharides, except when an α 1-3 Fuc is present on the core GlcNAc. As shown in Fig.

8B, the enzymatic digestions of the recombinant lectin domains caused a decrease in the molecular mass of all them. The molecular masses of the deglycosylated lectin domains were around 30 kDa as expected for their peptide chains (Supplementary Fig. S1 and Fig. S2). The elimination of the carbohydrates from the lectin domains was very similar regardless the type of enzyme, suggesting that the complex-class of *N*-glycans was not present.

The glycosylation status of the lectin domains was further examined using lectin-blotting revealed with the digoxigenin (DIG)-labelled lectins GNA (*Galanthus nivalis* agglutinin) and DSA (*Datura stramonium* agglutinin) suitable to identify different types of *N*-glycans (see Material and Methods). For LecRK-I.5::V5::6xHis, the two major bands were detected by both GNA and DSA (Fig. 8). A third band, detected by Coomassie colloidal blue staining and the anti-V5 antibody, was not revealed in the lectin blot experiment, probably due to a lower amount of protein or to a *N*-glycan unrecognised by GNA and DSA. For LecRK-I.9::V5::6xHis, more complex patterns were observed: the three upper bands were detected by both GNA and DSA, the fifth band was detected only by GNA while the two others were not detected. Finally, for Lec-I.3::V5::6xHis, the two bands were detected by both GNA and DSA. Taken together, the results indicate that the recombinant lectin domains produced in *N. benthamiana* leaves were *N*-glycosylated.

Finally, the recombinant lectin domain of LecRK-I.9 was obtained after stable genetic transformation of *A. thaliana* using the same construct as for transient transformation of *N. benthamiana*. After RuBisCO depletion of TSP, SDS-PAGE and immunoblotting, recombinant lectin domains produced in *A. thaliana* exhibited similar features to that produced in *N. benthamiana*, *i.e.* molecular mass heterogeneity with multiple closely spaced bands ranging from 55 to 40 kDa (Supplementary Fig. S4). After enzymatic deglycosylation with PNGaseF, the molecular mass of the lectin domains was around 30 kDa, thus showing the presence of *N*-glycans.

4. Discussion

In the present study, we present a simple and robust procedure to purify recombinant Legume lectin domains from transiently transformed *N. benthamiana* leaves. We found that the depletion of the high-abundant protein RuBisCO from protein extracts can be obtained by differential precipitation using citrate buffers. This allowed improving the efficiency of further chromatographic steps and the recovery of multiple *N*-glycosylation isoforms of lectin domains in the purified fractions.

Protein precipitation relies on the disruption of protein hydration shell that forms around proteins and which modifies the repulsive electrostatic forces between proteins to favour their aggregation. The most common contributing parameters include pH, ionic strength, identity of buffer species and additives, most of them being linked to each other. In particular, citrate buffers were considered in studies aiming at deciphering aggregation mechanisms and protein-protein interactions for immunoglobulins (Barnett et al. 2015; Roberts et al. 2015). At low citrate concentrations (below 50 mM) and mild acidic pH, the electrostatic repulsions between proteins were shown to be reduced because of the accumulation of citrate ions at the protein surface. Interestingly, the effectiveness of the buffer at reducing the repulsion forces in solutions at pH 5.0 is the highest for citrate buffers, followed by succinate buffers (Roberts et al. 2015). In addition, in the process of precipitating protein to form crystals, di and tri-carboxylic acids play a dual role as buffer components and protein cross-linking agents (McPherson et al. 2011). In our conditions, *i.e.* a salting-in step after ammonium sulphate precipitation, RuBisCO could preferentially bind citrate or succinate, thus leading to protein aggregates which were removed by centrifugation. Finally, the overall procedure including the chromatographic steps provided purification yields in the range of those frequently described in the literature, *i.e.* between 0.1% and 1% of TSP

(Lombardi et al. 2012; Nishimoto et al. 2014). High-level transient protein expression in *N. benthamiana* leaves was obtained by constructing hydrophobin fusion proteins with the exceptional yield of 51% of TSP (Joensuu et al. 2010). Recombinant proteins were targeted to the ER in which they accumulated into protein bodies. However, glycoproteins to be localised at the plasma membrane or in the cell wall require going through the secretory pathway in order to have fully processed *N*-glycans. Consequently, a proper *N*-glycosylation is not compatible with hydrophobin fusion constructs.

