the FEBS Journal

cDNA cloning and 1.75 Å crystal structure determination of PPL2, an endochitinase and *N*-acetylglucosamine-binding hemagglutinin from *Parkia platycephala* seeds

Benildo S. Cavada[1], Frederico B. B. Moreno[2], Bruno A. M. da Rocha[1,3], Walter F. de Azevedo Jr[4], Rolando E. R. Castellón[1], Georg V. Goersch[1], Celso S. Nagano[5], Emmanuel P. de Souza[1], Kyria S. Nascimento[1], Gandhi Radis-Baptista[1], Plínio Delatorre[3], Yves Leroy[6], Marcos H. Toyama[7], Vicente P. T. Pinto[8], Alexandre H. Sampaio[9], Domingo Barettino[5], Henri Debray[6], Juan J. Calvete[5] and Libia Sanz[5]

1 BioMol-Laboratory, Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil
2 Departamento de Física, Universidade Estadual Paulista, UNESP, São José do Rio Preto, São Paulo, Brazil
3 Departamento de Ciências Físicas e Biológicas, Universidade Regional do Cariri, Fortaleza, Ceará, Brazil
4 Faculdade de Biociências, Centro de Pesquisas em Biologia Molecular e Funcional, PUCRS, Porto Alegre, Rio Grande do Sul, Brazil
5 Instituto de Biomedicina de Valencia, CSIC, Spain
6 Laboratoire de Chimie Biologique et Unité Mixte de Recherche No. 8576 du CNRS, Université des Sciences et Technologies de Lille, France
7 Departamento de Bioquímica, Instituto de Biologia, Universidade Estadual de Campinas (UNICAMP), Campinas, SP, Brazil
8 Faculdade de Medicina, Universidade Federal do Ceará, Sobral, Brazil
9 Laboratorio de Bioquímica Marinha, Departamento de Engenharia de Pesca, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil

Keywords
endochitinase; glycosyl hydrolase family 18; Mimosoideae; *Parkia platycephala*; X-ray crystal structure

Correspondence
B. S. Cavada, BioMol-Laboratory, Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil
Fax/Tel: +55 8540089818
E-mail: bscavada@ufc.br
H. Debray, Laboratoire de Chimie Biologique et Unité Mixte de Recherche du CNRS N°8576, Université des Sciences et Technologies de Lille, bâtiment C-9, 59655 Villeneuve D'Ascq Cedex
Fax: +33 320436555
Tel: +33 320410108
E-mail: henri.debray@univ-lille1.fr

(Received 22 May 2006, revised 26 June 2006, accepted 28 June 2006)

doi:10.1111/j.1742-4658.2006.05400.x

Parkia platycephala lectin 2 was purified from *Parkia platycephala* (Leguminosae, Mimosoideae) seeds by affinity chromatography and RP-HPLC. Equilibrium sedimentation and MS showed that *Parkia platycephala* lectin 2 is a nonglycosylated monomeric protein of molecular mass 29 407 ± 15 Da, which contains six cysteine residues engaged in the formation of three intramolecular disulfide bonds. *Parkia platycephala* lectin 2 agglutinated rabbit erythrocytes, and this activity was specifically inhibited by *N*-acetylglucosamine. In addition, *Parkia platycephala* lectin 2 hydrolyzed β(1–4) glycosidic bonds linking 2-acetoamido-2-deoxy-β-D-glucopyranose units in chitin. The full-length amino acid sequence of *Parkia platycephala* lectin 2, determined by N-terminal sequencing and cDNA cloning, and its three-dimensional structure, established by X-ray crystallography at 1.75 Å resolution, showed that *Parkia platycephala* lectin 2 is homologous to endochitinases of the glycosyl hydrolase family 18, which share the $(\beta\alpha)_8$ barrel topology harboring the catalytic residues Asp125, Glu127, and Tyr182.

Abbreviations
CTAB, cetyl triethylammonium bromide; GlcNac, N-acetyl-D-glucosamine; GSP, gene-specific forward primer; HPAEC-PAD, high-pH anion exchange chromatography with pulsed amperometric detection; PE, pyridylethylated; PPL1, *Parkia platycephala* lectin 1; PPL2, *Parkia platycephala* lectin 2; PTC, phenylisothiocyanate; PTH, phenylthiohydantoin.

Lectins comprise a heterogeneous class of (glyco)proteins that possess one noncatalytic domain that binds carbohydrates in a specific and reversible manner without altering their covalent structure [1]. Lectins decipher the glycocodes encoded in the structure of glycans in processes such as cell communication, host defense, fertilization, development, parasitic infection, tumor metastasis, and plant defense against herbivores and pathogens [2]. Mechanisms for sugar recognition have evolved independently in a restricted number of protein folds (e.g. jelly roll domain, C-type lectin fold, β-propeller, β-trefoil motif, β-prism I and II domains, Ig domains, β-sandwich, mixed αβ structure, and hevein domain) [1,3] (for a complete catalog of carbohydrate-binding protein domains, please consult the 3D Lectin Database at http://www.cermav.cnrs.fr/lectines). In plants, most of the currently known lectins can be placed in seven families of structurally and evolutionarily related proteins [1]. The seed lectins of leguminous plants constitute the largest and most thoroughly studied lectin family. These lectins have represented paradigms for establishing the structural basis [4–9] and thermodynamics [10–13] of selective sugar recognition.

Most studies on lectins from Leguminosae involve members of the Papilionoideae subfamily, whereas investigations on lectins of the other two subfamilies, Caesalpinoideae and Mimosoideae, are scarce. Indeed, to date, the only lectins from the Mimosoideae that have been functionally and structurally characterized are those from seeds of species of the genus *Parkia*, including *Parkia speciosa* [14], *Parkia javanica* [15], *Parkia discolor* [16] and the glucose/mannose-specific lectin from *Parkia platycephala* seeds [17–21]. *Parkia* (Leguminosae, Mimosoideae), regarded as the most primitive group of leguminous plants [22], is a pantropical genus of trees comprising about 30 species found in the neotropics from Honduras to south-eastern Brazil, West Africa, the northern part of Malaysia and the south of Thailand. *Parkia platycephala* is an important forage tree growing in parts of north-eastern Brazil. The seed lectin from *Parkia platycephala* is a 47.9-kDa single-chain nonglycosylated mosaic protein composed of three tandemly arranged jacalin-related β-prism domains [19,20].

The sugar-binding specificity of *Parkia platycephala* lectin towards mannose, an abundant building block of surface-exposed glycoconjugates of viruses, bacteria, and fungi, suggests a role for the *Parkia platycephala* lectin in defense against plant pathogens [1]. Moreover, the *Parkia platycephala* lectin also shows sequence similarity with stress-upregulated and pathogen-upregulated defense genes of a number of different plants, suggesting a common ancestry for jacalin-related lectins and inducible defense proteins [19]. In addition to using lectins, whose precise role in plant defense remains to be determined [23,24, and references cited], plants defend themselves against pathogens (i.e. fungi) secreting pathogenesis-related enzymes, such as xylanases and chitinases, which degrade the pathogen's cell wall [25–27]. In a previous article we have reported the presence of an endochitinase in *Parkia platycephala* seeds [28]. Now, we have determined its complete amino acid sequence by a combination of Edman degradation and cDNA cloning, and report its biochemical characterization and the determination of its crystal structure. Our results show that this protein, termed *Parkia platycephala* lectin 2 (PPL2), is homologous to endochitinases of the glycosyl hydrolase family 18 that exhibit rabbit erythrocyte-agglutinating, *N*-acetylglucosamine-binding and chitin-hydrolyzing activities.

Results and Discusion

PPL2, a nonglycosylated and monomeric GlcNAc-binding hemagglutinin

PPL2 was purified from *Parkia platycephala* seeds by affinity chromatography on either Red-Sepharose (Fig. 1A) or chitin-Sepharose. The protein agglutinated trypsin-treated rabbit erythrocytes (128 hemagglutinating units mg^{-1}), and this activity was abolished by 19 mM *N*-acetyl-D-glucosamine (GlcNac). Other sugars, such as glucose, mannose, galactose, fucose and *N*-acetyl-D-galactosamine, displayed only partial hemagglutination inhibitory activity at much higher concentrations (> 75 mM) than GlcNac. Moreover, the glycoproteins bovine thyroglobulin, ovine submaxillary mucin, bovine fetuin and bovine asialofetuin were devoid of hemagglutination inhibitory activity. Bovine thyroglobulin contains nine complex glycosylation sites and four high-mannose oligosaccharides [29]. Ovine submaxillary mucin is a glycoprotein bearing a high density of O-linked oligosaccharides expressing sialyl Tn antigens and sialyl core 3 sequences [30]. Bovine fetuin contains three N-linked glycosylation sites occupied with trisialylated, tetrasialylated or pentasialylated trianntennary structures, and three monosialylated or disialylated O-linked saccharides [31–33]. We thus concluded that PPL2 represented an *N*-acetylglucosamine-binding hemagglutinin.

