Appl Biochem Biotechnol (2009) 152:383–393
DOI 10.1007/s12010-008-8334-9

Purification, Characterization, and Preliminary X-Ray Diffraction Analysis of a Lactose-Specific Lectin from *Cymbosema roseum* Seeds

Bruno A. M. Rocha · Frederico B. M. B. Moreno · Plínio Delatorre · Emmanuel P. Souza · Emmanuel S. Marinho · Raquel G. Benevides · Joane Kathelen Rodrigues Rustiguel · Luis A. G. Souza · Celso S. Nagano · Henri Debray · Alexandre H. Sampaio · Walter F. de Azevedo Jr. · Benildo S. Cavada

Received: 8 January 2008 / Accepted: 29 July 2008 /
Published online: 19 August 2008

Abstract The unique carbohydrate-binding property of lectins makes them invaluable tools in biomedical research. Here, we report the purification, partial primary structure, carbohydrate affinity characterization, crystallization, and preliminary X-ray diffraction analysis of a lactose-specific lectin from *Cymbosema roseum* seeds (CRLII). Isolation and purification of CRLII was performed by a single step using a Sepharose-4B-lactose affinity chromatography column. The carbohydrate affinity characterization was carried using

B. A. M. Rocha · E. P. Souza · E. S. Marinho · R. G. Benevides · B. S. Cavada (✉)
BioMol-Lab, Departmento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará,
P. O. Box 6043, 60.455-970 Fortaleza, Ceará, Brazil
e-mail: bscavada@ufc.br

B. A. M. Rocha · P. Delatorre
Departamento de Ciências Físicas e Biológicas, Universidade Regional do Cariri,
Fortaleza, Ceará, Brazil

F. B. M. B. Moreno · J. K. R. Rustiguel
Departamento de Física, UNESP, São José do Rio Preto, São Paulo, Brazil

L. A. G. Souza
Instituto Nacional de Pesquisas da Amazônia—INPA, Manaus, Amazonas, Brazil

C. S. Nagano · A. H. Sampaio
Biomol-Mar, Departamento de Engenharia de Pesca, Universidade Federal do Ceará, Fortaleza, Brazil

H. Debray
Laboratoire de Chimie Biologique et Unité Mixte de Recherche N° 8576 du CNRS,
Université des Sciences et Technologies de Lille, Lille, France

W. F. de Azevedo Jr.
Faculdade de Biociências, Centro de Pesquisas em Biologia Molecular e Funcional, PUCRS,
Porto Alegre, Brazil

assays for hemagglutination activity and inhibition. CRLII showed hemagglutinating activity toward rabbit erythrocytes. *O*-glycoproteins from mucine mucopolysaccharides showed the most potent inhibition capacity at a minimum concentration of 1.2 μg mL^{-1}. Protein sequencing by mass spectrometry was obtained by the digestion of CRLII with trypsin, Glu-C, and AspN. CRLII partial protein sequence exhibits 46% similarity with the ConA-like α chain precursor. Suitable protein crystals were obtained with the hanging-drop vapor-diffusion method with 8% ethylene glycol, 0.1 M Tris-HCl pH 8.5, and 11% PEG 8,000. The monoclinic crystals belong to space group $P2_1$ with unit cell parameters a=49.4, b=89.6, and c=100.8 Å.

Keywords *Cymbosema roseum* · Crystallization · Tandem mass spectrometry · Lectins · Lactose-specific lectin

Introduction

Lectins are glycoproteins that interact reversibly with sugars or glycoconjugates [1] and have been purified from organisms as diverse as microbes, animals, and plants [2]. Although the endogenous function of plant lectins remains unknown, interest in their study has increased because of their various biological and biotechnological applications. Furthermore, some lectins are excellent tools for studies of structure or function relationships. The unique carbohydrate-binding property of lectins makes them invaluable tools in biomedical research [2].

Lectins isolated from the same species usually have the same sugar specificity, although different lectins isolated from the same plant (e.g., *Dioclea lehmanni*) have shown distinct biological properties, such as sugar specificities. These proteins have distinct polypeptide sequences [3–5] or even distinct quaternary assembles [6, 7] when compared with ConA-like lectins despite the high gene homology among legume lectins [5].

