Characterization and sugar-binding properties of arcelin-1, an insecticidal lectin-like protein isolated from kidney bean (*Phaseolus vulgaris* L. cv. RAZ-2) seeds

Christine FABRE*, Henri CAUSSE*, Lionel MOUREY*, Jos KONINKX†, Michel RIVIÈRE*, Henno HENDRIKS†, Germain PUZO*, Jean-Pierre SAMAMA* and Pierre ROUGÉ*¹

*Institut de Pharmacologie et Biologie Structurale, UPR CNRS no. 9062, 205 Route de Narbonne, 31077 Toulouse, France, and †Department of Veterinary Pathology, Faculty of Veterinary Medicine, University of Utrecht, Yalelaan 1, Postbus 80158, 3508 TD Utrecht, The Netherlands

Arcelin-1 is a lectin-like protein found in the seeds of wild varieties of the kidney bean (*Phaseolus vulgaris*). This protein displays insecticidal properties, but the mechanism of action is as yet unknown. In the present study we investigated the biochemical and biophysical properties of arcelin-1 from *Phaseolus vulgaris* cv. RAZ-2. Native arcelin-1 is a dimeric glycoprotein of 60 kDa, built from the non-covalent association of two identical monomers. This dimer resists dissociation by chaotropic agents and is highly resistant to proteolytic enzymes. Each subunit contains

10% (w/w) neutral sugars which belong to the high-mannose and complex-type glycans attached to three glycosylation sites. No interaction of the protein with simple sugars could be detected, but arcelin-1 displays an intrinsic specificity in binding complex glycans. Arcelin-1 therefore differs from the closely related phytohaemagglutinin lectins and α -amylase inhibitor in several respects: oligomerization states, sugar-binding affinities and the type and number of glycan chains. These features may be related to the toxicity of arcelin-1.

INTRODUCTION

Seeds of kidney bean (*Phaseolus vulgaris*) express the closely related lectins PHA-E and PHA-L (erythroagglutinating and leucoagglutinating subunits of phytohaemagglutinin) and the lectin-like proteins arcelins and α -amylase inhibitors [1]. All these proteins, which are encoded by a multigenic family, are likely involved in the defence of the plant against insects and pathogenic fungi or bacteria.

Arcelin has insecticidal properties against the larvae of various bruchids, e.g. the Mexican bean weevil (*Zabrotes subfasciatus*) [2]. Five variants of arcelin were previously identified [3], and most of them have been isolated and partially characterized [4–7]. Although the insecticidal properties of the various arcelin variants against bruchid pests are now well documented [6,8], their mechanism of action is controversial. According to Osborn et al. [8], arcelin behaves as a toxic protein, whereas Minney et al. [9] consider arcelin to be an indigestible protein which induces the starvation of the larvae. A better knowledge of the biochemical properties and of the mechanism of action of arcelin appears, nevertheless, as a prerequisite to authorize their use in creating insect-resistant transgenic crops. This strategy was successfully used to protect transgenic peas (*Pisum sativum*) harbouring the cDNA encoding the α -amylase inhibitor (α -AI) of *P. vulgaris* against various predatory insects [10,11]. In the present study we characterized the oligomerization state, the extent and type of the N-glycosylation, and the specific glycan-binding properties of arcelin-1 devoid of contamination by PHA and α -AI.

EXPERIMENTAL

Materials

Mature seeds of kidney bean [*Phaseolus vulgaris* L., cv. RAZ-2 of the Centro Internacional de Agricultura Tropical (CIAT)] were used for arcelin-1 isolation and purification.

Sugars, sugar derivatives, saccharides and glycoproteins used in hapten inhibition of haemagglutination and biosensor measurements were purchased from Sigma. PHA from lima bean (*P. lunatus*), and unlabelled and FITC-labelled PHA-E and PHA-L were obtained from Vector Laboratories (Burlingame, CA, U.S.A.). Snowdrop (*Galanthus nivalis*) agglutinin (GNA) and elderberry (*Sambucus nigra*) agglutinin (SNAI) were gifts from Dr. W. J. Peumans (Katholieke Universiteit Leuven, Heverlee-Leuven, Belgium). FITC used for labelling of arcelin-1 was from Sigma. Trypsin, chymotrypsin and pepsin used to check the resistance of arcelin-1, PHA-E and PHA-L against proteolysis were purchased from Sigma. Radioactive precursors were purchased from Amersham Nederland B. V. (Houten, The Netherlands). Peptide N-glycosidase F (PNGase F) was from Boehringer-Mannheim.

Sensor chips (CM 5) and all the chemicals required for the activation of the carboxymethylated dextran and the immobilization of glycoproteins [100 mM *N*-hydroxysuccinimide, 400 mM *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodi-imide hydrochloride and 1 M ethanolamine hydrochloride adjusted to pH 8.5 with NaOH] were obtained from Pharmacia Biosensor AB. Hepes-buffered saline [HBS; 10 mM Hepes (pH 7.4)/150 mM

Abbreviations used: GNA, *Galanthus nivalis* (snowdrop) agglutinin; HBS, Hepes-buffered saline; HPAEC-PAD, high-pH anion-exchange chromatography with pulsed amperometric detection; LoLI, isolectin I from *Lathyrus ochrus* (vetch); PHA-E and PHA-L, erythroagglutinating and leucoagglutinating subunits of the phytohaemagglutinin from *Phaseolus vulgaris* (kidney bean); PNGase F, peptide N-glycosidase F from *Flavobacterium meningosepticum*; PsA, *Pisum sativum* (garden pea) agglutinin; RIP, ribosome-inactivating protein; SAXS, small-angle X-ray scattering; SNAI, *Sambucus nigra* (elderberry) agglutinin; SPR, surface plasmon resonance; TFA, trifluoroacetic acid; α -AI, α -amylase inhibitor; IEF, isoelectric focusing; *a*, Stokes radius; *R*_g, radius of gyration; PGK, 3-phosphoglycerate kinase; O Rh+, blood-group-O Rhesus-positive; HU, haemagglutination units; DMEM, Dulbecco's modified Eagle medium; CIAT, Centro Internacional de Agricultura Tropical.

¹ To whom correspondence should be addressed.

NaCl, containing 0.05 % *BIA*core surfactant P20] used for the biosensor measurements was from Pharmacia Biosensor AB.

Materials (tissue-culture plates and flasks) and chemicals (culture medium, non-essential amino acids, gentamycin, buffers, fetal-calf serum) used for Caco-2 cell cultures were purchased from Greiner (Alphen a/d Rijn, The Netherlands) and Flow Laboratories (Amstelslad B. V., Amsterdam, The Netherlands) respectively.