Glycosylation is a very common post-translational modification for plant receptors and cell wall proteins (Song et al. 2013; Canut et al. 2016). Glycans contribute to protein folding and function, and take part in many physiological processes (Strasser 2014): thus, the activity of receptors and cell wall proteins may be dependent on their glycosylation status (Häweker et al. 2010; Yamamoto et al. 2014; Méndez-Yañez et al. 2017). In particular, proteins harbouring a N-terminal signal peptide will enter the sequential biosynthesis of *N*-glycans within the endomembrane system. The lectin domains of the two receptors LecRK-I.5 and LecRK-I.9 but also that of the cell wall protein Lec-I.3 display several predicted *N*-glycosylation sites. Our results indicate that most of the predicted *N*-glycosylation sites were occupied by carbohydrates as revealed by (i) the presence of multiple glycoforms, (ii) the decrease in molecular mass of the lectin domains after EndoH and PNGaseF digestions and (iii) the recognition by DSA and GNA which are specific for different *N*-glycan structures. The enzymatic deglycosylation indicated that the complex-class of *N*-glycans is not present within the recombinant lectin domains. DSA preferentially binds to tri- or higher-branched *N*-glycans containing the disaccharide GalGlcNAc with an unsubstituted Gal (Nishimoto et al. 2015): for instance, the GalGlcNAcMan₇GlcNAc₂ *N*-glycan structure belonging to the hybrid-class is a possible target. GNA strongly interacts with the high Man-class of *N*-glycans and preferentially recognition of terminal Man residues, (1-3), (1-6) or (1-2) linked to Man

(Fouquaert et al. 2009) including the $\text{Man}_{5-9}\text{GlcNAc}_2$ structures. For LecRK-I.9, six isoforms of the recombinant lectin domain were detected in a purified fraction. Starting from the lowest molecular mass isoform, an incremental increase of about 2 kDa separated the five other bands suggesting that the lectin domain could be decorated by either only one or up to six *N*-glycan moieties. However, we cannot rule out the possibility that some *N*-glycosylation sites are occupied by low molecular mass *N*-glycans such as the GlcNAc-only class (Zeng et al. 2018). A similar conclusion can be drawn for the recombinant lectin domains of LecRK-I.5 and Lec-I.3.

Heterogeneity within glycoforms of the recombinant lectin domains could also result from differences in the structure of the *N*-glycan moieties. Indeed, a number of bands were revealed both by DSA and GNA meaning that the *N*-glycan hybrid-class with terminal unsubstituted Gal and high Man-class could be present on the same lectin domain. This applied particularly for the two major isoforms of LecRK-I.5 and Lec-I.3. The example of LecRK-I.9 was even more intricate. The three first isoforms were revealed by both DSA and GNA. The fifth isoform was only detected by GNA meaning that (i) it did not contain the complex-class of *N*-glycans harbouring Gal such as the Lewis epitope, and (ii) the high Man-class was present, without excluding other structures. Finally, the fourth isoform was labelled neither by DSA nor by GNA: the presence of a structure belonging to the hybrid-class with a substituted Gal could be assumed. Recently, a comprehensive analysis of *N*-glycopeptides using advanced mass spectrometry has revealed the *N*-glycan micro-heterogeneity present in *A. thaliana* *N*-glycoproteins (Zeng et al. 2018). A total of 1110 *N*-glycopeptides were identified from 324 proteins exhibiting 492 *N*-glycosites, showing that one specific *N*-glycosite can be decorated by different *N*-glycans. In particular, Lec-I.3 was examined in this study. It reveals two Asn residues providing sites for *N*-glycosylation, namely Asn³³ and Asn⁸⁴. Both were predicted to be *N*-glycosites (Supplementary Fig. S2). At the Asn³³ position,

two different structures of *N*-glycans were identified, HexNAc(2)Hex(3)Fuc(1)Pent(1) and HexNAc(2)Hex(4)Fuc(1)Pent(1) belonging to the paucimannose- and hybrid-class, respectively. At Asn⁸⁴, three different structures were reported, HexNAc(2)Hex(3)Fuc(1)Pent(1), HexNAc(2)Hex(4)Pent(1) and HexNAc(3)Hex(3)Pent(1). Our study also revealed that the recombinant lectin domains of LecRK-I.9 produced in *A. thaliana* were highly *N*-glycosylated. There is still the question of whether this high glycosylation status for the recombinant lectin domains is of biological relevance. It will also be necessary to demonstrate the presence of *N*-glycans in the native proteins.