The posttranslational sequence circular permutation, as described to ConA-like lectins by Carrington and coworkers [8], probably does not occur in all legume lectins, even when the lectins are from the same species. For example, it does not occur in galactose-specific lectin from *D. lehmanni* (DLL II) [5].

The *Papilionoideae* subfamily, especially the *Erythrina* genus, contains galactose-specific lectins. Some examples are *Erythrina velutina* [9], *Erythrina coralodendrum* [10], *Erythrina cristagalli* [11], *Erythrina indica* [12], and *Erythrina speciosa* [13]. Crystal X-ray structures of *E. coralodendrum* and *E. cristagalli* lectins in complex with carbohydrates have been solved, confirming their lactose-specific affinity. Galactose-specific legume lectins are also found in the subfamily *Caesalpinoideae*, for example in *Bauhinia variegata* candida (BVcL). Carbohydrate inhibition analysis indicated that BvcL is inhibited by lactose, galactose, galactosamine, and other galactoside derivates [14].

Recently, we purified and crystallized a mannose-specific lectin (CRLI) from *Cymbosema roseum* seeds [15], an Amazonian species from Papilionoideae subfamily. The present paper reports the purification, the partial primary structure obtained by tandem mass spectrometry, the protein characterization by carbohydrate affinity, and the crystallization and preliminary X-ray diffraction analysis of a second lectin (CRLII - lactose-specific) from *C. roseum* seeds.

The study was developed in order to address structural features about circular permutation in posttranslational process in ConA-like lectins and its relationship with carbohydrate interactions and lectins specificity. This work is a preliminary structural

characterization of a lectin and the complete crystal structures of both lectins isolated from this plant (CRLI e CRLII) are been solved by our group and constitute a valuable model to understand the differences in carbohydrate recognition.

Material and Methods

CRLII Purification

Isolation and purification of *C. roseum* lectin II (CRLII; lactose-specific) was performed as previously described for CRLI but with lactose instead of a mannose affinity column [15]. *C. roseum* seeds were ground to a fine powder in a coffee mill. The powder was mixed with 0.15 M NaCl (1:10, w/v) at room temperature for 4 h and then centrifuged at 10,000 *g* for 20 min at 278 K. The resultant supernatant was applied to a Sepharose-4B-lactose column equilibrated with a solution of 0.15 M NaCl, 5 mM $CaCl_2$ and 5 mM $MnCl_2$. After removing unbound material, the lectin was eluted with 0.1 M glycine, 0.15 M NaCl, pH 2.6. The eluted sample was exhaustively dialyzed against water and lyophilized. Purified CRLII was monitored by SDS-PAGE as described by Laemmli [16] and was used for inhibition assays, mass spectrometric analysis, and crystallization trials.

Hemagglutination Activity and Inhibition Assays

Hemagglutinating activity and inhibition tests were carried out according to standard procedure [17] using native and enzyme-treated rabbit erythrocytes. Inhibition tests were carried out using monosaccharides and dissaccharides (α-D-glucose, α-D-mannose, α-D-galactose, α-D-lactose, α-L-fucose, N-acetyl-α-D-galactosamine, N-acetyl-α-D-glucosamine and N-acetyl-α-D-lactosamine), N-glycoproteins (human serotransferrin, desialylated human serotransferrin, bovine lactotransferrin and desialylated bovine lactotransferrin) and O-glycoproteins (bovine fetuin, bovine asialofetuin, bovine submaxillary mucin, porcine stomach mucin, ovine submaxillary mucin and desialylated ovine submaxillary mucin). The saccharides were purchased from Sigma-Aldrich and the glycoproteins were isolated by Laboratoire de Chimie Biologique et Unité Mixte de Recherche (Université des Sciences et Technologies de Lille, France). The CRLII samples were diluted in 0.15 M NaCl to give a concentration corresponding to 4 units of hemagglutinating activity per mL. The titre was defined as 2^{10} (the highest dilution of the CRLII that can agglutinate erythrocytes, i.e., the titre, defined as containing one hemagglutinating unit per milliliter). Aliquots (0.2 mL) of the diluted lectin (1/256) were added to each well of the diluted inhibitor series. The plates were then left at room temperature for 1 h before 0.2 mL of 2% native or enzyme-treated rabbit erythrocytes was added to each well. After 1 h the plates were examined for hemagglutination by microscopic observation. The hemagglutination inhibition titre was recorded as the highest dilution of inhibitor which inhibited agglutination produced by 4 hemagglutinating units of CRLII sample.