Isolation of arcelin-1

Arcelin-1 was isolated using a combination of classical protein-purification techniques, including ion-exchange chromatography and gel chromatofocusing. A 10 g portion of finely powdered *P. vulgaris* seed flour was extracted in 100 ml of 25 mM acetate buffer, pH 4.6, by stirring overnight at 4 °C. The homogenate was centrifuged at 15000 *g* for 25 min at 4 °C, and the resulting supernatant was filtered through cottonwool to remove the floating particles. All the following operations were conducted at room temperature. A 50 ml portion of the cleared extract was filtered through a column (40 cm × 1.6 cm, 80 ml bed volume) of CM-Trisacryl M (Sepracor, Villeneuve la Garenne, France) previously equilibrated with acetate buffer at a flow rate of 10 ml/min. The column was rinsed overnight with acetate buffer and eluted with 1 M NaCl at a flow rate of 30 ml/min. The pooled saline fractions were adjusted to 90 % (w/v) (NH₄)₂SO₄ and the resulting protein precipitate was dialysed against running tap water for 72 h. After centrifugation at 15000 *g* for 25 min at 4 °C, the water extract (5 ml) was loaded on to a column (40 cm × 1 cm, 28 ml bed volume) of PBE 94 (Pharmacia) equilibrated with 25 mM imidazole/HCl, pH 7.35, and elution was performed at a flow rate of 20 ml/min with a tenfold (v/v) dilution of Polybuffer 74 (Pharmacia) adjusted to pH 5.1 with 1 M HCl. The arcelin-containing fractions eluted at about pH 6.4 were pooled, precipitated with 90 % (NH₄)₂SO₄ and resuspended in 5 ml of distilled water. The protein was dialysed against running tap water for 48 h and distilled water for 24 h. After centrifugation at 15000 *g* for 25 min at 4 °C, arcelin-1 was freeze-dried and stored dry until used. Following this purification procedure 10 g of seed flour yielded about 50 mg of arcelin-1.

Labelling of arcelin-1 with FITC

For labelling of arcelin-1 with FITC, all the operations were carried out in the dark at room temperature. A 50 ml portion of a solution prepared by dissolving 5 mg of FITC in *N,N*-dimethylformamide was added under continuous stirring to 0.5 ml of a solution containing 5 mg of arcelin-1 in 0.1 M Na₂CO₃/NaHCO₃ buffer, pH 9.0. After 1 h of incubation at room temperature, the mixture was loaded on a column (20 cm × 1.6 cm, 30 ml bed volume) of Ultrogel AcA 202 (IBF, Villeneuve la Garenne, France) equilibrated with 10 mM PBS, pH 7.4, containing 150 mM NaCl. Elution, monitored at 280 nm, was performed at a flow rate of 20 ml/min, and the fractions eluted in the first peak were pooled and then dialysed against the same buffer. FITC-labelled arcelin-1 was kept in the dark at 4 °C in the presence of 0.01 % (w/v) NaN₃.

Analytical methods

The total neutral-sugar content of arcelin-1 was estimated colorimetrically by the phenol/H₂SO₄ method [12], using D-glucose as standard.

SDS/PAGE was performed on the Phast System (Pharmacia) using ready-made 12.5 and 20 % acrylamide gels. Isoelectric

focusing (IEF) was performed over the pH range 3.0–9.0 using ready-made IEF 3–9 Phastgels. The Coomassie Blue staining method used to reveal the separated protein bands can detect 20–30 ng of protein.

The bicinchoninic acid method [13] and the micro-biuret method [14] were carried out for protein determinations (Pierce), with BSA as standard. The tyrosine content was estimated by the modified Folin method [15], with tyrosine as reference.

Glycan-chain analysis

Neutral sugars of arcelin-1 were hydrolysed in 4 M trifluoroacetic acid (TFA) at 100 °C for 4 h. The monosaccharides were separated isocratically with 50 mM sodium acetate/1 mM NaOH on a CarboPac PA1 analytical column (Dionex) and analysed by high-pH anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD). This detection method allows carbohydrate analysis to be carried out at the picomolar level without any chemical derivatization.

Arcelin-1 (1 mg/ml) was dissolved in 20 mM sodium phosphate buffer, pH 7.5, containing 50 mM EDTA, and was denatured at 100 °C for 2 min. After cooling, PNGase F from *Flavobacterium meningosepticum* (100 µg/ml) was added and the mixture was incubated at 37 °C for 18 h. The oligosaccharides of the PNGase F digest were separated on a CarboPac PA1 column at a flow rate of 1 ml/min with 500 mM sodium acetate/100 mM NaOH (1:24, v/v) over 25 min, followed by a first gradient from 4 to 20 % (v/v) 500 mM sodium acetate over 25 min and a second gradient from 20 to 60 % (v/v) 500 mM sodium acetate over 10 min, and analysed for oligosaccharide components by HPAEC-PAD. (Man)₅-(GlcNAc)₂ to (Man)₆-(GlcNAc)₂ standards (Oxford Glycosystem) were used to identify the oligosaccharides of the high-mannose type. Digestion by PNGase A (Boehringer-Mannheim) was performed according to the manufacturer's instructions and the resulting oligosaccharides were separated as described for the PNGase F digest.

Analytical size-exclusion chromatography

This was performed at 4 °C on Superdex 75 10/30 (optimal separation range 30–70 kDa) and Superdex 200 10/30 (optimal separation range 10–600 kDa) (Pharmacia). The columns were equilibrated with 100 mM Tris/150 mM NaCl (pH 7.6)/50 mM Mes/150 mM NaCl, pH 4.9. Samples (100 µl) with protein concentrations ranging from 1 to 2 mg/ml were injected at a flow rate of 0.5 ml/min and fractions of 0.5 ml corresponding to arcelin-1 were collected. The columns were calibrated using BSA [67 kDa, Stokes radius (*a*) 3.55 nm (35.5 Å)], ovalbumin [43 kDa, *a* 3.05 nm (30.5 Å)] and chymotrypsinogen A (25 kDa, *a* 2.09 nm (20.9 Å)] (Pharmacia). Elution volumes were measured for standard proteins and arcelin-1. The *K_{av}* values were calculated for each protein by using the equation:

$$K_{av} = (V_e - V_0)/(V_t - V_0)$$

The *a* values were calculated from the plot of (–log *K_{av}*)^{1/2} versus *a*, and the molecular mass was calculated from a plot of *K_{av}* versus log(molecular mass).

Measurements under denaturing conditions were similarly performed on the Superdex 75 column in the presence of increasing concentrations of urea (1 M, 3 M and 5 M) or guanidinium chloride (3 M, 4 M, 5 M and 6 M). BSA was used as standard.

Small-angle X-ray scattering (SAXS)

For SAXS measurements, freeze-dried arcelin-1 was dissolved in pure water at a concentration of 10 mg/ml and then centrifuged for 1 h at 4 °C and 10000 *g* to remove insoluble material. Before SAXS measurements, protein aliquots were mixed with the buffer to give the following final conditions: 5 mg/ml arcelin-1, 100 mM Mes, 300 mM sodium citrate, pH 4.9. The samples contained 1.3 mM dithiothreitol to eliminate the free radicals formed in solution under X-ray irradiation.

X-ray scattering spectra were recorded at 4 °C on the small-angle scattering beam line D24 using synchrotron radiation at Laboratoire pour l'Utilisation du Rayonnement Electromagnétique (Orsay, France). The instrument [16] and the data-acquisition system [17] have already been described. The wavelength (λ) of the X-rays was 0.1488 nm (1.488 Å) (K-edge of Ni) and the sample-to-detector distance was set to 1316 mm. Scattered intensities were recorded in the angular range $0.00183 < s$ (Å⁻¹) < 0.04118 , where s is the scattering vector [$(2\sin\theta)/\lambda$, where 2θ and λ are the scattering angle and wavelength of the X-rays respectively]. Eight successive frames of 100 s each were collected for each sample. The scattering intensity of a reference sample of carbon black recorded immediately before and after each sample was used to normalize all data to the transmitted intensity. The scattering contribution from the medium was subtracted from each sample before further analysis.