The apparent molecular masses of both native and reduced PPL2 determined by SDS/PAGE were 30 kDa (Fig. 1A, insert). The molecular mass of native PPL2, measured by MALDI-TOF MS, was 29 407 ± 15 Da (Fig. 1A). This value was not altered upon incubation of the denatured, but nonreduced,

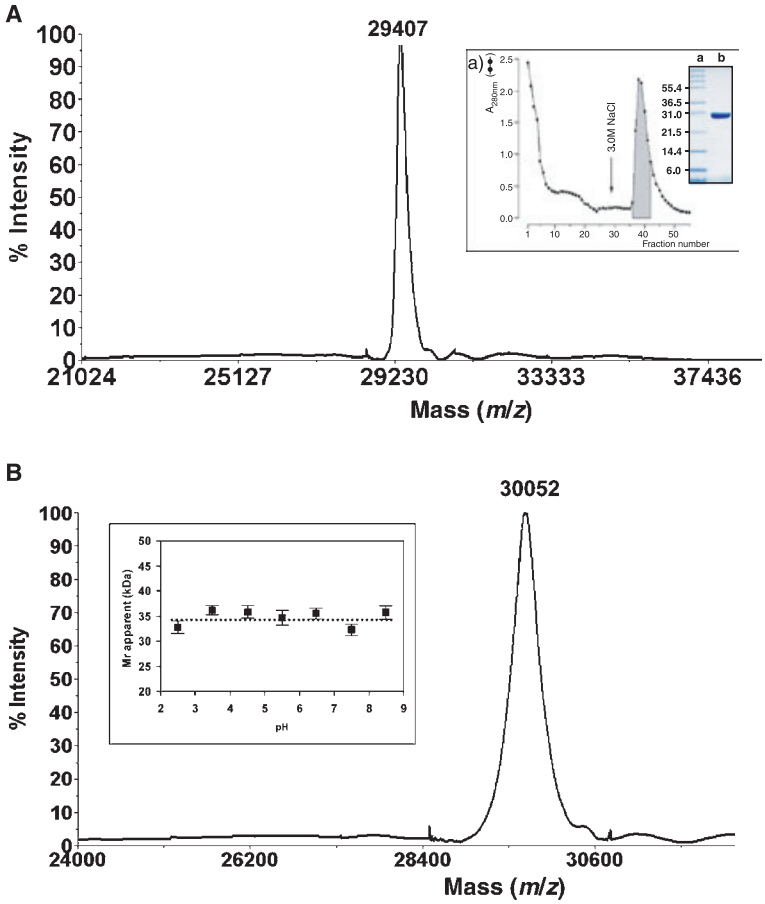

Fig. 1. Purification and molecular mass determination of PPL2. (A) MALDI-TOF mass determination of native PPL2 purified by affinity chromatography as illustrated in the insert. Insert: the fraction of a *Parkia platycephala* seed homogenate precipated with 60% saturation ammonium sulfate was resuspended in 50 mM Tris, pH 7.0, containing 100 mM NaCl, and applied to a Red-Sepharose column. Retained material was eluted with 3 M NaCl. Fractions exhibiting hemagglutinating activity (gray area) were pooled. Right panel: SDS/PAGE of the pooled hemagglutinin termed PPL2. Lane a, molecular mass makers: glutamic dehydrogenase (55.4 kDa), lactate dehydrogenase (36.5 kDa), carbonic anhydrase (31.0 kDa), trypsin inhibitor (21.5 kDa), lysozyme (14.4 kDa), aprotinin (6.0 kDa). Lane b, reduced PPL2. (B) MALDI-TOF mass determination of reduced and pyridylethylated PPL2. Insert: apparent molecular masses of native PPL2 determined by equilibrium sedimentation analytic centrifugation in solutions with different pH values.

protein with the alkylating reagent 4-vinylpyridine. On the other hand, the same treatment after reduction of the protein with dithiothreitol changed the molecular mass of PPL2 to 30 052 ± 15 Da (Fig. 1B). The mass increment of about 645 Da indicated that PPL2 had incorporated six pyridylethyl groups. The combined data clearly showed that PPL2 contained six cysteine residues engaged in the formation of three intramolecular disulfide bonds. Amino acid compositional analysis of the purified protein (Table 1) was in agreement with this conclusion.

The estimated apparent molecular mass for PPL2 on a calibrated size-exclusion chromatographic column was 12 kDa, indicating that the protein had an anomalous elution profile. Molecular mass determinations by size-exclusion chromatography are dependent on the hydrodynamic properties of the molecule, and, in addition, interaction of the protein with the matrix may also introduce large errors into the estimated molecular mass. Thus, we carried out a more rigorous analysis of the aggregation state of PPL2 employing

Table 1. Amino acid composition [mol·(mol protein)$^{-1}$] of PPL2. Asx, aspartic acid and asparagine; Glx, glutamic acid and glutamine.

Amino acid	PPL2
Asx	34
Glx	16
Gly	22
Ser	27
His	2
Arg	5
Thr	13
Ala	20
Pro	11
Tyr	9
Val	12
Met	1
Cys	6
Ile	13
Leu	23
Phe	11
Lys	9
Trp	7
Total	241

analytic ultracentrifugation equilibrium sedimentation, a technique that is firmly based in thermodynamics and does not therefore rely on calibration or on making assumptions concerning the shape of the protein. Using this approach, the apparent molecular mass of the PPL2 lectin in solutions with pH in the range 2.5–8.5 was 34 ± 3 kDa (Fig. 1B, insert). This figure, in conjunction with the MS analyses, showed that the protein behaved as a pH-independent monomeric protein.

Carbohydrate analysis performed by GLC (data not shown) failed to show the presence of any amino or neutral monosaccharide, strongly indicating that PPL2 was a nonglycosylated protein.

PPL2 displays chitinase activity

Edman degradation analysis of reduced and pyridylethylated protein yielded the first 42 amino acid residues of PPL2: GGIVVYWGQNGGEGTLTSTCESGL YQIVNIAFLSQFGGGRRP. A BLAST analysis (http://www.ncbi.nlm.nih.gov/blast/) revealed extensive (up to approximately 75%) similarity with a large number of plant chitinase sequences deposited in the publicly accessible protein databases, such as the basic chitinase III from *Nicotiana tabacum* (P29061), an acidic chitinase from *Glycine max* (BAA77677), chitinase b from *Phytolacca americana* (Q9S9F7), chitinase from *Psophocarpus tetragonolobus* (BAA08708), chitinases from *Vitis vinifera* (CAC14014), basic chitinase from *Vigna unguiculata* (Q43684), and chitinase B from leaves of pokeweed (Q9S9F7). All of these proteins are poly [1,4-(*N*-acetyl-β-D-glucosaminide)] glycanhydrolases of the glycosyl hydrolase family 18 (EC 3.2.1.14) [34] (http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00704), whose prototype is hevamine, isolated from the rubber tree [35,36].