Our work also noted a marked difference between receptors and cell wall proteins in the number of their lectin domains glycoforms (Fig. 1). Furthermore, using bioinformatic prediction, more *N*-glycosites were detected in the Legume lectin domains of *A. thaliana* and *Medicago truncatula* receptors (Supplementary Fig. S5). That pointed out, respectively, an average of 5.5 and 5.6 predicted *N*-glycosites for receptors, and an average of 2.5 and 2.1 for cell wall proteins. Plant receptor-like kinases were already noted as heavily *N*-glycosylated proteins (Schoberer and Strasser 2018) but the biological function of the complex *N*-glycans is largely unexplored. The example of the S-locus receptor kinase involved in the self-incompatibility response of the *Brassicaceae* indicated that *N*-glycosylation functions primarily to ensure the proper and efficient subcellular trafficking to the plasma membrane through the ER quality control machinery (Yamamoto et al. 2014). The two closely related pattern-recognition receptors EFR and FLS2 were distinguished by a differential requirement for ER quality control components that could be linked to *N*-glycosylation (*e.g.* position, number of glycosylated sites, glycan structures) (Nekrasov et al. 2009). In particular, a single *N*-glycosite plays a critical role for EFR abundance and ligand recognition (Häweker et al. 2010). Noteworthy, in mammals, the number of *N*-glycans per protein and the branching of their structures provided mechanisms to regulate developmental processes (Lau et al. 2007;

Hang et al. 2016; Medina-Cano et al. 2018; Chandler et al. 2019). It is then possible that the lectin domains being part of receptor-like kinases have evolved to a high number of *N*-glycosites in order to ensure regulatory mechanisms through the ER quality control machinery.

In summary, our present results provide a purification scheme for the multiple *N*-glycoforms of recombinant Legume lectin domains produced by *N. benthamiana* leaves. Through depletion of the high-abundant protein RuBisCO from protein samples, the citrate buffer served as an effective non-chromatographic step for the initial fractionation and clarification of protein extracts. The next two chromatographic steps performed in the negative mode were dedicated to the enrichment in lectin domains before a Ni-NTA final affinity chromatography using the 6xHis tag: the multiple *N*-glycosylation lectins have the same chromatographic behaviour and were recovered as proteins purified near to homogeneity. The procedure can be readily adapted to other plant species. Despite being designed for the purification of recombinant proteins, the procedure is fully compatible with proteomic studies. It reveals the high degree of heterogeneity in the *N*-glycosylation patterns of the Legume lectin domains, and particularly of those of LecRKs. We believe that our procedure holds significant promise to reveal the functional characterization of *A. thaliana* Legume lectins in cell signalling and cell wall organisation.

Acknowledgements

The authors are thankful to the Paul Sabatier-Toulouse III University (France) and CNRS for supporting their research work. KB has been granted by the Paul Sabatier-Toulouse III University for his PhD work. AL and LL have been supported by the ANR grant

WALLARRAY (ANR-2010-Genom-BTV-010). This work was also supported by the ANR French Laboratory of Excellence project entitled "TULIP" (ANR-10-LABX-41; ANR-11-IDEX-0002-02).

Abbreviations

DSA, *Datura stramonium* agglutinin ; DIG, digoxigenin ; eGFP, enhanced green-fluorescent protein; GNA, *Galanthus nivalis* agglutinin; LB, Luria-Bertani medium; TSP, total soluble proteins; PEI, polyethyleneimine; PEG, polyethyleneglycol; PS, protamine sulphate; RuBisCO, ribulose biphosphate carboxylase/oxygenase.

References

Adar R, Sharon N. 1996. Mutational studies of the amino acid residues in the combining site of *Erythrina corallodendron* lectin. *Eur. J. Biochem.* 239:668-674.

André S, Siebert H, Nishiguchi M, Tazaki K, Gabius H. 2005. Evidence for lectin activity of a plant receptor-like protein kinase by application of neoglycoproteins and bioinformatic algorithms. *Biochim. Biophys. Acta* 1725:222-232.

Arfi ZA, Hellwig S, Drossard J, Fischer R, Buyel JF. 2016. Polyclonal antibodies for specific detection of tobacco host cell proteins can be efficiently generated following RuBisCO depletion and the removal of endotoxins. *Biotechnol. J.* 11:507-518.

Balagué C, Gouget A, Bouchez O, Souriac C, Haget N, Boutet-Mercey S, Govers F, Roby D, Canut H. 2017. The *Arabidopsis thaliana* lectin receptor kinase LecRK-I.9 is required for full resistance to *Pseudomonas syringae* and affects jasmonate signalling. *Mol. Plant Pathol.* 18:937-948.

Barnett GV, Razinkov VI, Kerwin BA, Laue TM, Woodka AH, Butler PD, Perevozchikova T, Roberts CJ. 2015. Specific-ion effects on the aggregation mechanisms and protein-protein interactions for anti-streptavidin immunoglobulin gamma-1. *J. Phys. Chem. B* 119:5793-5804.