The radius of gyration (R_g) was deduced from Guinier's law:

$$I(s) = I(0)\exp(-4\pi^2 R_g^2 s^2/3)$$

where $I(s)$ is the SAXS intensity and $I(0)$ is the value of the extrapolated scattered intensity at $s = 0$ [18].

For monodisperse solutions, the Guinier plot of $\ln I(s)$ versus s^2 approximates to a straight line which yields the R_g in the small-angle region. The value of $I(0)/c$, with c the protein concentration, is proportional to the molecular mass of the scattering entity. It was calibrated using a monodisperse solution of 3-phosphoglycerate kinase (PGK, 44.5 kDa). Measurements with PGK were carried out in 20 mM Tris buffer, pH 7.5.

Amino-acid-sequence analysis

The N-terminal sequencing of arcelin-1 was performed on an Applied Biosystems model 477A protein sequencer using the Edman-degradation procedure.

Resistance to proteolysis

In order to check the resistance of arcelin-1, PHA-E and PHA-L towards proteolytic enzymes, each protein was digested under the appropriate conditions of pH (7.6, 7.8 and 2.0 for trypsin, α -chymotrypsin and pepsin digestions respectively) and temperature (25 °C for trypsin and α -chymotrypsin digestions, 37 °C for pepsin digestion) in the presence of tosylphenylalanylchloromethane ('TPCK')-treated trypsin, α -chymotrypsin and pepsin, and the time course of the proteolysis was monitored by measuring, after 5 h of incubation, the remaining amounts of the protein and the increased amounts of released tyrosine. Proteins were used at a constant concentration of 5 mg/ml in the appropriate buffer (10 mM PBS, pH 7.6, and 10 mM PBS, pH 7.8, for trypsin and α -chymotrypsin respectively; 0.2 M KCl adjusted at pH 2.0 with 1 M HCl was used for pepsin) and the ratio protease/protein was fixed at 1/50 (w/w). Casein was used as standard protein for trypsin and α -chymotrypsin digestions, whereas BSA was taken as reference for pepsin digestion. Results are the means for three separate analyses.

Haemagglutination assays

The haemagglutinating activity of arcelin-1 was measured as previously described [19] in U-bottomed micro-titration plates (Flow Laboratories). Briefly, 25 μ l of two-fold serial dilutions of arcelin-1 in 10 mM PBS, pH 7.4, were mixed at room temperature with an equal volume of a 1% suspension of thrice-washed untreated or trypsin-treated rabbit or human blood-group-O Rhesus-positive (O Rh+) erythrocytes in the same buffer. Haemagglutination was read 2 h later at room temperature and, for the control, after the mixture had been left for at 4 °C for 12 h. The haemagglutination titre corresponds to the higher dilution still giving a complete haemagglutination pattern.

Trypsin-treated human or rabbit erythrocytes were prepared by adding 1 mg of trypsin (Merck) to 1 ml of thrice-washed erythrocytes in 10 mM PBS, pH 7.4. After 1 h of incubation at 37 °C, erythrocytes were washed three times with PBS to eliminate trypsin.

Hapten inhibition of haemagglutination

To measure hapten inhibition of haemagglutination, 25 μ l of a twofold dilution of sugar in 10 mM PBS, pH 7.4, was mixed with an equal volume of arcelin-1 solution corresponding to 4 haemagglutination units (4 HU) and the plate was incubated at room temperature for 1 h. 25 μ l of a 1% suspension of thrice-washed untreated or trypsin-treated rabbit or human O Rh+ erythrocytes in the same buffer were added and inhibition of haemagglutination was read 4 h later at room temperature. Results are expressed as the minimum concentration (mM) of monosaccharide, oligosaccharide or glycoprotein required to completely inhibit 4 HU. Account was taken of the threefold dilution caused by the addition of arcelin-1 and erythrocytes.

ELISA measurements

The pentadecapeptide corresponding to the sequence stretch Arg¹⁰¹-Tyr-Leu-Gly-Leu-Phe-Asn-Thr-Thr-Asn-Tyr-Asp-Arg-Asp-Ala¹¹⁵ of arcelin-1 was synthesized by a solid-phase-synthesis method using fluoren-9-ylmethoxycarbonyl ('Fmoc') polyamide active ester chemistry [20]. For coupling to BSA, equal amounts of peptide (3 mg) and BSA (3 mg) were mixed in 6 ml of 0.1 M NaHCO₃ in the presence of 3 ml of 25% (w/v) glutaraldehyde and the mixture was stirred for 24 h. After extensive dialysis against distilled water the resulting conjugate was freeze-dried. The percentage of coupling was near 100%, as estimated by weighing. Polyclonal antibodies were raised in rabbits according to the following immunization schedule. A first injection of 200 μ g of conjugate into a mixture (1:1, v/v) of saline and modified Freund's complete adjuvant (Calbiochem, France Biochem, Meudon, France) was followed by several weekly injections of 200 μ g of conjugate in saline alone. Blood was collected by cardiac puncture 10 days after the eighth booster injection and immune sera obtained after blood clotting for 4 h at room temperature and overnight at 4 °C, were stored freeze-dried at -30 °C until used.

ELISA measurements were performed as previously described [21]. Briefly, the wells of a microtitration plate (Flow Laboratories) were loaded with 150 μ l of antigen (arcelin-1, PHA-E, PHA-L or PHA from lima bean) solution in 0.1 M Na₂CO₃/NaHCO₃ buffer, pH 9.6, used at various concentrations ranging from 0.625 to 50 μ g/ml. After overnight incubation at 4 °C, the wells were flooded with 10 mM PBS, pH 7.4, containing 0.05% (v/v) Tween-20 (PBST1) and 200 μ l of a BSA solution at 10 mg/ml in PBS, pH 7.4, used as blocking reagent, were added. After 1 h of incubation at 37 °C, the wells were flooded with

PBST1 and polyclonal antisera (150 μ l) diluted 1:5000 in 10 mM PBS, pH 7.6, containing 0.025% (v/v) Tween-20 (PBST2) were added. After 1 h of incubation at 37 °C, the wells were rinsed with PBST1 and goat anti-rabbit globulin (150 μ l) labelled with alkaline phosphatase (Sigma), diluted 1:2000 in PBST2, was added. After 1 h of incubation at 37 °C, the wells were flooded with PBST1 and 150 μ l of substrate solution (1 mg/ml of *p*-nitrophenyl phosphate in 0.1 M $\text{Na}_2\text{CO}_3/\text{NaHCO}_3$ /1 mM MgCl_2 , pH 9.8) were added and the plates were incubated at 37 °C for 30 min. The A_{405} was recorded with a Titertek spectrophotometer (Flow Laboratories) after the development of the colour was stopped by adding 50 μ l of 4.0 M NaOH.