The possible chitinase activity of PPL2 was investigated by quantitative GC determination of the amount of GlcNac released using chitin as substrate. PPL2 released 3 μg of GlcNac·h^{-1}·(mg protein)$^{-1}$. In comparison, commercial *Streptomyces griseus* chitinase exhibited an activity of 80 μg of GlcNac·h^{-1}·(mg protein)$^{-1}$, and the GlcNac-specific agglutinins from wheat germ (WGA) and *Urtica dioica* (UDA) did not show any chitinase activity. Peracetylated GlcNac (retention time 33.60 min) was observed in the reaction mixtures containing PPL2 or *Streptomyces griseus* chitinase but not in those reaction mixtures to which WGA or UDA were added. These results demonstrated that PPL2 was indeed an active chitinase able to hydrolyze the β(1–4) glycosidic bond linking the GlcNac units of chitin. In order to determine whether PPL2 presented chitinase activity only for the nonreducing end of chitin (exochitinase activity) or also had the ability to hydrolyze internal β(1–4) glycosidic linkages (an endochitinase activity), 40 μL of the reaction mixture used for the chitinase assay were analyzed by Dionex high-pH anion exchange chromatography using a CarboPac PA-100 column. The elution times of three major analytes present in the reaction mixture (3.93, 4.84 and 5.58 min) matched those of the standard carbohydrates GlcNac, $(GlcNac)_2$ and $(GlcNac)_3$ (3.86, 4.84 and 5.58 min, respectively). This result demonstrated an endochitinase activity for PPL2. The exact mechanism of glycoside hydrolysis (e.g. with retention or not of the β-anomeric configuration of the products) remains to be established, however.

The finding that PPL2 exhibited GlcNac-dependent hemagglutination and endochitinase activities was striking but not without precedent. The acidic chitinase BjCHI1 from *Brassica juncea* showed hemagglutination ability [37]. However, BjCHI1 is a unique chitinase with two chitin-binding domains, and both chitin-binding domains are essential for agglutination [38]. On the other hand, PPL2 is a single-domain protein. Hence, PPL2 may possess at least two carbohydrate-binding sites. One of them probably corresponds to the catalytic site, whereas the other one(s) remain to be characterized.

Plant chitinases constitute a class of pathogenesis-related proteins that play an important role in defense against pathogens through degradation of chitin present in the fungal cell wall and in insect cuticles [37,39]. The first characterization of a chitinase in the Mimosoideae subtribe, an antifungal chitinase from *Leucaena leucocephala* has been reported only recently [40]. This protein belongs to the class I chitinases of the glycosyl hydrolase family 19, and is, thus, structurally unrelated to PPL2.

It is noteworthy that the seeds of *Parkia platycephala* contain two different lectins: the mannose/glucose-specific PPL1 [19,21] and the GlcNac-binding lectin with chitinase activity, PPL2, described here. The fact that mannose is an abundant building block of surface-exposed glycoconjugates of viruses, bacteria and fungi supports the view that PPL, and other mannose-recognizing lectins, play a role in plant defense against pathogens [1]. Specifically, the planar array of carbohydrate-binding sites on the rim of the toroid-shaped structure of the *Parkia platycephala* lectin dimer [21] immediately suggested a mechanism to promote multivalent interactions leading to cross-linking of carbohydrate ligands as part of the host strategy against phytopredators and pathogens. The presence of two unrelated lectins in plant seeds has

been also reported in *Canavalia ensiformis* (Leguminosae): concanavalin A, a prototypic glucose/mannose-specific legume lectin built by the jellyroll fold [1,7], and concanavalin B, which, although it shares about 40% sequence identity with plant chitinases belonging to glycosyl hydrolase family 18, has not been shown to have any chitinase activity [41]. The lack of chitinase activity of concanavalin B can be explained by differences in the loops that form the substrate-binding cleft [42].

Sequencing of cDNA and genomic DNA for PPL2

Conserved amino acid sequences from glycosyl hydrolase family 18 were used to design two degenerate primers that allowed us to PCR-amplify a specific product of approximately 500 bp (pPPL2). Its sequence was then used to design a gene-specific forward primer (GSP-PPL2) to extend the sequence analysis of the PPL2 cDNA by 3′RACE. Using the GSP-PPL2 and Qo primers, the sequence was extended in the 3′ direction by PCR walking. From these sequences (pPPL2 and 3′RACE), two specific primers (PPL2f and PPL2r) were designed that amplified a fragment of 800 bp corresponding to the stretch between the conserved N-terminal sequence ^{6}YWGQNGG^{12} and the STOP codon (Fig. 2). Using primers designed from the cDNA sequence, the PPL2 gene was amplified from genomic DNA of *Parkia platycephala* seedlings. The size of the amplified genomic DNA was identical to that of the cDNA, indicating that the PPL2 gene was devoid of introns, as observed for other class III chitinase genes [43].

The complete amino acid sequence of PPL2 determined by the combination of N-terminal sequencing and cDNA cloning contains 271 amino acid residues, including the six conserved cysteine residues of class III chitinases, and the putative catalytic residues of class III plant chitinases, which in PPL2 correspond to amino acid positions 125 (Asp) and 127 (Glu). The calculated isotope-averaged molecular mass of the PPL2 sequence is 29 490.1 Da, which is about 86 ± 15 Da greater than the molecular mass determined by MALDI-TOF MS, suggesting that the native protein may lack the C-terminal valine residue.

Overall three-dimensional structure of PPL2

Figure 3 displays the structure of PPL2. The $2F_o - F_c$ density map contoured at 1σ showed that, with the exception of a small loop between the α_4 and β_5 regions corresponding to residues from Asn144 to Lys149, the majority of the protein residues were well

```
|---------------------------- N-terminal-------------------------------------
 G   G   I   V   V   Y   W   G   Q   N   G   G   E   G   T   L   T   S   T   C      20
                     TAT TGG GGC CAG AAT GGA GGA GAA GGC ACT TTG ACA TCA ACA TGT     60
                     |PPL2f ->

--------------------------------N-terminal-------------------------------------
 E   S   G   L   Y   Q   I   V   N   I   A   F   L   S   Q   F   G   G   G   R      40
GAA AGT GGC CTC TAT CAA ATT GTT AAC ATA GCT TTT CTA TCA CAA TTT GGG GGT GGT CGA    120
------->
 R   P   Q   I   N   L   A   G   H   C   D   P   A   N   N   G   C   R   T   V      60
CGA CCT CAA ATA AAC CTC GCT GGC CAC TGT GAT CCT GCT AAT AAT GGT TGC CGC ACG GTG    180

 S   D   G   I   R   A   C   Q   R   R   G   I   K   V   M   L   S   I   G   G      80
AGC GAT GGC ATC AGA GCC TGC CAA AGA AGA GGC ATC AAG GTC ATG CTC TCC ATC GGA GGC    240

 G   A   G   S   Y   S   L   S   S   V   Q   D   A   R   S   V   A   D   Y   I     100
GGC GCC GGA AGC TAC TCG TTG TCG TCG GTC CAG GAT GCC AGA AGC GTA GCA GAT TAC ATA    300

 W   N   N   F   L   G   G   R   S   S   S   R   P   L   G   D   A   V   L   D     120
TGG AAC AAC TTT CTA GGT GGA CGA TCA TCC TCA AGA CCT CTA GGC GAC GCC GTT TTG GAT    360

 G   V   D   F   D   I   E   H   G   G   A   Y   Y   D   A   L   A   R   R   L     140
GGA GTA GAT TTT GAC ATT GAA CAT GGT GGT GCA TAC TAT GAT GCT CTT GCA AGA AGA CTA    420

 S   E   H   N   R   G   G   K   K   V   F   L   S   A   A   P   Q   C   P   F     160
TCA GAG CAT AAC CGA GGA GGT AAA AAA GTG TTC TTA TCT GCT GCA CCA CAA TGT CCT TTT    480
                                                |-GSP-PPL2->

 P   D   Q   S   L   N   K   A   L   S   T   G   L   F   D   Y   V   W   V   Q     180
CCA GAT CAG TCT CTC AAT AAG GCA CTT TCC ACT GGG CTC TTT GAT TAT GTG TGG GTT CAA    540

 F   Y   N   N   P   Q   C   E   F   N   S   G   N   P   S   N   F   R   N   S     200
TTC TAT AAC AAC CCT CAG TGT GAG TTT AAT TCT GGA AAC CCT AGC AAC TTT CGG AAC TCA    600

 W   N   K   W   T   S   S   F   N   A   K   F   Y   V   G   L   P   A   S   P     220
TGG AAC AAG TGG ACT TCA AGT TTC AAT GCC AAG TTT TAT GTT GGG CTT CCT GCC TCA CCT    660

 E   A   A   G   S   G   Y   V   P   P   Q   Q   L   I   N   Q   V   L   P   F     240
GAA GCA GCT GGG AGT GGG TAT GTG CCA CCA CAA CAA CTT ATA AAT CAA GTG TTG CCT TTT    720

 V   K   R   S   P   K   Y   G   G   V   M   L   W   D   R   F   N   D   L   K     260
GTT AAA AGA TCT CCC AAG TAT GGA GGA GTC ATG CTT TGG GAT AGG TTC AAT GAT CTG AAA    780

 T   K   Y   S   S   K   I   K   P   S   V   *                                     271
ACT AAG TAC AGC TCC AAA ATT AAG CCC AGT GTT TGA                                   816
                                    <- PPL2r -|
```

Fig. 2. cDNA and amino acid sequence of PPL2. The nucleotide and the amino acid sequences are numbered on the right side. The underlined nucleotide sequences correspond to primers used to clone and sequence the full-length PPL2. The underlined amino acid sequences 6–12 and 178–185 represent the conserved polypeptide stretches from which degenerate primers were initially designed. The N-terminal amino acid sequence determined by Edman degradation is labeled. The six conserved cysteines of class III chitinases are shadowed, and the conserved residues of the active site of family 18 of glycosyl hydrolases are boxed.