Bellande K, Bono JJ, Savelli B, Jamet E, Canut H. 2017 Plant lectins and lectin receptor-like kinases: how do they sense the outside? *Int. J. Mol. Sci.* 18:1164.

Canut H, Albenne C, Jamet E. 2016. Post-translational modifications of plant cell wall proteins and peptides: a survey from a proteomics point of view. *Biochim. Biophys. Acta* 1864:983-990.

Cellar NA, Kuppannan K, Langhorst M, Ni W, Xu P, Young SA. 2008. Cross species availability of abundant protein depletion columns for ribulose-1,5-biphosphate carboxylase/oxygenase. *J. Chromatogr. B* 861:29-39.

Chandler KB, Leon DR, Kuang J, Meyer RD, Rahimi N, Costello CE. 2019. N-Glycosylation regulates ligand-dependent activation and signaling of vascular endothelial growth factor receptor 2 (VEGFR2). *J. Biol. Chem.* 294:13117-13130.

Choi J, Tanaka K, Cao Y, Qi Y, Qiu J, Liang Y, Lee S, Stacey G. 2014. Identification of a plant receptor for extracellular ATP. *Science* 343:290-294.

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

Gouget A, Senchou V, Govers F, Sanson A, Barre A, Rougé P, Pont-Lezica R, Canut H. 2006. Lectin receptor kinases participate in protein-protein interactions to mediate plasma membrane-cell wall adhesions in Arabidopsis. *Plant Physiol.* 140:81-90.

Hang Q, Isaji T, Hou S, Zhou Y, Fukuda T, Gu J. 2016. N-Glycosylation of integrin $\alpha 5$ acts as a switch for EGFR-mediated complex formation of integrin $\alpha 5\beta 1$ to $\alpha 6\beta 4$. *Sci Rep.* 6:33507.

Häweker H, Rips S, Koiwa H, Salomon S, Saijo Y, Chinchilla D, Robatzek S, von Schaewen A. 2010. Pattern recognition receptors require N-glycosylation to mediate plant immunity. *J. Biol. Chem.* 285:4629-4636.

Hijazi M, Roujol D, Nguyen-Kim H, Del Rocio Cisneros Castillo L, Saland E, Jamet E, Albenne C. 2014. Arabinogalactan protein 31 (AGP31), a putative network-forming protein in *Arabidopsis thaliana* cell walls? *Ann. Bot.* 114:1087-97.

Holler C, Zhang C. 2008. Purification of an acidic recombinant protein from transgenic tobacco. *Biotechnol. Bioeng.* 99:902-909.

Joensuu JJ, Conley AJ, Lienemann M, Brandle JE, Linder MB, Menassa R. 2010. Hydrophobin fusions for high-level transient protein expression and purification in *Nicotiana benthamiana*. *Plant Physiol.* 152:622-633.

Kim YJ, Lee HM, Wang Y, Wu J, Kim SG, Kang KY, Park KH, Kim YC, Choi IS, Agrawal GK, Rakwal R, Kim ST. 2013. Depletion of abundant plant RuBisCO protein using the protamine sulfate precipitation method. *Proteomics* 13:2176-2179.

Lau KS, Partridge EA, Grigorian A, Silvescu CI, Reinhold VN, Demetriou M, Dennis JW. 2007. Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. *Cell* 129:123-134.

Lombardi R, Donini M, Villani ME, Brunetti P, Fujiyama K, Kajiura H, Paul M, Ma JKC, Benvenuto E. 2012. Production of different glycosylation variants of the tumour-targeting mAb H10 in *Nicotiana benthamiana*: influence on expression yield and antibody degradation. *Transgenic Res.* 21:1005-1021.

McPherson A, Nguyen C, Cudney R, Larson SB. 2011. The role of small molecule additives and chemical modification in protein crystallization. *Cryst. Growth Des.* 11:1469–1474.

Medina-Cano D, Ucuncu E, Nguyen LS, Nicouleau M, Lipecka J, Bizot JC, Thiel C, Foulquier F, Lefort N, Faivre-Sarrailh C, Colleaux L, Guerrera IC, Cantagrel V. 2018. High N-glycan multiplicity is critical for neuronal adhesion and sensitizes the developing cerebellum to N-glycosylation defect. *eLife* 7: e38309.

Méndez-Yañez Á, Beltrán D, Campano-Romero C, Molinett S, Herrera R, Moya-León MA, Morales-Quintana L. 2017. Glycosylation is important for FcXTH1 activity as judged by its structural and biochemical characterization. *Plant Physiol. Biochem.* 119:200-210.