Surface plasmon resonance (SPR) measurements

Specific-interaction analyses of arcelin-1 with fetuin, asialofetuin and thyroglobulin were performed by SPR using the biosensor *BIACore* (Pharmacia Biosensor AB). Arcelin-1, used at concentrations ranging from 0.25 up to 1.50 mg/ml in HBS, pH 7.4, was injected for 5 min on to the glycoprotein-bound surface of the sensor chip at a flow rate of 5 μ l/min. The change in the SPR response [in resonance units (RU)] was monitored at 25 °C for 9.30 min. The same glycoprotein sensor chip surface was used repeatedly after removing the remaining immobilized arcelin by two successive washes with 10 mM HCl and 10 mM NaOH for 2 min each. Asialofetuin contains six oligosaccharide chains, namely three disaccharides (T antigen) O-linked to Thr or Ser residues, and three complex glycans, N-linked to Asn residues [22,23]. In fetuin, the exposed Gal residues of both O-linked and N-linked saccharides are masked by sialic acid residues.

Interaction between the glycan chains of arcelin-1 and lectins of various specificities [mannose-specific lectins from the snow-drop (GNA), *Listera ovata* (twayblade) (LOA), the vetch *Lathyrus ochrus* (LoLI), pea (PsA)] or the ribosome-inactivating protein (RIP) of type 2 from elderberry (SNAI) were similarly measured by SRP. The lectins and type-2 RIP, used at a constant concentration of 100 μ g/ml in HBS, pH 7.4, were injected for 5 min on to the arcelin-bound surface of the sensor chip at a flow rate of 5 μ l/min. The change in the SPR response (in RU) was monitored at 25 °C for 9.30 min.

For immobilization, both arcelin-1 and glycoproteins were used at a concentration of 1 mg/ml in 5 mM sodium acetate buffer, pH 4.0. According to the change in SPR response, expressed in RU, as a result of the immobilization of arcelin-1 on the carboxymethylated dextran layer covering the sensor chip, an estimated surface concentration of 10 ng/mm² of dextran was obtained for the immobilized proteins.

Experiments with Caco-2 cells

The Caco-2 cells were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 1% (v/v) non-essential amino acids, 50 μ g of gentamycin, 10 mM NaHCO_3 , 25 mM Hepes and 20% (v/v) fetal-calf serum and cultured at 37 °C in a humidified atmosphere of CO_2 /air (1:19). The cells were seeded at 4×10^4 cells/cm² on six-well tissue-culture plates or in a 175 cm² tissue-culture flask containing respectively 2 or 35 ml of culture medium. The culture medium was changed three times a week [24].

Differentiated Caco-2 cells used for binding experiments with FITC-conjugated PHA-E, PHA-L and arcelin-1 were grown on glass coverslips on six-well tissue-culture plates. Cells were washed three times with PBS and incubated with 100 μ g/ml of FITC-PHA-E, FITC-PHA-L or FITC-arcelin-1 in DMEM for 10 min at 37 °C, washed three times with PBS, fixed with 4%

(w/v) paraformaldehyde in PBS for 15 min, and washed with distilled water. The glass coverslips were covered with fluorescent mounting solution (Gur BDH Ltd., Poole, U.K.) and examined using a Zeiss fluorescence microscope with an excitatory wavelength of 470–490 nm and an emission wavelength of 515 nm.

Changes in the relative incorporation of the radioactive precursors ($[^{14}\text{C}]$ thymidine, $[^{14}\text{C}]$ uridine, $[^3\text{H}]$ methionine and $[^3\text{H}]$ glucosamine) were measured after exposure of Caco-2 cells for 44 h to graded amounts of PHA-E, PHA-L and arcelin-1. The incorporated radioactivity was calculated as radioactivity (c.p.m.)/mg of DNA [25] and expressed as the relative incorporation, i.e. the factor by which the incorporation was increased or decreased compared with cell cultures that were not exposed to lectins or arcelin-1. Experiments were performed in quadruplicate, and the results are expressed as the mean relative incorporation $\pm$ S.D. In the relevant graphs the hatched area represents the mean relative incorporation $\pm$ S.D. for cell cultures not exposed to lectins or arcelin-1.

RESULTS

Isolation and characterization of arcelin-1

Our two-step procedure based on ion-exchange chromatography and gel chromatofocusing gave pure arcelin-1 with a yield of 50 mg of protein/10 g of seeds. The peak corresponding to arcelin-1 was eluted at pH 6.4 from the chromatofocusing column (Figure 1) and was readily separated from α -AI, PHA-E and PHA-L proteins, which appeared at pH 4.6 and beyond pH 6.0 respectively. The pooled arcelin-1 fractions gave a single protein peak from a second chromatofocusing run.

The protein migrated as a single protein band on SDS/PAGE with an apparent molecular mass of 31 kDa (Figure 2A). Gel-filtration analysis on a calibrated Superdex 75 column gave an apparent molecular mass of 59 kDa and an a value of 3.38 nm (33.8 Å) for native arcelin-1. The Guinier plot (Figure 3), computed from the SAXS measurements, indicated that the arcelin-1 solutions were monodisperse and provided the R_g of the protein [2.88 ± 0.02 nm (28.8 ± 0.2 Å)]. The extrapolated intensities, $I(0)$, were brought to absolute scale from the SAXS data of a solution of PGK of known concentration. It allowed us to calculate the molecular mass of the scattering species to be 60.4 ± 0.6 kDa. These experiments indicate that arcelin-1 is built from the non-covalent association of two identical monomers. The protein dimer was insensitive to the presence of 5 M urea, whereas a partial dissociation into monomeric species could be

Figure 1 Isolation of arcelin-1 by chromatofocusing on a PBE 94 column in the pH range between pH 7.35 and pH 5.10

Arcelin-1 was isolated as a symmetrical peak (bounded by the rectangle) eluted at about pH 6.40.

Figure 2 SDS/PAGE (A) and IEF (B) of purified arcelin-1

(A) SDS/PAGE of purified arcelin-1 (lanes 4 and 5) compared with PHA-E (lane 2) and PHA-L (lane 3). Molecular-mass reference proteins (lanes 1 and 6) were lysozyme (14 kDa), soybean trypsin inhibitor (20.1 kDa), carbonic anhydrase (30 kDa), ovalbumin (43 kDa), BSA (67 kDa) and phosphorylase *b* (94 kDa). A 1 μ l portion of a solution containing 2.5 mg/ml arcelin-1 was placed on the gel. (B) IEF of purified arcelin-1 (lane 2). pI reference proteins (lanes 1 and 3) were amyloglucosidase (3.50), soybean trypsin inhibitor (4.55), β -lactoglobulin A (5.20), bovine carbonic anhydrase (5.85), human carbonic anhydrase (6.55), acidic lentil lectin (8.15) and basic lentil lectin (8.65). A 1 μ l portion of a solution containing 1.0 mg/ml arcelin-1 was placed on the gel.

detected in the presence of 6 M guanidinium chloride, according to analytical-gel-filtration analysis.

A major protein band (pI 6.80) followed by a minor (pI 6.65) and a very weak (pI 6.50) protein band were detected by IEF/PAGE analysis (Figure 2B). However, N-terminal sequencing of the first 15 amino acid residues of arcelin-1 produced a sequence stretch identical with that deduced from its cDNA [8] (Figure 4), and we suggest that the heterogeneity observed on IEF/PAGE may result from C-terminal ‘ragged end’ of the polypeptide chain that is frequently observed with such proteins [26].