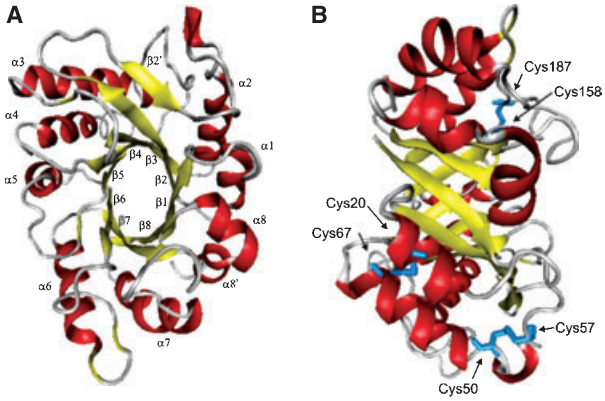

Fig. 3. Crystal structure of PPL2. (A) and (B) show two views of the $(\alpha\beta)_8$ barrel fold of PPL2. The α-helices (red) and β-strands (yellow) are labeled from 1 to 8. Disulfide bonds are depicted in blue. In (B), the active site cleft loops are located at the right face of the model.

fitted. The PPL2 model has good overall stereochemistry (Table 2), with no amino acid residues in the disallowed region of the Ramachandran plot. The PPL2 structure consists of a compact $(\beta/\alpha)_8$ barrel with dimensions of approximately $50 \times 40 \times 25$ Å, including three disulfide bonds (Cys20–Cys67, Cys50–Cys57 and Cys158–Cys187) and five *cis* peptide bonds. Two of the *cis* peptide bonds of PPL2 (Gly147–Lys148 and Lys148–Lys149) are located in a region of poor density, whereas the remaining three (Ala31–Phe32, Phe160–Pro161 and Trp253–Asp254) are well defined at the electron density. With the exception of four sulfate ions (Fig. 4), which presumably remained bound to PPL2 throughout its purification protocol, as the protein was precipitated by ammonium sulfate to separate it from pigments, no metal ions or ligands were detected. Sulfate ions were assigned according to Copley and Barton [44].

Structural comparison and analysis of conserved motifs

The overall structural features of the PPL2 model are conserved in other GH18 plant chitinases, i.e. hevamine (*Hevea brasiliensis*) (PDB code 2HVM), the

Table 2. Statistics of data collection, refinement and quality of the structure.

	Overall resolution dataset	Highest resolution dataset
Data collection		
Total number of observations	95 262	12 669
Total number of unique observations	25 805	3521
R_{merge}	0.040	0.228
Highest resolution limit (Å)	1.73	1.73
Lowest resolution limit (Å)	32.31	1.83
Completeness (%)	95.5	90.4
Multiplicity	3.7	3.6
I/σ (I)	13.1	2.4
Wavelength (Å)	1.431	
Space group	$P2_12_12_1$	
Cell parameters (Å)	$a = 55.19$, $b = 59.95$, $c = 76.70$	
Refinement		
Resolution range (Å)	1.73–32.31	
R_{factor} (%)	16.88	
R_{free} (%)	19.87	
Number of nonhydrogen atoms in protein structure	2086	
Number of sulfate ions	4	
Number of water molecules	249	
Root mean square deviations from ideal values		
Bond lengths (Å)	0.012	
Bond angles (degrees)	1.48	
Temperature factors		
Average *B*-value for whole protein chain (Å^2)	13.26	
Average *B*-value for sulfate ions (Å^2)	41.97	
Average *B*-values for water molecules (Å^2)	24.29	
Ramachandran plot		
Residues in most favored regions	195 (87.8%)	
Residues in additional allowed regions	26 (11.7%)	
Residues in generously allowed regions	1 (0.5%)	

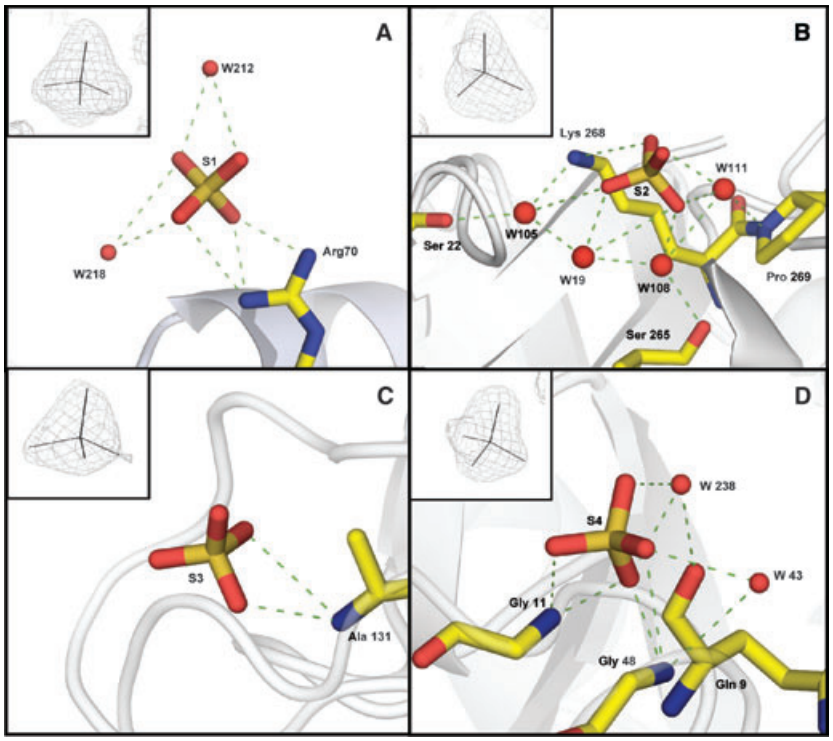

Fig. 4. Sulfate ions bound to crystallized PPL2. (A)–(D) display details of the binding of sulfate ions (S) 1–4 within the crystal structure of PPL2. In each panel, the electron density assigned to the sulfate ions is displayed in an insert. W, water molecule.

xylanase inhibitor XIP-I from *Triticum aestivum* (1TE1), and ConB (*Canavalia ensiformis*) (1CNV), with which PPL2 shares 68%, 40% and 40% sequence similarity, respectively (Fig. 4A). The three-dimensional structure of PPL2 can be superimposed onto those of hevamine, XIP-I and ConB, with root mean square deviation (r.m.s.d.) for all Cα atoms of 0.90 Å, 1.01 Å and 1.14 Å, respectively. In particular, the two consensus motifs described for the glycosyl hydrolase family 18, e.g. the presence of the absolutely conserved strands β_3 and β_4 (Fig. 4A, boxed), and the hydrogen bond network between residues Asp120 and Gly121 and Val74 (Fig. 4A,B) [33], are also conserved in PPL2. On the other hand, the largest structural divergence is associated with the active site cleft loops, which comprise the residues linking neighbor β-strands in the $(\alpha\beta)_8$ barrel. Thus, whereas with the exception of the $\beta_6\alpha_6$ loop, all the active site cleft loops of PPL2 are highly conserved in hevamine, and only few structural differences are evident when comparing the $\beta_2\alpha_2$ and $\beta_7\alpha_7$ loops from PPL2 and ConB, the active site cleft loops from XIP-I significantly depart from those of PPL2.