Nekrasov V, Li J, Batoux M, Roux M, Chu ZH, Lacombe S, Rougon A, Bittel P, Kiss-Papp M, Chinchilla D, van Esse HP, Jorda L, Schwessinger B, Nicaise V, Thomma BP, Molina A, Jones JD, Zipfel C. 2009. Control of the pattern-recognition receptor EFR by an ER protein complex in plant immunity. *EMBO J.* 28:3428-3438.

Nguyen C, Tanaka K, Cao Y, Cho S, Xu D, Stacey G. 2016. Computational analysis of the ligand binding site of the extracellular ATP receptor, DORN1. *PLoS One* 11:e0161894.

Nishimoto K, Tanaka K, Murakami T, Nakashita H, Sakamoto H, Oguri S. 2015. *Datura stramonium* agglutinin: cloning, molecular characterization and recombinant production in *Arabidopsis thaliana*. *Glycobiology* 25:157-169.

Roberts D, Keeling R, Tracka M, van der Walle CF, Uddin S, Warwicker J, Curtis R. 2015. Specific ion and buffer effects on protein-protein interactions of a monoclonal antibody. *Mol. Pharmaceutics* 12:179-193.

Romanini D, Braia MJ, Porfiri MC. 2013. Applications of calorimetric techniques in the formation of protein-polyelectrolytes complexes. In: Elkordi AA, editor. *Applications of calorimetry in a wide context*. London (UK): IntechOpen. p. 105-128.

Schoberer J, Strasser R. 2018. Plant glyco-biotechnology. *Semin. Cell Dev. Biol.* 80:133-141.

Singh P, Zimmerli L. 2013. Lectin receptor kinases in plant innate immunity. *Front Plant Sci.* 4:124.

Song W, Mentink RA, Henquet MG, Cordewener JH, van Dijk AD, Bosch D, America AH, van der Krol AR. 2013. N-Glycan occupancy of *Arabidopsis* N-glycoproteins. *J. Proteomics* 93:343–355.

Stephan A, Hahn-Löbmann S, Rosche F, Buchholz M, Giritch A, Gleba Y. 2018. Simple purification of *Nicotiana benthamiana*-produced recombinant colicins: high-yield recovery of purified proteins with minimum alkaloid content supports the suitability of the host for manufacturing food additives. *Int. J. Mol. Sci.* 19:95-107.

Strasser R. 2014. Biological significance of complex N-glycans in plants and their impact on plant physiology. *Front. Plant Sci.* 5:363.

Van Damme E, Lannoo N, Peumans W. 2008. Plant lectins. *Adv. Bot. Res.* 48:107-209.

Wang C, Zhou M, Zhang X, Yao J, Zhang Y, Mou Z. 2017. A lectin receptor kinase as a potential sensor for extracellular nicotinamide adenine dinucleotide in *Arabidopsis thaliana*. *eLife* 6: e25474.

Wang C, Huang X, Li Q, Zhang Y, Li JL, Mou Z. 2019. Extracellular pyridine nucleotides trigger plant systemic immunity through a lectin receptor kinase/BAK1 complex. *Nat. Commun.* 10:4810.

Widjaja I, Naumann K, Roth U, Wolf N, Mackey D, Dangl JL, Scheel D, Lee J. 2009. Combining subproteome enrichment and Rubisco depletion enables identification of low abundance proteins differentially regulated during plant defense. *Proteomics* 9:138-147.

Xi J, Wang X, Li S, Zhou X, Yue L, Fan J, Hao D. 2006. Polyethylene glycol fractionation improved detection of low-abundant proteins by two-dimensional electrophoresis analysis of plant proteome. *Phytochemistry* 67:2341-2348.

Yamamoto M, Tantikanjana T, Nishio T, Nasrallah ME, Nasrallah JB. 2014. Site-specific *N*-glycosylation of the *S*-locus receptor kinase and its role in the self-incompatibility response of the brassicaceae. *Plant Cell* 26:4749-4762.

Yusibov V, Kushnir N, Streatfield SJ. 2016. Antibody production in plants and green algae. *Annu. Rev. Plant Biol.* 67:669-701.

Zeng W, Ford KL, Bacic A, Heazlewood JL. 2018. *N*-linked glycan micro-heterogeneity in glycoproteins of *Arabidopsis*. *Mol. Cell. Proteomics* 17:413-421.

Zhang C, Lillie R, Cotter J, Vaughan D. 2005. Lysozyme purification from tobacco extract by polyelectrolyte precipitation. *J. Chromatogr. A* 1069:107-112.