Arcelin-1 is a glycoprotein

The total carbohydrate content of purified arcelin-1 was found to be 10 % (w/w) and the monosaccharidic composition was 68 %

Figure 3 Small-angle-X-ray-scattering data of arcelin-1 at 5 mg/ml in 50 mM Mes/300 mM sodium citrate, pH 4.9

(A) Experimental curve; (B) Guinier plot. Note: 1 Å = 0.1 nm.

mannose, 24% *N*-glucosamine, 5% fucose and 3% xylose (Figure 5A). Although some discrepancies could result from the different susceptibility of monosaccharides to hydrolysis [27], the presence of fucose and xylose, in addition to the prevailing mannose and *N*-glucosamine monosaccharides, suggested the occurrence of both glycans of the high-mannose and complex

1	11	21	31	41
SNDASFNVET	<u>F</u> NKTNLILQG	DATVSSEHGL	LLTNVKGNEE	DSMGRAFYSA
SNDASFNVET	F*KTN			
1	11 15			
51	61	71	81	91
PIQINDRTID	NLASFST <u>N</u> FT	FRINAKNIEN	SAYGLAFALV	PVGSRPKLKG
101	111	121	131	141
RYLGLF <u>N</u> TTN	YDRDAHTVAV	VFDTVSNRIE	IDVNSIRPIA	TESCNFGHNN
151	161	171	181	191
GEKAEVRITY	DSPKNDLRVS	LLYPSSEKCK	HVSATVPLEK	EVEDWVSVGF
201	211	221	231	241 244
SATSGSKKET	TETHNVLSWS	FSSNFINFKG	KKSEERSNILL	NKIL
			+ + -+	+

Figure 4 N-terminal sequence of arcelin-1 compared with the amino acid sequence deduced from cDNA of arcelin-1

Residue 12 did not give a phenylhydantoin derivative upon sequencing and is indicated by an asterisk. Positively (+) and negatively (–) charged amino acid residues of the C-terminal end of arcelin-1 are indicated.

Figure 5 Monosaccharides (A) and glycans (B) of arcelin-1

(A) Monosaccharides of arcelin-1 eluted from a CarboPac PA1 column after acid hydrolysis with TFA and analysis by HPAEC-PAD. Fuc, GlcN, Xyl and Man correspond respectively to fucose, *N*-acetylglucosamine, xylose and mannose. (B) Glycans of the high-mannose type eluted from a CarboPac PA1 column after hydrolysis with PNGase F (upper) and PNGase A (lower) and analysis by HPAEC-PAD. Man6, Man7a, Man7b, Man8a, Man8b and Man9 correspond respectively to $(\text{Man})_6$ -(GlcNAc)₂, the two isomeric forms of $(\text{Man})_7$ -(GlcNAc)₂ and $(\text{Man})_8$ -(GlcNAc)₂, and $(\text{Man})_9$ -(GlcNAc)₂.

types, as previously reported for PHA [28,29]. PNGase F digestion of the N-glycan chains of arcelin-1 analysed by HPAEC-PAD indicated the presence of high-mannose-type glycans with mainly $(\text{Man})_6$ and $(\text{Man})_7$, and a lower amount of $(\text{Man})_8$ and $(\text{Man})_9$ (Figure 5B). This result is in agreement with the fact that plant complex glycans (carrying an α 1-3 fucose on the *N*-acetylglucosamine linked to Asn) are not cleaved by PNGase F. Digestion by PNGase A (known to be active towards these structures) gave a rather similar HPAEC-PAD profile (Figure 5B), but the identification of fucose after hydrolysis of the peak corresponding to $(\text{Man})_6$ eluted between 17 and 19 min supports the occurrence of glycans of the complex type. All these results were in agreement with binding experiments using SPR. Immobilized arcelin-1 strongly reacted with GNA and LOA1 (Figure 6), two lectins which exhibit high affinities for the high-mannose-type glycans (P. Rougé, unpublished work). Interactions with LoLI and PsA, which recognize the α 1-3, α 1-6

Figure 6 Sensorgrams showing the interaction of GNA (trace A), LoLI (trace B) and SNA-I (trace C) with surface-bound arcelin-1

Lectins were injected at a concentration of 100 $\mu\text{g/ml}$, and arrows indicate the beginning and the end of the injections. Each sensorgram is divided into an association (ascending slope) and a dissociation (descending slope) phases.

Figure 7 ELISA measurements at 405 nm of the interaction between PHA-E (●), PHA-L (▲) and arcelin-1 (■) used at concentrations ranging from 0.625 up to 50 $\mu\text{g/ml}$, and antibodies raised against the pentadecapeptide Arg¹⁰¹-Tyr-Leu-Gly-Leu-Phe-Asn-Thr-Thr-Asn-Tyr-Asp-Arg-Asp-Ala¹¹⁵

mannotriose inner core of the high-mannose-type glycans, were of much smaller magnitude. There was no interaction between arcelin-1 and SNA I, which binds the terminal α (2-6)-linked sialyl residues of the *N*-acetyl-lactosaminic-type glycans.

The amino acid sequence of arcelin-1 [8] revealed three putative N-glycosylation sites, namely Asn¹²-Lys¹³-Thr¹⁴, Asn⁶⁸-Phe⁶⁹-Thr⁷⁰ and Asn¹⁰⁷-Thr¹⁰⁸-Thr¹⁰⁹. No phenylhydantoin derivative was obtained in position 12 of the N-terminal amino acid sequence of our purified arcelin-1 preparation (see Figure 4), which strongly suggested N-glycosylation at Asn¹². Antibodies raised against the synthetic pentadecapeptide Arg¹⁰¹-Ala¹¹⁵ did not react in ELISA with arcelin-1, whereas they readily reacted with the related lectins (PHA-E, PHA-L from *P. vulgaris* and PHA from the lima bean), which do not exhibit N-glycosylation sites in the corresponding region (Figure 7). These observations supported the hypothesis that Asn¹⁰⁷ is also glycosylated. Evidence for glycosylation of the third site, Asn⁶⁸-Phe⁶⁹-Thr⁷⁰, could not be provided biochemically, but came from the X-ray-structure determination of this protein (see the Discussion).

Table 1 Haemagglutinating activities of arcelin-1, PHA-E and PHA-L towards untreated or trypsin-treated human and rabbit erythrocytes

A 4 HU portion of all proteins was used and results are expressed as haemagglutination titres corresponding to the highest twofold dilutions still giving a clear agglutination. The lowest concentrations of proteins corresponding to the highest twofold dilutions are given in parentheses.

Erythrocytes	Arcelin-1	PHA-E	PHA-L
Human O Rh +	4 (1 mg/ml)	512 (0.9 µg/ml)	4 (0.125 mg/ml)
Human O Rh +, trypsin-treated	128 (0.03 mg/ml)	65 536 (0.007 µg/ml)	64 (0.007 mg/ml)
Rabbit	16 (0.25 mg/ml)	4 096 (0.12 µg/ml)	8 (0.062 mg/ml)
Rabbit, trypsin-treated	32 (0.125 mg/ml)	524 288 (0.001 µg/ml)	32 (0.015 mg/ml)

Table 2 Inhibition of the haemagglutinating activity of arcelin-1, PHA-E and PHA-L by various monosaccharides* and glycoproteins

The inhibiting concentrations (mg/ml) correspond to the lower concentrations of glycoproteins able to completely inhibit 4 HU of arcelin-1 or lectins. All the data are the means for two separate determinations.