The PPL2 chitin-binding site

X-ray studies have suggested that enzymes of the GH18 family showing chitinase activity have conserved Asp125, Glu127 and Tyr183 amino acids (hevamine numbering) in their active sites. Their significance for catalysis is not well understood, although it has been suggested that Glu127 may act as a proton donor to the cleavable glycosidic bond, and Asp125 and Tyr183 would contribute to the stabilization of the oxazolinium intermediate [45]. In PPL2, these residues correspond to Asp125, Glu127 and Tyr182 (Figs 2 and 4A). Asp125 and Glu127 are located in the $\beta_4\alpha_4$ loop, and Tyr182 at loop $\beta_6\alpha_6$. The highly conserved, functionally relevant, structural features that are common to PPL2 and hevamine suggest that these two chitinases may share essentially the same catalytic mechanism. In addition, our data showing that PPL2 strongly binds GlcNac would support a hypothetical mechanism by which the lectin hydrolyzes a chitin polymer by cycles of anchoring, cleavage and being released from a GlcNac-binding site, and anchoring to another GlcNac-binding site. Clearly, detailed molecular and structural studies are required to investigate this.

Experimental procedures

Isolation of PPL2

Mature seeds from *Parkia platycephala* were collected in the state of Ceará (north-eastern Brazil) and ground in a coffee mill. The flour was defatted with *n*-hexane, air-dried at room

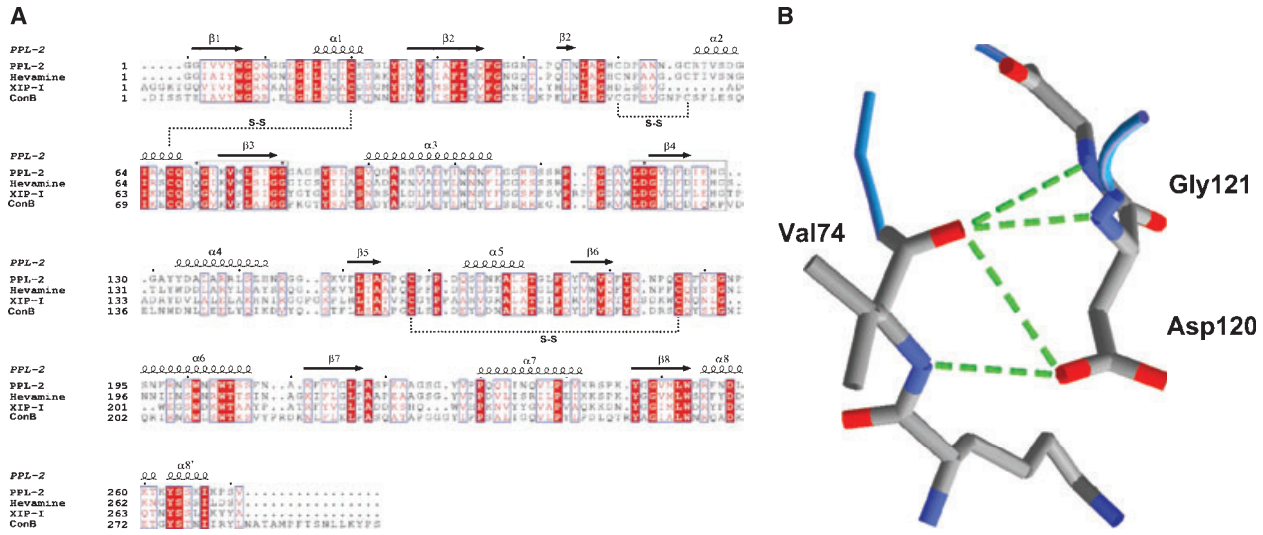

Fig. 5. Structural features of PPL2 and the GH18 family. (A) Multiple sequence alignment of PPL2, hevamine, XIP-I and ConB. Absolutely conserved residues in the four proteins are shown in white over a red background. Conservative substitutions or residues conserved in at least two proteins are depicted in pale red and boxed. Cysteine residues engaged in the formation of disulfide bonds (S–S) are conected by discontinuous lines. The secondary structure elements of PPL2 are shown on top of the sequence alignment: arrows represent β-strands and springs denote α-helices. (B) Detail of the network of hydrogen bonds between PPL2 residues Asp120, Gly121 and Val74, which represent a conserved structural motif of the GH18 family.

temperature and kept dry for further use. Soluble proteins were extracted overnight at room temperature by continuous stirring with 1 : 15 (w/v) 500 mM HCl solution, containing 150 mM NaCl. Insoluble material was separated by centrifugation (Ultracentrifuga Beckman modelo XL-1, Palo Alto, CA) at 10 000 ***g*** for 20 min at 5 °C. The supernatant was adjusted to pH 7.0 and left for 12 h at 4 °C. Precipitated pigments were removed by centrifugation (Ultracentrifuga Beckman modelo XL-1), and the supernatant was subjected to precipitation with 60% saturated ammonium sulfate. After centrifugation (Ultracentrifuga Beckman modelo XL-1), the pellet was resuspended in a small volume of 50 mM Tris, pH 7.0, containing 100 mM NaCl, dialyzed against this buffer, and subjected to affinity chromatography on a Red-Sepharose CL-4B column (26 × 1.5 cm) (Sigma-Aldrich, São Paulo, Brazil) equilibrated with the same buffer as described previously for GlcNAc-specific enzymes [46]. Unbound material was eluted by washing the column with equilibration buffer, and the retained fraction was desorbed with 3 M NaCl in buffer, dialyzed against equilibrium buffer, and assayed for hemagglutinating activity following a standard procedure with trypsin-treated rabbit red blood cells [47]. To this end, a two-fold dilution was prepared for each sugar (1 M starting concentration) solution in 0.15 M NaCl containing 5 mM $CaCl_2$ and 5 mM $MnCl_2$. Each dilution had a final volume of 0.2 mL. The purified lectin was diluted in 0.15 M NaCl to achieve 4 units of hemagglutinating activity per mL. The lowest concentration of inhibitor exhibiting agglutinating activity was termed the minimum inhibitory concentration. Aliquots of 0.2 mL of the 4 unit solution of the lectin were used for hemagglutination inhibition assay. Monosaccharides (mannose, glucose, galactose, *N*-acetylglucosamine, *N*-acetylgalactosamine, fucose) and glycoproteins (bovine thyroglobulin, ovine submaxillary mucin, bovine fetuin, and asialofetuin) were tested for hemagglutination inhibitory activity.

Purification of PPL2

The protein fraction retained in the Red-Sepharose CL-4B column was further fractionated by RP-HPLC and by chitin affinity chromatography. For RP-HPLC, 3 mg of total proteins was dissolved in 250 μL of 0.1% trifluoroacetic acid (solution A) and centrifuged (Ultracentrifuga Beckman modelo XL-1) at 4500 ***g*** for 2 min. The supernatant was applied on a μBondapack C18 analytic column (3.9 × 300 mm) (Waters, Milford, MA, USA) equilibrated in solution A, and the column was developed using the following chromatographic conditions: 100% buffer A for 5 min, followed by gradients of 0–30% of solution B (66.6% acetonitrile in A) for 5 min, 30–40% B for 30 min, 40–70% B for 5 min, 70–80% B for 10 min, 80–100% B for 5 min, and 100% B for 10 min. The elution was monitored at 280 nm. Fractions were collected manually, lyophilized and stored at −70 °C until used. For affinity chromatography, the protein fraction retained in the Red-Sepharose column was applied overnight to a chitin column (2 × 5 cm) (Sigma-Aldrich) equilibrated in 50 mM Tris/HCl, 150 mM NaCl, pH 7.2. Unbound material was eluted by washing the column with equilibration buffer,

and the retained fraction was desorbed with 50 mM Tris/HCl, 3 M NaCl, pH 7.2.

Molecular mass determinations

Tricine-PAGE in a discontinuous gel and buffer system [48] was used to estimate the apparent molecular mass of the proteins. Samples were denatured for 10 min in sample buffer containing 2.5% (w/v) SDS before electrophoresis. After the run, the gels were stained with Coomassie Brilliant Blue G (0.2%) in methanol/acetic acid/water (4 : 1 : 6, v/v) and destained in the same solution. Protein molecular weight markers (GE Healthcare Biosciences AB, Uppsala, Sweden) were included in each run.