Legends for supplementary material:

Fig. S1. The lectin domains of LecRK-I.9 and LecRK-I.5 are potentially highly glycosylated.

Fig. S2. The lectin domains of Lec-I.3 and Lec-I.6 are potentially glycosylated.

Fig. S3. Effect of polyethyleneglycol (PEG) on RuBisCO precipitation and recombinant lectin domain recovery.

Fig. S4. Deglycosylation of the partially purified recombinant lectin domains of LecRK-I.9 produced in *A. thaliana*.

Fig. S5. Putative *N*-glycosites of Legume lectin domains occur more frequently in LecRKs than in Lecs.

A - LecRK-I.9::V5::6xHis

1 MARWLLQILI ISSLHLSSVS SQQETSFVYE SFLDRQNLVYL DKSAIVLPSG
 51 LLQLTNASEH QMGHAFHKKP IEFSSSGPLS FSTHFVCALV PKPGFEGGHG
 101 IVFVLSPSMD FTHAESTRYL GIFNASTNGS SSYHVLAVEL DTIWNPDFKD
 151 IDHNVGIDV NSPISVAIAS ASYYS DMKGS NESINLLSGN PIQVWVDYEG
 201 TLLNVS VAPL EVQKPTRPLL SHPINLTEL PNRSSLFAGF SAATGTAISD
 251 QYILWWSFSI DRGSLQRLDI SKLPEVPHPR APHKVSKGE LEGKPIP NPL
 301 LGLDSTRTG HHHHH

| Position | Potential | Jury agreement | N-Glyc result | |
|----------|-----------|----------------|---------------|----|
| 56 | NASE | 0.6493 | (7/9) | + |
| 124 | NAST | 0.5673 | (6/9) | + |
| 128 | NGSS | 0.6333 | (8/9) | + |
| 181 | NESI | 0.5841 | (8/9) | + |
| 204 | NVSV | 0.6798 | (9/9) | ++ |
| 225 | NLTE | 0.5311 | (5/9) | + |
| 232 | NRSS | 0.5540 | (6/9) | + |

B - LecRK-I.5::V5::6xHis

1 MSKGLFLIWL ISSFHLISFS TSSKDTSFVF NGFGQSNLAL DGSATLLPNG
 51 LLQLAKDSQH QMGHAFIKKP IDFSSSKPLS FSTHFVCALV PKPGFEGGHG
 101 ITFVISPTVD FTRAQPTRYM GIFNASTNGS PSSHLFAVEL DTVRNPDFRE
 151 TNNNHIGIDV NNPISVESAP ASYFSKTAQK NVSINLSSGK PIQVWVDYHG
 201 NVLNVS VAPL EAEKPSLPLL SRSMNLSEIF SRRRLFVGFA AATGTSISYH
 251 YLLGWSFSKG ELEGKPIP NPL LLGLDSTRTG HHHHHH

| Position | Potential | Jury agreement | N-Glyc result | |
|----------|-----------|----------------|---------------|-----|
| 124 | NAST | 0.5387 | (6/9) | + |
| 128 | NGSP | 0.1248 | (9/9) | --- |
| 181 | NVSI | 0.5299 | (7/9) | + |
| 185 | NLSS | 0.4806 | (5/9) | - |
| 204 | NVSV | 0.5641 | (6/9) | + |
| 225 | NLSE | 0.5382 | (5/9) | + |

Fig. S1. The lectin domains of LecRK-I.9 and LecRK-I.5 are potentially highly glycosylated. Amino acid sequences of the recombinant lectin domain of LecRK-I.9 (A) and LecRK-I.5 (B) showing: predicted signal peptides (1-21 and 1-22), predicted N-glycosylation sites (Asn-Xaa-Ser/Thr sequons in the sequence are highlighted in blue; Asn residues predicted to be N-glycosylated are highlighted in red), the V5 epitope and the 6xHis tag (red italics). The results for the prediction of N-glycosylation sites in the recombinant lectin domains were performed using the NetNGlyc server (<http://www.cbs.dtu.dk/services/NetNGlyc/>). The sequons that are predicted to be N-glycosylated are highlighted in yellow boxes (threshold 0.5).