Glycoprotein	Arcelin-1	PHA-E	PHA-L
Fetuin	0.052	0.208	0.052
Asialofetuin	0.013	0.006	0.013
Thyroglobulin	0.05	20.052	0.104
Ovomucoid	0.833	0.833	0.833
Orosomucoid	0.208	1.667	0.416
Ovalbumin	3.333	> 10	1.667
Lactotransferrin	0.208	0.208	0.052
Apotransferrin	> 10	> 10	> 10
Serotransferrin	> 10	> 10	> 10
Mucin	> 10	> 10	> 10
Asialomucin	> 10	> 10	> 10

* The following monosaccharides and derivatives were not inhibitory at concentrations > 100 mM: arabinose, ribose, xylose, glucose, α -glucose 1-phosphate, β -glucose 1-phosphate, glucose 6-phosphate, methyl α -glucopyranoside, methyl β -glucopyranoside, *p*-nitrophenyl α -glucopyranoside, *p*-nitrophenyl β -glucopyranoside, *p*-nitrophenyl 1-thio- β -glucopyranoside, *o*-nitrophenyl β -glucopyranoside, glucuronic acid, *n*-dodecyl β -glucopyranoside, glucosamine, *N*-acetylglucosamine, glucosaminic acid, glucosamine 2-sulphate, glucosamine 3-sulphate, glucosamine-6-phosphate, *N*-acetylglucosamine 3-sulphate, *N*-acetylglucosamine 6-sulphate, *p*-nitrophenyl *N*-acetyl- α -glucosaminide, *p*-nitrophenyl *N*-acetyl- β -glucosaminide, *o*-nitrophenyl-*N*-acetyl- α -glucosamine, mannose, mannosamine, methyl α -mannopyranoside, methyl β -mannopyranoside, *N*-acetylmannosamine, mannose 6-phosphate, mannose 6-sulphate, *p*-nitrophenyl α -mannopyranoside, *p*-nitrophenyl β -mannopyranoside, rhamnose, galactose, galactosamine, *N*-acetylglactosamine, galacturonic acid, methyl α -galactopyranoside, methyl β -galactopyranoside, *p*-nitrophenyl α -galactopyranoside, *p*-nitrophenyl β -galactopyranoside, *p*-nitrophenyl 1-thio- β -galactopyranoside, *p*-nitrophenyl β -galactopyranoside, fucose, fructose, sorbose, sorbitol, *N*-acetylneuraminic acid, muramic acid, *N*-acetylmuramic acid, 2-*O*-methyl-*N*-acetylneuraminic acid, cellobiose, gentiobiose, lactose, lactulose, maltose, melibiose, sucrose, 6-*O*-methyl α -*D*-mannopyranosyl- α -*D*-mannopyranoside, trehalose, chitotriose and raffinose.

Sugar-binding specificity of arcelin-1

Arcelin-1 gave reproducible haemagglutination titres of 4 (1 mg/ml arcelin-1) and 16 (0.25 mg/ml arcelin-1) with human O Rh + and rabbit erythrocytes respectively. Higher titres of 128 (0.03 mg/ml arcelin-1) and 32 (0.125 mg/ml arcelin-1) were obtained with the corresponding trypsin-treated cells. The haemagglutinating activity of the protein is at least 100-fold weaker than that of typical lectins (Table 1). This suggested, however, the occurrence of a complex-oligosaccharide-binding

Figure 8 Binding (expressed as RU) of fetuin (◆), asialofetuin (■) and thyroglobulin (▲) to surface-bound arcelin-1

Glycoproteins were injected at concentrations ranging from 0.25 to 1.50 mg/ml. No binding was obtained with human serotransferrin.

site in arcelin-1. In line with this hypothesis, we found that glycoproteins such as fetuin, asialofetuin and thyroglobulin were potent inhibitors of the haemagglutination (Table 2) and that these glycoproteins readily interacted with arcelin-1 used at concentrations ranging from 0.25 to 1.5 mg/ml in SPR experiments (Figure 8). According to the inhibition of haemagglutination, arcelin-1, PHA-E and PHA-L [30], exhibit very different affinities towards the same glycoproteins, e.g. fetuin, orosomucoid and lactotransferrin. By contrast, there was no inhibition, by simple sugars, sugar derivatives and short oligosaccharides, of the haemagglutination caused by arcelin-1.

These different affinities towards complex glycans may explain the intense fluorescent staining of the surface of the Caco-2 cell monolayer which occurred with PHA-E and PHA-L, in contrast with the very weak staining observed with arcelin-1 (Figure 9). Considerable differences in the relative incorporation of [14 C]-thymidine, [14 C]uridine, [3 H]glucosamine and [3 H]methionine were observed when differentiated Caco-2 cells were incubated for 48 h in the presence of PHA-E, PHA-L or arcelin-1 (Figure 10). Arcelin-1 had little or no effect on the stimulation of cellular metabolism, in contrast with the effect observed with PHA-E or PHA-L. Preincubation of the lectins with fetuin (1 mg/ml) was shown to abolish their stimulatory effect. Neither the lectins nor arcelin-1 appeared to be capable of interfering with RNA synthesis. These experiments suggest that the oligosaccharides exposed by the brush-border membrane are involved in the different extent by which these proteins recognize and stimulate differentiated Caco-2 cells.

Arcelin-1 resistance to proteolysis

Arcelin-1 and the two lectins present in kidney-bean seeds, PHA-E and PHA-L, were similarly highly resistant to trypsin ($\approx 15\%$ digested after 5 h of incubation) and chymotrypsin ($\approx 10\%$ digested after 5 h of incubation) and, to a lesser extent, to pepsin ($\approx 30\%$ digested after 5 h of incubation). Under the same conditions, standard proteins (casein and BSA) were completely digested after 30 min of incubation with proteinases.

DISCUSSION

The reliability of the biochemical properties of arcelin-1 from natural sources required no contamination by the closely related PHA-E, PHA-L and α -AI proteins. This requirement was met

Figure 9 Fluorescence labelling of differentiated Caco-2 cells with FITC-conjugated PHA-E (A), PHA-L (B) or arcelin-1 (C)

according to the following criteria: (i) even when using large amounts of arcelin-1, no reaction was observed with antibodies raised against the synthetic pentadecapeptide Arg¹⁰¹–Ala¹¹⁵ which readily reacted with both PHA-E and PHA-L, (ii) the sugar-binding activity exhibited by arcelin-1 towards complex glycans significantly differs from those of PHA-E and PHA-L, and (iii) no inhibition of the pig pancreatic α -amylase activity could be detected, even though the K_i value of α -AI is in the picomolar range (result not shown).

Although one can assume that some arcelin-1 was lost during purification, the yield of 50 mg of arcelin-1 for 10 g of seeds is surprisingly low for the selected RAZ-2 arcelin-1-containing line developed at CIAT [31]. Indeed, Osborn et al. [8] reported an approximate concentration of 10% (w/w) of arcelin-1 in seed of the SARC1-7 line, whereas Minney et al. [9] showed that arcelin-1, expressed as proportion of total proteins, is close to 55% in

the seed of variety (CIAT accession) no. G12882. These discrepancies could, however, be explained by assuming that aggregated arcelin molecules occurring in the protein bodies of seeds were not solubilized in our mild protein-extraction procedure. The solutions of native arcelin-1 purified in this way were monodisperse according to SAXS experiments. SAXS and biochemical experiments revealed the dimeric organization of the protein (molecular mass 60.4 kDa) and the absence of tetrameric species for arcelin-1, in contrast with previous reports [32]. This dimer, of globular molecular shape according to its R_g [2.88 nm (28.8 Å)], was found to be very stable. This result pointed out the significant difference between arcelin-1 and arcelin-5, for which the concomitant occurrence of monomers and dimers has been recently reported [33].