The molecular masses of the native, reduced and carbamidomethylated proteins were determined by MALDI-TOF MS using an Applied Biosystems (Foster City, CA, USA) Voyager PRO-STR instrument operating at an accelerating voltage of 25 kV in the linear mode and using 3,5-dimethoxy-4-hydroxycinnamic acid (10 mg·mL^{-1} in 50% acetonitrile) as the matrix.

The apparent molecular mass of the *Parkia platycephala* lectin 2 in solutions of different pH was determined by size-exclusion chromatography and by analytic ultracentrifugation equilibrium sedimentation using a Beckman XL-A centrifuge with UV absorption scanner optics. For size-exclusion chromatography, PPL2 (2 mg·mL^{-1}) was applied to a Superose-12 HR10/30 column connected to an ÄKTA HPLC system (GE-Healthcare Bioscience). The column was equilibrated and eluted with 20 mM sodium phosphate buffer, pH 7.2, containing 150 mM NaCl at a flow rate of 0.5 mL·min^{-1}. Elution was monitored at 280 nm. Equilibrium sedimentation experiments were carried out at 20 °C and 13 000 r.p.m. using an AN-50 Ti rotor. The protein was dissolved at about 0.1 mg·mL^{-1} in the following buffers, each containing 100 mM NaCl, 1 mM Cl_2Mn, and 1 mM Cl_2Ca: 20 mM sodium citrate pH 2.5; 20 mM sodium citrate, pH 3.5; 20 mM sodium citrate, pH 4.5; 20 mM Mes, pH 5.5; 20 mM Mes, pH 6.5; 20 mM Tris/HCl, pH 7.5; and 20 mM Tris/HCl, pH 8.5.

Quantitation of free cysteine residues and disulfide bonds

For quantitation of free cysteine residues and disulfide bonds, the purified proteins dissolved in 10 μL of 50 mM Hepes, pH 9.0, 5 M guanidine hydrochloride containing 1 mM EDTA were heat-denatured at 85 °C for 15 min, allowed to cool at room temperature, and incubated with either 10 mM 4-vinylpyridine for 15 min at room temperature, or with 10 mM 1,4-dithioerythritol (Sigma-Aldrich) for 15 min at 80 °C; this was followed by addition of 4-vinylpyridine at 25 mM final concentration and incubation for 1 h at room temperature. The pyridylethylated (PE) protein was freed from reagents using a C18 Zip-Tip pipette tip (Millipore Ibérica S.A., Madrid, Spain) after activation with 70% acetonitrile (ACN) and equilibration in 0.1% trifluoroacetic acid. Following protein adsorption and washing with 0.1% trifluoroacetic acid, the PE-protein was eluted onto the MALDI-TOF plate with 1 μL of 70% ACN and 0.1% trifluoroacetic acid and subjected to MS analysis as above.

The number of free cysteine residues (N_{SH}) was determined from Eqn (1):

$$N_{SH} = (M_{PE} - M_{NAT})/105.1 \quad (1)$$

where M_{PE} is the mass of the denatured but nonreduced protein incubated in the presence of 4-vinylpyridine, M_{NAT} is the mass of the native, HPLC-isolated protein, and 105.1 is the mass increment (in Da) due to the pyridylethylation of one thiol group.

The number of total cysteine residues (N_{Cys}) can be calculated from Eqn (2):

$$N_{Cys} = (M_{Alk} - M_{NAT})/105.1 \quad (2)$$

where M_{Alk} is the mass (in Da) of the fully reduced and alkylated protein.

Finally, the number of disulfide bonds N_{S-S} can be calculated from Eqn (3):

$$N_{S-S} = (N_{cys} - N_{SH})/2 \quad (3)$$

Amino acid analysis and N-terminal amino acid sequence determination

Amino acid analysis was performed on a Pico-Tag amino acid analyzer (Waters) as described [49]. One nanomole of purified protein was hydrolyzed in 6 M HCl/1% phenol at 106 °C for 24 h. The hydrolyzate was reacted with 20 μL of fresh derivatization solution (methanol/triethylamine/water/phenylisothiocyanate, 7 : 1 : 1 : 1, v/v) for 1 h at room temperature, and the phenylisothiocyanate (PTC)-amino acids were identified and quantitated on an RP-HPLC column calibrated with a mixture of standard PTC-amino acids (Pierce, Rockford, IL, USA). Cysteine residues were determined as cysteic acid.

N-terminal sequencing of reduced and carboxymethylated proteins was performed in an Applied Biosystems model Procise 491 gas–liquid protein sequencer. The phenylthiohydantoin (PTH) derivatives of the amino acids were identified with an Applied Biosystems model 450 microgradient PTH analyzer.

Genomic DNA and RNA isolation, and cDNA cloning

Genomic DNA from fresh leaves of 2-week-old seedlings of *Parkia platycephala* grown from mature seeds was extracted using the cetyl triethylammonium bromide (CTAB) procedure [50].

For RNA isolation, young *Parkia platycephala* buds were immediately ground to a powder with a pestle in liquid nitrogen. Total cellular RNA was isolated with Concert Plant RNA reagent (Invitrogen S.A., Barcelona, Spain). Single-stranded cDNAs were synthesized by reverse transcription using oligo-dT_{17} and MMLV reverse transcriptase (Promega Biotech Ibérica, Madrid, Spain). Degenerated primers were designed from conserved amino acid sequences of plant chitinases YWGQNGG and WVQFY NNP (sense primer 5′-TAY TGG GAR AAY GGN GG-3′, and antisense primer 5′-GG RTT RTT RTA RAA YTG NAC CCA-3′; the nomenclature follows the IUPAC code for degeneracies). PCR amplification was performed with 1 U (International unit) of Taq DNA polymerase (HF, Roche Diagnostics S.L., Barcelona, Spain) using the following conditions: DNA was denatured at 94 °C for 4 min, and this was followed by 30 cycles of denaturation (30 s at 94 °C), annealing (30 s at 50 °C) and extension (30 s at 72 °C), followed by a final extension for 10 min at 72 °C. The amplified DNA fragment was cloned into the pGEM-T vector (Invitrogen). The inserted DNA fragments were subjected to sequencing on an Applied Biosystems model 377 DNA sequencing system using T7 and SP6 primers, and this sequence was used for designing specific oligonucleotides for completing the sequence by 3′RACE. 3′RACE was done as described [51] using the Qt primer (5′-CCA GTG AGC AGA GTG ACG AGG ACT CGA GCT CAA GCT_{16}-3′) for reverse transcription, and the sense primer GSP-PPL2 (5′-CTG CTG CAC CAC AAT GTC CTT TTC-3′) and the antisense primer Qo (5′-CCA GTG AGC AGA GTG ACG-3′) for PCR amplification. The 3′RACE reaction conditions were as those for cDNA amplification, except that annealing was done at 60 °C. Using this information, two specific primers were designed, PPL2-forward (5′-TAT TGG GGC CAG AAT GGA G-3′) and PPL2-reverse (5′-TCAA ACA CTG GGC TTA ATT TTG G-3′) for amplifying and sequencing the full-length ORF of PPL2.

Assay for chitinase activity

Chitinase enzymatic assays were performed in Pyrex tubes (7 mL) with Teflon-lined screw caps. The reaction mixtures (total 1250 μL) contained 0.05 M sodium acetate buffer (pH 5.5), 5 mg of washed chitin powder (blank), and either 25 μL of a PPL2 solution (1 mg·mL^{-1}) or 10 μL (0.5 μU) of *Streptomyces griseus* family 19 chitinase (Sigma) (one unit will liberate 1.0 mg of GlcNac from chitin per hour at pH 6.0 at 25 °C in a 2 h assay) as positive control, both in sodium acetate buffer. The negative control consisted of the same reaction mixture, except that sodium acetate buffer replaced the protein sample. Twenty-five microliters of 1 mg·mL^{-1} solutions of two GlcNac-specific lectins, the agglutinins from wheat germ (WGA) and *Urtica dioica* (UDA), which are devoid of chitinase activity, were also included in the assays as specificity controls. For calibration and quantitation, a mixture of 1 μg of each, mannose and GlcNac in sodium acetate buffer was used. The reaction mixtures were incubated at 37 °C for 3 h and lyophilized. GlcNac production was monitored and quantitated as peracetylated GlcNac by GC [Varian 3400 gas chromatograph equipped with a flame ionization detector, a Ross injector and a 30 m × 0.25 mm capillary column EC.Tm^{-1} (100% methylsilicone apolar phase of column, EC.Tm^{-1}, 0.25 μm film phase, Altech), 0.25 μm film phase (Altech, Flemington, NJ, USA)]. The injector and detector temperature was 250 °C, and the oven temperature program was 3 C°·min^{-1} from 120 to 250 °C. The carrier gas helium pressure was 1 bar. Briefly, released GlcNac was peracetylated by addition of 0.5 mL of acetic anhydride to the lyophilized samples, followed by incubation for 4 h at 100 °C. Samples were then evaporated to dryness under a stream of nitrogen and mild heating with a hair dryer. To eliminate salts and proteins from the reaction mixture, 1.5 mL of chloroform and 1 mL of distilled water were added to each tube. After thorough vortexing, the aqueous upper phase was discarded and the lower chloroform phase was extracted four times with 1 mL of distilled water. The chloroform phases were freed of water by filtration through small columns made of a Pasteur pipette filled with anhydrous sodium sulfate. The filtrates were collected in Pyrex tubes (7 mL) and evaporated to dryness under a stream of nitrogen. Chloroform (40 μL) was added to each tube, and 4 μL was injected in the gas chromatograph for analysis.