A - Lec-I.3::V5::6xHis

1 MKIQILCF^{TT} LFLAIF^{TSQV} TTAYKFKFDY FG^{NGT}DPISE HGDAEYGPDT
 51 DGKSRSGAIA LTRDNIPFSH GRAIFTTPIT FKP^{NAS}ALYP FKTSFTFSIT
 101 PKTNP^{NQGHG} LAFIVVPSN^Q NDAGSGLGYL SLL^{NRT}NNGN PNNHLFAVEF
 151 DVFQDKSLGD MNDNHVGIDI NSVDSVSVK SGYWV^{MTRSG} WLFKDLKLS
 201 GDRYKAWIEY NNNYKVSVT IGLAHLK^{KPN} RPLIEAKFDL SKVIHEV^{MYT}
 251 GFAGSMGRGV ERHEIWDWTF QNKGELE^{GKP} IPN^{PLLGLDS} TRT^{GHHHHH}

| Position | Potential | Jury agreement | N-Glyc result | |
|----------|-----------|----------------|---------------|----|
| 33 | NGTD | 0.7247 | (9/9) | ++ |
| 84 | NASA | 0.5043 | (5/9) | + |
| 134 | NRTN | 0.7301 | (9/9) | ++ |

B - Lec-I.6::V5::6xHis

1 MQIHKLCFLV LFLANAFAV KFNFD^{SFDGS} NLLFLGDAEL GPSSDGVSR
 51 GALSMTRDEN PFSHGQGLYI NQIPFKPS^{NT} SSPFSFETS^F TFSITPR^{TKP}
 101 NSGQGF^{AFII} TPEADNSGAS DGGYL^{GILNK} TNDGK^{PENHI} LAIEFD^{TFQN}
 151 KEFLDISGNH VGVNINSMTS LVAEKAGYVW QTRVGK^{RKVW} SFKDV^{NLSSG}
 201 ERFKAWVEFR NKDSTITVTL APENVK^{KPKR} ALIEAP^{RVLN} EVLLQ^{NMYAG}
 251 FAGSMGRAVE RHD^{IWSWFE} NAAKNN^{KGEL} EGK^{PIPPLL} GLD^{STRTGH}
 301 HHHH

| Position | Potential | Jury agreement | N-Glyc result | |
|----------|-----------|----------------|---------------|-----|
| 79 | NTSS | 0.7094 | (9/9) | ++ |
| 129 | NKTN | 0.7736 | (9/9) | +++ |
| 196 | NLSS | 0.6688 | (9/9) | ++ |

Fig. S2. The lectin domains of Lec-I.3 and Lec-I.6 are potentially glycosylated. Amino acid sequences of the recombinant lectin domain of Lec-I.3 (A) and Lec-I.6 (B) showing: predicted signal peptides (1-23 and 1-19), predicted N-glycosylation sites (Asn-Xaa-Ser/Thr sequons in the sequence are highlighted in blue; Asn residues predicted to be N-glycosylated are highlighted in red), the V5 epitope and the 6xHis tag (red italics). The results for the prediction of N-glycosylation sites in the recombinant lectin domains were performed using the NetNGlyc server (<http://www.cbs.dtu.dk/services/NetNGlyc/>). All the sequons are predicted to be N-glycosylated and highlighted in yellow boxes (threshold 0.5).

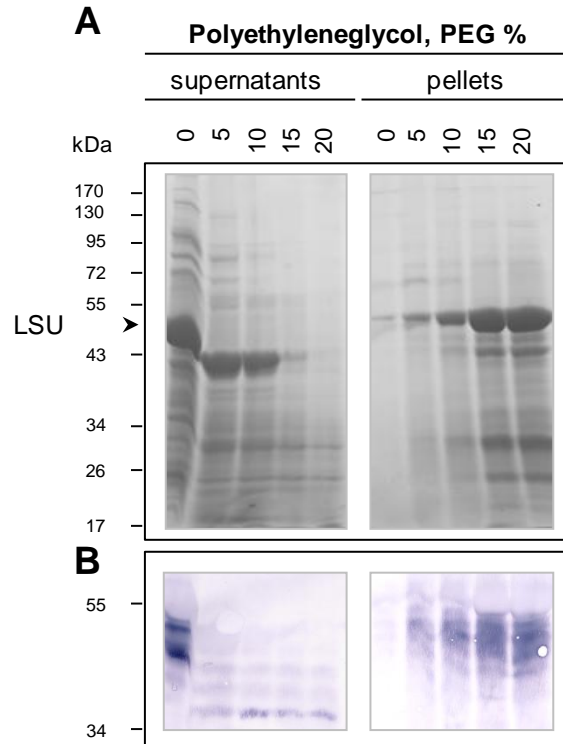


Fig. S3. Effect of polyethyleneglycol (PEG) on RuBisCO precipitation and recombinant lectin domain recovery. (A) Total soluble proteins of *N. benthamiana* leaves after transient transformation for the production of recombinant lectin domains of LecRK-I.9 were subjected to different concentrations of PEG ranging from 5% to 20%. After centrifugation, proteins from supernatants and pellets were analysed by SDS-PAGE and stained with Coomassie colloidal blue. Abundant RuBisCO large subunit (LSU) is marked by black arrowhead. (B) Immunoblots corresponding to (A) were revealed with anti-V5 antibody to detect recombinant lectin domains of LecRK-I.9.