The apparent molecular mass of the arcelin-1 monomer (31 kDa) is lower than that previously reported for the monomeric forms of the arcelin variants 1 (36.1 kDa), 2 (34.3 kDa), 3 (36.9 kDa) and 4 (36.9 kDa) isolated from various *P. vulgaris* accessions [32]. This discrepancy does not result from a proteolytic cleavage of some 45 residues of the polypeptide chain, since preliminary results from X-ray-structure determination of this protein show electron density to at least residue 226. However, a differential processing of the last C-terminal residues, shown to occur during the post-translational modifications of lectins into the ripening seeds [26], may account for the two minor bands of pI 6.65 and 6.50 observed on IEF/PAGE. Several possible proteolytic cleavage sites identified at the C-terminus of arcelin-1 could indeed lead to protein species of lower pI according to the calculations performed with the GCG (University of Wisconsin Genetics Computer Group) software [34].

The oligosaccharidic chains, which represent 10% (w/w) of each monomer, belong to the complex- and to the high-mannose-type N-glycans. The latter mainly contain (Man)₇ and (Man)₆ and, to a smaller extent, (Man)₈ and (Man)₉. The three putative N-glycosylation sites are glycosylated. The presence of glycan chains at Asn¹² and Asn¹⁰⁷ was demonstrated biochemically, but glycosylation of Asn⁶⁸ could only be assessed from preliminary X-ray-structure refinement which clearly indicated continuous electron density extending away from this asparagine side chain. Assuming two high-mannose and one complex glycan chains ($\approx$ 4.25 kDa), a polypeptide chain of arcelin-1 consisting of 244 residues ($\approx$ 26.9 kDa) would have a molecular mass of 31.2 kDa. This value is in good agreement with the SAXS and SDS/PAGE molecular-mass determinations. Moreover, the monosaccharidic composition and the glycan analysis suggest that the glycan moiety of arcelin-1 is mainly composed of high-mannose type oligosaccharides. This hypothesis would also be in line with the BIAcore experiments, which indicated that GNA binds to a greater extent to arcelin-1 than to PHA (one high-mannose plus one complex glycan chains) (results not shown). However, the small amounts of fucose identified after PNGase A digestion hamper any attempt to draw any definite conclusion about the ratio of glycans of both types in arcelin-1.

α -AI forms a 1:2 complex with pig pancreatic α -amylase (V. Le Berre-Anton, C. Bompard-Gilles, F. Payan and P. Rougé, unpublished work). The X-ray-structure determination of this complex revealed the molecular mechanism of inhibition of the enzyme [35] which accounts for the protection of kidney-bean seeds against predation by insect pests. An other mechanism must be invoked for arcelin-1, which does not bind to α -amylase.

It was postulated that arcelin may induce bruchid-larvae starvation [9]. However, the resistance of arcelin-1 to proteolytic enzymes should not account for the entire toxicity of this protein, since the protein dimer was as insensitive to proteolysis as PHA-

Figure 10 Incorporation of precursors by differentiated Caco-2 cells after incubation with PHA-isolectins or arcelin-1

Experiments were performed in quadruplicate and the results are expressed as the mean relative incorporation $\pm$ S.D. In all curves, the hatched area represents the mean relative incorporation $\pm$ S.D. of cell cultures not exposed to lectins or arcelin-1. Symbols used: ▲, PHA-L; ■, PHA-E; ○, arcelin-1.

E or PHA-L, which are apparently devoid of any toxicity towards the larvae [36]. The finding of a specific carbohydrate-binding activity of arcelin-1 significantly different from that exhibited by PHA-E should now be considered with respect to toxicity. Indeed, arcelin-1 binds to the complex glycans carried by fetuin, asialofetuin and thyroglobulin and, accordingly, displays haemagglutinating activity towards Pronase-treated erythrocytes. In addition, the complex glycans of the Caco-2 cells, which are of particular interest for *in vitro* studies of the binding of these proteins to intestinal cells, were recognized to quite different extents by PHA and arcelin-1. Finally, the type of glycosylation in arcelin-1, responsible for its binding to mannose-specific lectins such as GNA from snowdrop or LoLI from *Lathyrus ochrus*, may account for its toxicity, since mannose-specific gut lectins were shown to occur in various insects [37].

The specificity in complex-glycan-binding properties of arcelin-1 likely arises from what remains of the extended carbohydrate-binding site on this protein. At the amino acid sequence level, PHA, arcelin-1 and α -AII markedly differ in the canonical monosaccharide-binding-site structure, made of four loop regions in legume lectins [38]. One loop (corresponding to residues 125A–135A of LoLI) and two loops (corresponding to residues 99A–113A and residues 125A–135A of LoLI) are deleted in arcelin-1 and α -AII respectively [39]. As a consequence, these two proteins lack an asparagine residue (corresponding to Asn¹²⁵ or Asn¹²⁸ of LoLI and PHA respectively) which is one of the seven amino acids involved in sugar binding in the single- [40] and in the two-chain [41] legume lectins shown to play a key role in the sugar-binding activity [42,43]. Its deletion in arcelin-1 likely explains why, according to hapten inhibition of haem-

agglutination, the protein is devoid of any monosaccharide-binding activity.

A better insight into the structure–function relationship of arcelin should be provided by a knowledge of the three-dimensional structure of this protein. This is a prerequisite for describing, at the molecular level, dimeric organization, conformation of the glycan chains and topology of the extended carbohydrate binding-site which is defined by a few surface-exposed residues in close proximity to the monosaccharide-binding site of typical lectins [44,45]. The detailed comparison, at the atomic level, of the PHA-L, α -amylase inhibitor and arcelin-1 three-dimensional structures may also contribute to a better understanding of the insecticidal role of arcelin-1.

We are indebted to Dr. C. Cardona (CIAT, Cali, Colombia) for providing us with the kidney-bean seeds (*P. vulgaris* L. cv. RAZ-2) used for the isolation and purification of arcelin-1. We also thank Dr. W. Peumans (Katholieke Universiteit Leuven, Leuven, Belgium) for his gift of GNA and SNAI, and Dr. H. Mazarguil (Institut de Pharmacologie et Biologie Structurale, UPR CNRS no. 9062, Toulouse, France) for the synthesis of oligopeptides. We are also grateful to Professor Joel Janin (CNRS, Gif-sur-Yvette, France) for providing us with pure PGK and to Dr. Patrice Vachette (Laboratoire pour l'Utilisation du Rayonnement Electromagnétique, Orsay, France) for excellent data-collection facilities. This work was supported by CNRS and the Conseil Régional de Midi-Pyrénées.