GlcNac production (retention time 33.60 min) was also monitored by GC/MS analysis performed on a Carlo Erba GC 8000 gas chromatograph equipped with a 25 m × 0.32 mm CP-Sil 5CB low-bleed MS capillary column, 0.25 μm film phase (Chrompack France, Les Ullis, France). The temperature of the Ross injector was 250 °C and the samples were analyzed using the following temperature program: 120 °C for 3 min, then 3 C°·min^{-1} until 250 °C. The column was coupled to a Finnigan Automass II mass spectrometer. The analyses were performed either in the electron impact mode (ionization energy 70 eV, source temperature 150 °C) or in the chemical ionization mode in the presence of ammonia (ionization energy 150 eV, source temperature 100 °C). Detection was performed for positive ions.

High-pH anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD)

HPAEC-PAD was performed with a Dionex Series DX30 HPLC system (Dionex Corporation, Voisins Le Bretonneux, France) equipped with a pulsed electrochemical detector, operating in the pulsed amperometric detection mode with a gold working electrode and an Ag/AgCl reference electrode. Electrode potential settings were E1 + 0.05 V, E2 + 0.6 V and E3 − 0.6 V, with 500, 3 and 7 ms applied durations, respectively, and an integrated time period of

0.10–0.48 s. Detection was set with a range of detection of 300 nC. Detector response was analyzed with a C-R8A chromatopac integrator (Shimadzu, Kyoto, Japan). A standard sample consisted of 0.1 $\mu g \cdot \mu L^{-1}$ of each GlcNac, *N*,*N*′-diacetylchitobiose or *N*,*N*′,*N*″-triacetylchitotriose dissolved in water. Samples (12.5 μL) were injected in a Dionex CarboPac PA-100 pellicular anion exchange column running at a flow rate of 0.8 $mL \cdot min^{-1}$. Elution was performed with buffer A (100 mM NaOH) for 1 min followed by a linear gradient of 0–40% buffer B (100 mM NaOH and 1 M sodium acetate) over 40 min.

Crystallization and structure determination

PPL2 was crystallized by the hanging drop vapor diffusion method at 20 °C as described [28]. The crystals belong to the $P2_12_12_1$ space group with one monomer in the asymmetric unit. Crystals soaked in a cryoprotectant solution containing 75% of mother liquor [0.2 M ammonium acetate, 0.1 M trisodium citrate dehydrate, pH 5.6, and 30% (w/v) PEG 4000] and 25% of glycerol were flash-frozen at 100 K in a liquid nitrogen stream. X-ray diffraction data were collected at 1.73 Å at the synchrotron radiation source of Cpr station Laboratório Nacional de Luz Síncrotron (Campinas, Brazil). The data were processed and scaled using MOSFLM and SCALA [52], respectively. Crystallographic data are summarized in Table 2.

The PPL2 crystal structure was determined by molecular replacement using the AMORE software [52], using data in the resolution range 15–3.0 Å, and the hevamine coordinates (PDB accession code 2HVM) as the search model. Rotation and translation functions revealed one molecule in the asymmetric unit. The position and orientation of the molecule, as a single rigid body entity, were refined for 20 cycles with REFMAC [52], using reflections in the resolution range 32–1.73 Å. Appropriate amino acid changes were carried out to convert the molecular model of hevamine into PPL2. Several steps of rebuilding, interspersed with restrained refinement, using REFMAC, yielded the current model at 1.73 Å resolution. Sulfate ion molecules were placed by inspection of the $F_o - F_c$ map. For each cycle of refinement, the stereochemistry of the model was monitored with the PROCHECK incorporated into the CCP4 package [52]. Finally, water molecules were placed in the model over several steps of refinement with ARP/WARP and inspected manually. The atomic coordinates, fitted with XTALVIEW [52], are accesible from the Protein DataBank (http://www.rcsb.org/pdb/) under code 2GSJ.

Acknowledgements

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), CAPES, FUNCAP, PADCT and Program CAPES/COFECUB no. 336/01, and grant BFU2004-01432/BMC from the Ministerio de Educación y Ciencia, Madrid, Spain. B. S. Cavada, W. F. De Azevedo Jr and A. H. Sampaio are senior investigators of CNPq.

References

1 Van Damme EJM, Peumans WJ, Barre A & Rougé P (1998) Plant lectins: a composite of several distinct families of structurally and evolutionary related proteins with diverse biological roles. *Crit Rev Plant Sci* **17**, 575–692.
2 Gabius H-J & Gabius S (1997) *Glycoscience. Status and Perspectives*. Chapman & Hall, Weinheim.
3 Dodd RB & Drickamer K (2001) Lectin-like proteins in model organisms: implications for evolution of carbohydrate-binding activity. *Glycobiology* **11**, 71–79.
4 Rini JM (1995) Lectin structure. *Annu Rev Biomol Struct* **24**, 551–577.
5 Weis WI & Drickamer K (1996) Structural basis of lectin-carbohydrate recognition. *Annu Rev Biochem* **65**, 441–473.
6 Elgavish S & Shaanan B (1997) Lectin–carbohydrate interactions: different folds, common recognition principles. *Trends Biochem Sci* **22**, 462–467.
7 Loris R, Hamelryck T, Bouckaert J & Wyns L (1998) Legume lectin structure. *Biochim Biophys Acta* **1383**, 9–36.
8 Bouckaert J, Hamelryck T, Wyns L & Loris R (1999) Novel structures of plant lectins and their complexes with carbohydrates. *Curr Opin Struct Biol* **9**, 572–577.
9 Vijayan M & Chandra N (1999) Lectins. *Curr Opin Struct Biol* **9**, 707–714.
10 Chervenak MC & Toone EJ (1995) Calorimetric analysis of the binding of lectins with overlapping carbohydrate binding. *Biochemistry* **34**, 5685–5695.
11 Dam TK, Cavada BS, Grangeiro TB, Santos CF, de Sousa FAM, Oscarson S & Brewer CF (1998) Diocleinae lectins are a group of proteins with conserved binding sites for the core trimannoside of asparagine-linked oligosaccharides and differential specificities for complex carbohydrates. *J Biol Chem* **273**, 12082–12088.
12 Dam TK, Cavada BS, Grangeiro TB, Santos CF, Ceccatto VM, de Sousa FAM, Oscarson S & Brewer CF (2000) Thermodynamic binding studies of lectins from the diocleinae subtribe to deoxy analogs of the core trimannoside of asparagine-linked oligosaccharides. *J Biol Chem* **275**, 16119–16126.
13 Dam TK, Roy R, Das SK, Oscarson S & Brewer CF (2000) Binding of multivalent carbohydrates to concanavalin A and Dioclea grandiflora lectin. Thermodynamic analysis of the 'multivalency effect'. *J Biol Chem* **275**, 14223–14230.