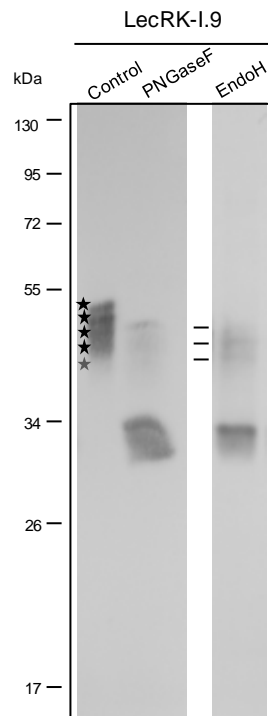


Fig. S4. Deglycosylation of the partially purified recombinant lectin domains of LecRK-I.9 produced in *A. thaliana*. After RuBisCO depletion of protein extracts from transformed *A. thaliana* plants, the recombinant lectin domains of LecRK-I.9 were deglycosylated using Peptide-*N*-Glycosidase F (lane PNGaseF) or Endoglycosidase H (lane EndoH). Stars indicate isoforms of the LecRK-I.9 lectin domains in the control lane (without enzyme). Dashes indicate remaining high molecular mass bands after enzymatic digestions.

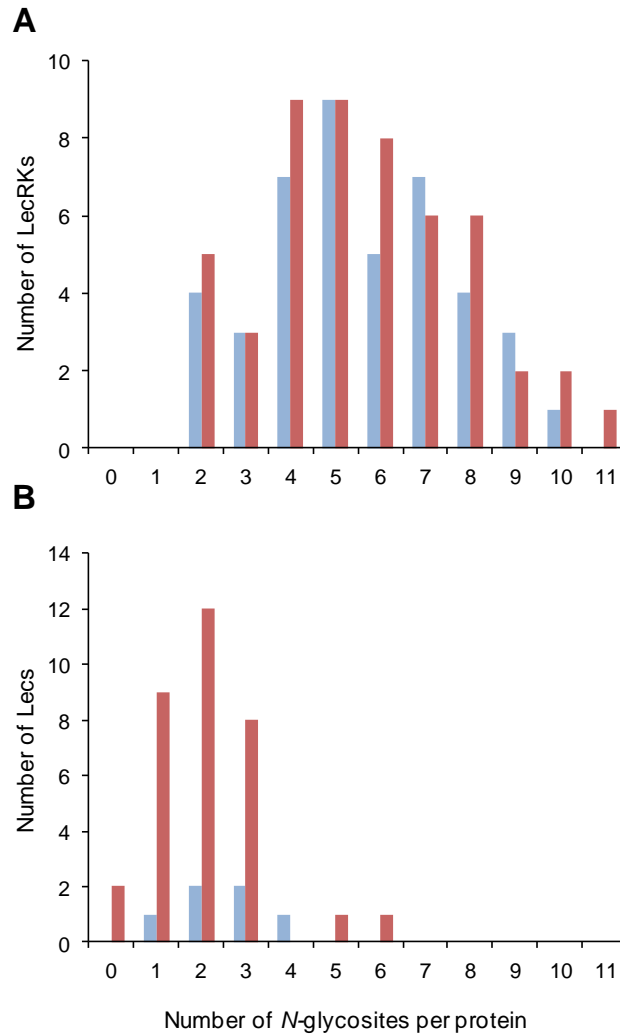


Fig. S5. Putative *N*-glycosites of Legume lectin domains occur more frequently in LecRKs than in Lecs. Distribution of LecRKs (**A**) and Lecs (**B**) according to the number of predicted *N*-glycosites in their lectin domains. Two plant species were considered *A. thaliana* (blue bars) and *M. truncatula* (red bars). Data were retrieved from Pfam (www.pfam.xfam.org – version 32, September 2018): the number of sequences including a predicted Legume lectin domain (PF00139) was of 43 and 51 LecRKs, 6 and 33 Lecs for *A. thaliana* and *M. truncatula* respectively. The prediction of *N*-glycosites for the lectin domains was performed using the NetNGlyc server (<http://www.cbs.dtu.dk/services/NetNGlyc/>).