REFERENCES

- Chrispeels, M. J. and Raikhel, N. V. (1991) *Plant Cell* **3**, 1–9
- Cardona, C., Posso, C. E., Kornegay, J., Valor, J. and Serrano, M. (1989) *J. Econ. Entomol.* **82**, 310–315
- Cardona, C., Kornegay, J., Posso, C. E., Morales, F. and Ramirez, H. (1990) *Entomol. Exp. Appl.* **56**, 197–206

- 4 Osborn, T. C., Burow, M. and Bliss, F. A. (1988) *Plant Physiol.* **86**, 399–405
- 5 John, M. E. and Long, C. M. (1990) *Gene* **86**, 171–176
- 6 Mirkov, T. E., Wahlstrom, J., Hagiwara, K., Finardi-Filho, F., Kjemtrup, S. and Chrispeels, M. J. (1994) *Plant Mol. Biol.* **26**, 1103–1113
- 7 Goossens, A., Geremia, R., Bauw, G., Van Montagu, M. and Angenon, G. (1994) *Eur. J. Biochem.* **225**, 787–795
- 8 Osborn, T. C., Alexander, D. C., Sun, S. S. M., Cardona, C. and Bliss, F. A. (1988) *Science* **240**, 207–210
- 9 Minney, B. H. P., Gatehouse, A. M. R., Dobie, P., Dendy, J., Cardona, C. and Gatehouse, J. A. (1990) *J. Insect Physiol.* **36**, 757–767
- 10 Schade, R. E., Schroeder, H. E., Pueyo, J. J., Tabe, L. M., Murdock, L. L., Higgins, T. J. V. and Chrispeels, M. J. (1994) *Bio/Technology* **12**, 793–796
- 11 Schroeder, H. E., Gollasch, S., Moore, A., Tabe, L. M., Craig, S., Hardie, D. C., Chrispeels, M. J., Spencer, D. and Higgins, T. J. V. (1995) *Plant Physiol.* **107**, 1233–1239
- 12 Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. and Smith, F. (1956) *Anal. Chem.* **28**, 350–356
- 13 Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. and Klenk, D. C. (1985) *Anal. Biochem.* **150**, 76–85
- 14 Goa, J. (1953) *Scand. J. Clin. Lab. Invest.* **5**, 218–222
- 15 Ladd, J. N. and Butler, J. H. A. (1972) *Soil Biol. Biochem.* **4**, 19–30
- 16 Depautex, C., Desvignes, C., Leboucher, P., Lemonnier, M., Dagneaux, D., Benoit, J. P. and Vachette, P. (1987) in *LURE Annual Report*, pp. 75, Université Paris–Sud, Paris
- 17 Bordas, J., Koch, M. H. J., Clout, P. N., Dorrington, E., Boulton, C. and Gabriel, A. (1980) *J. Phys. E: Sci. Instrum.* **13**, 938–944
- 18 Guinier, A. and Fournet, G. (1955) *Small Angle X-Ray Scattering*, Wiley, New York
- 19 Debray, H. and Rougé, P. (1984) *FEBS Lett.* **176**, 120–124
- 20 Atherton, E., Clive, D. L. J. and Sheppard, R. C. (1975) *J. Am. Chem. Soc.* **97**, 6584–6585
- 21 Lueken, K., Mazarguil, H. and Rougé, P. (1988) *Immunol. Lett.* **19**, 309–312
- 22 Shinohara, Y., Kim, F., Shimizu, M., Goto, M., Tosu, M. and Hasegawa, Y. (1994) *Eur. J. Biochem.* **223**, 189–194
- 23 Dill, K. and Olson, J. D. (1995) *Glycoconj. J.* **12**, 660–663
- 24 Koninkx, J. F. J. G., Hendriks, H. G. C. J. M., Van Rossum, J. M. A., Van den Ingh, T. S. G. A. M. and Mouwen, J. M. V. M. (1992) *Gastroenterol.* **102**, 1516–1523
- 25 Labarca, C. and Paigen, K. (1980) *Anal. Biochem.* **102**, 344–352
- 26 Young, N. M., Watson, D. C., Yaguchi, M., Adar, R., Arango, R., Rodriguez-Arango, E., Sharon, N., Blay, P. S. K. and Thibault, P. (1995) *J. Biol. Chem.* **270**, 2563–2570
- 27 Manzi, A. E. and Varki, A. (1993) in *Glycobiology: A Practical Approach* (Fukuda, M. and Kobata, A., eds.), pp. 27–77, IRL Press, Oxford.
- 28 Rayon, C., Gomord, V., Faye, L. and Lerouge, P. (1996) *Plant Physiol. Biochem.* **34**, 273–281
- 29 Sturm, A., Bergwerff, A. A. and Vliegthart, F. G. (1992) *Eur. J. Biochem.* **204**, 313–316
- 30 Goldstein, I. J. and Poretz, R. D. (1986) in *The Lectins: Properties, Functions, and Applications in Biology and Medicine* (Liener, I. E., Sharon, N. and Goldstein, I. J., eds.), pp. 33–247, Academic Press, New York
- 31 Cardona, C., Dick, K., Posso, C. E., Ampofo, K. and Nadhy, S. M. (1992) *Trop. Pest. Manage.* **38**, 173–175
- 32 Hartweck, L. M., Vogelzang, R. D. and Osborn, T. C. (1991) *Plant Physiol.* **97**, 204–211
- 33 Hamelryck, T. W., Poortmans, F., Goossens, A., Angenon, G., Van Montagu, M., Wyns, L. and Loris, R. (1996) *J. Biol. Chem.* **271**, 32796–32802
- 34 Devereux, J., Haerberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395
- 35 Bompard-Gilles, C., Rousseau, P., Rougé, P. and Payan, F. (1996) *Structure* **4**, 1441–1452
- 36 Huesing, J. E., Shade, R. E., Chrispeels, M. J. and Murdock, L. L. (1991) *Plant Physiol.* **96**, 993–996
- 37 Ingram, G. A. and Molyneux, D. H. (1991) in *Lectin Reviews*, vol. 1 (Kilpatrick, D. C., Van Driessche, E. and Bøg-Hansen, T. C., eds.), pp. 103–127, Sigma Chemical Co., St. Louis
- 38 Rougé, P., Cambillau, C. and Bourne, Y. (1991) in *Lectin Reviews*, vol. 1 (Kilpatrick, D. C., Van Driessche, E. and Bøg-Hansen, T. C., eds.), pp. 143–159, Sigma Chemicals Co., St. Louis
- 39 Rougé, P., Barre, A., Causse, H., Chatelain, C. and Porthé, G. (1993) *Biochem. Syst. Ecol.* **21**, 695–703
- 40 Derewenda, Z., Yariv, J., Helliwell, J. R., Kalb (Gilboa), A. J., Dodson, E. J., Papiz, M. Z., Wan, T. and Campbell, J. (1989) *EMBO J.* **8**, 2189–2193
- 41 Bourne, Y., Roussel, A., Frey, M., Rougé, P., Fontecilla-Camps, J. C. and Cambillau, C. (1990b) *Proteins: Structure, Function and Genetics* **8**, 365–376
- 42 Van Eijsden, R. R., Hoedemaeker, F. J., Diaz, C. L., Lugtenberg, B. J. J., De Pater, B. S. and Kijne, J. W. (1992) *Plant Mol. Biol.* **20**, 1049–1058
- 43 Mirkov, T. E. and Chrispeels, M. J. (1993) *Glycobiology* **3**, 581–587
- 44 Bourne, Y., Mazurier, J., Legrand, D., Rougé, P., Montreuil, J., Spik, G. and Cambillau, C. (1994) *Structure* **2**, 209–219
- 45 Drickamer, K. (1995) *Nat. Struct. Biol.* **2**, 437–439