14 Suvachittanont W & Peutpaiboon A (1992) Lectin from *Parkia speciosa* seeds. *Phytochemistry* **31**, 4065–4070.
15 Utarabhand P & Akkayanont P (1995) Purification of a lectin from *Parkia javanica* beans. *Phytochemistry* **38**, 281–285.
16 Cavada BS, Madeira SVF, Calvete JJ, Sousa LAG, Bomfim LR, Dantas AR, Lopes MC, Grangeiro TB, Freitas BT, Pinto VPT *et al.* (2000) Purification, chemical, and immunochemical properties of a new lectin from Mimosoideae (*Parkia discolor*). *Prep Biochem Biotech* **30**, 271–280.
17 Cavada BS, Santos CF, Grangeiro TB, Moreira da Silva LIM, Campos MJO, de Sousa FAM & Calvete JJ (1997) Isolation and partial characterization of a lectin from *Parkia platycephala* Benth seeds. *Physiol Mol Biol Plant* **3**, 109–115.
18 Ramos MV, Cavada BS, Bomfim LR, Debray H, Mazard A-M, Calvete JJ, Grangeiro TB & Rougé P (1999) Interaction of the seed lectin from *Parkia platycephala* (*Mimosoideae*) with carbohydrates and complex glycans. *Prot Pept Lett* **6**, 215–222.
19 Mann K, Farias CM, Gallego del Sol FG, Santos CF, Grangeiro TB, Nagano CS, Cavada BS & Calvete JJ (2001) The amino-acid sequence of the glucose/mannose-specific lectin isolated from *Parkia platycephala* seeds reveals three tandemly arranged jacalin-related domains. *Eur J Biochem* **268**, 4414–4422.
20 Gallego del Sol F, Gómez J, Hoos C, Nagano CS, Cavada BS, England P & Calvete JJ (2005) Energetics of 5-bromo-4-chloro-3-indolyl-α-D-mannose binding to the *Parkia platycephala* seed lectin and its use for MAD phasing. *Acta Cryst F* **61**, 326–331.
21 Gallego del Sol F, Nagano CS, Cavada BS & Calvete JJ (2005) The first crystal structure of a *Mimosoideae* lectin reveals a novel quaternary arrangement of a widespread domain. *J Mol Biol* **353**, 574–583.
22 Heywood VH (1971) Chemotaxonomy of the *Leguminosae* (Harborne JB & Boulter D, eds), pp. 1–29. Academic Press, London.
23 Chrispeels MJ & Raikhel NV (1991) Lectins, lectin genes, and their role in plant defense. *Plant Cell* **3**, 1–9.
24 Wang X & Ma Q (2005) Characterization of a jasmonate-regulated wheat protein related to a β-glucosidase-aggregating factor. *Plant Physiol Biochem* **43**, 185–192.
25 Collinge DB, Kragh KM, Mikkelsen JD, Nielsen KK, Rasmussen U & Vad K (1993) Plant chitinases. *Plant J* **3**, 31–40.
26 Hamel F, Boivin R, Tremblay C & Bellemare G (1997) Structural and evolutionary relationships among chitinases of flowering plants. *J Mol Evol* **44**, 614–624.
27 Kasprzewska A (2003) Plant chitinases – regulation and function. *Cell Mol Biol Lett* **8**, 809–824.
28 Cavada BS, Castellón RER, Vasconcelos GG, Rocha BAM, Bezerra GA, Debray H, Delatorre P, Nagano CS, Toyama M, Pinto VPT *et al.* (2005) Crystallization and preliminary X-ray diffraction analysis of a new chitin-binding protein from *Parkia platycephala* seeds. *Acta Crystallogr F* **61**, 841–843.
29 Rawitch AB, Pollock HG & Yang S-X (1993) Thyroglobulin glycosylation: location and nature of the N-linked oligosaccharide units in bovine thyroglobulin. *Arch Biochem Biophys* **300**, 271–279.
30 Hill HD Jr, Reynolds JA & Hill RL (1977) Purification, composition, molecular weight, and subunit structure of ovine submaxillary mucin. *J Biol Chem* **252**, 3791–3798.
31 Spiro RG & Bhoyroo D (1974) Structure of the O-glycosidically linked carbohydrate units of fetuin. *J Biol Chem* **249**, 5704–5717.
32 Green ED, Adelt G, Baenziger JU, Wilson S & Van Halbeek H (1988) The asparagine-linked oligosaccharides on bovine fetuin. Structural analysis of N-glycanase-released oligosaccharide by 500-megahertz ^{1}H NMR spectroscopy. *J Biol Chem* **263**, 18253–18268.
33 Rohrer JS, Cooper GA & Townsend RR (1993) Identification, quantitation, and characterization of glycopeptides in reversed-phase HPLC separations of glycoprotein proteolytic digests. *Anal Biochem* **212**, 7–16.
34 Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* **280**, 309–316.
35 Jekel PA, Hartmann BH & Beintema JJ (1991) The primary structure of hevamine, an enzyme with lysozyme/chitinase activity from *Hevea brasiliensis* latex. *Eur J Biochem* **200**, 123–130.
36 Van Scheltinga ACT, Kalk KH, Beintema JJ & Dijkstra BW (1994) Crystal structures of hevamine, a plant defense protein with chitinase and lysozyme activity, and its complex with an inhibitor. *Structure* **2**, 1181–1189.
37 Chye ML, Zhao KJ, He ZM, Ramalingam S & Fung KL (2005) An agglutinating chitinase with two chitin-binding domains confers fungal protection in transgenic potato. *Planta* **220**, 717–730.
38 Tang CM, Chye ML, Ramalingam S, Ouyang SW, Zhao KJ, Ubhayasekera W & Mowbray SL (2004) Functional analyses of the chitin-binding domains and the catalytic domain of *Brassica juncea* chitinase BjCHI1. *Plant Mol Biol* **56**, 285–298.
39 Robertus JD & Monzingo AF (1999) The structure and action of chitinases. *EXS* **87**, 125–135.
40 Kaomek M, Mizuno K, Fujimura T, Sriyotha P & Cairns JR (2003) Cloning, expression, and characterization of an antifungal chitinase from *Leucaena leucocephala* de Wit. *Biosci Biotechnol Biochem* **67**, 667–676.
41 Hennig M, Jansonius JN, Van Scheltinga ACT, Dijkstra BW & Schlesier BJ (1995) Crystal structure of concanavalin B at 1.65 A resolution. An 'inactivated' chitinase

from seeds of *Canavalia ensiformis*. *J Mol Biol* **254**, 237–246.
42 Van Scheltinga ACT, Hennig M & Dijkstra BW (1996) The 1.8 Å resolution structure of hevamine, a plant chitinase/lysozyme, and analysis of the conserved sequence and structure motifs of glycosyl hydrolase family 18. *J Mol Biol* **262**, 243–257.
43 Lawton KA, Beck J, Potter S, Ward E & Ryals J (1994) Regulation of cucumber class III chitinase gene expression. *Mol Plant-Microbe Interact* **7**, 48–57.
44 Copley RR & Barton GJ (1994) A structural analysis of phosphate and sulphate binding sites in proteins. Estimation of propensities for binding and conservation of phosphate binding sites. *J Mol Biol* **242**, 321–329.
45 Bokma E, Rozeboom HJ, Sibbald M, Dijkstra BW & Beintema JJ (2002) Expression and characterization of active site mutants of hevamine, a chitinase from the rubber tree *Hevea brasiliensis*. *Eur J Biochem* **269**, 893–901.
46 Pastuszak I, Drake R & Elbein AD (1996) Kidney *N*-acetylgalactosamine (GalNAc)-1-phosphate kinase, a new pathway of GalNAc activation. *J Biol Chem* **271**, 20776–20782.
47 Ainouz IL, Sampaio AH, Benevides NMB, Freitas ALP, Costa FHF, Carvalho MR & Pinheirojoventino F (1992) Agglutination of enzyme treated erythrocytes by Brazilian marine algal extracts. *Bot Mar* **35**, 475–479.
48 Schägger H & von Jagow G (1987) Tricine–sodium dodecyl sulfate–polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**, 368–379.
49 Henrikson RL & Meredith SC (1984) Amino acid analysis by reversed-phase high-performance liquid chromatography: precolumn derivatization with phenylisothiocyanate. *Anal Biochem* **136**, 65–71.
50 Steenkamp J, Wiid I, Lourens A & van Helden P (1994) Improved method for DNA extraction from Vitis vinifera. *Am J Enol Vitic* **45**, 102–106.
51 Frohman MA & Martin GR (1989) Rapid amplification of cDNA ends using nested primers. *Techniques* **1**, 165–170.
52 Collaborative Computational Project Number 4 (1994). *Acta Cryst* **D50**, 760–763.