Protein Cleavages and Isolation of Peptides

One mg of lyophilized CRLII was dissolved in 100 mM ammonium bicarbonate, after which it was mixed with 25 μL of 2 mg/mL bovine pancreatic trypsin (Sigma-Aldrich), and incubated for 3 h at 310 K. A second aliquot, also in 100 mM ammonium bicarbonate pH 8.6 was digested with endoproteinase Asp-N (Sigma-Aldrich) from *Pseudomonas fragi*

for 8 h at 310 K. A third aliquot was dissolved in 50 mM Tris-HCl pH 8.0 and digested with Glu-C (Sigma-Aldrich) from *Staphylococcus aureus* for 16 h at 298 K. The enzyme: substrate ratio for AspN and Glu-C was 1:50 and 1:20 (w/w), respectively. All the reaction mixtures were centrifuged at 13,000 *g* for 10 min, supernatants were dried in a Speed-Vac. The peptides were loaded on a reversed-phase HPLC C18 (5 μm particle size) column and eluted at 1 mL/min with a gradient of 0.1% trifluoroacetic acid in water (solution A) and acetonitrile (solution B), isocratically (5% B) for 5 min, followed by 5–40% B for 60 min, and finally 40–70% B for 20 min.

Protein Sequencing by Tandem Mass Spectrometry

The digested peptides samples isolated by reverse phase were diluted 100x with Milli-Q™ water, injected into a nano-electrospray ionization source and analyzed in Micromass™ quadrupole time of flight instrument from Waters™(Massachusetts, USA) with a resolution of 8,000 and an accuracy of 10 ppm. The data were processed using the manufacturer's Biolynx 4.0 program. The peptides were sequenced by ProteinLynx (Micromass™) and MASCOT (Matrix Science) programs. An alignment and phylogenetic analysis of the CRLII sequence with all the non-redundant proteins deposited in the National Center of Biotechnology Information (NCBI) was performed with ClustalW [18] and BLAST [19].

Crystallization and X-ray Data Collection

The lyophilized, purified CRLII was dissolved at a concentration of 12 $mg.mL^{-1}$ in 10 mM Tris-HCl pH 7.6, 1.0 mM $CaCl_2$ and 1.0 mM $MnCl_2$ for use in crystallization trials. Crystallization screening by the hanging-drop vapor-diffusion method was performed in Linbro plates at 293 K using Crystal Screen and Crystal Screen II (Hampton Research, Aliso Viejo, CA, USA). The drops were composed of equal volumes (2 μL) of protein solution and reservoir solution and were equilibrated against 500 μL reservoir solution. Crystals grew in 0.1 M Tris-HCl, containing 10% (w/v) PEG 8,000. A crystal was transferred to a cryoprotectant solution consisting of 30% glycerol and 70% reservoir solution The data were collected at 1.42 Å wavelength at beamline MX1 station (Laboratório Nacional da Luz Síncrotron – LNLS, Campinas, Brazil) using a CCD (MAR research) imaging plate at 70 mm crystal to detector distance. A set of 120 images (1° oscillation) was recorded. Diffraction data were indexed, integrated and scaled using MOSFLM [20] and SCALA [21]. Full amino acid sequence by Edman's degradation and 3D structure determination by X-ray crystallography are in progress.

Results and Discussion

Purification and Affinity Characterization

Cymbosema roseum lectin II (CRLII) was purified by a single step using a Sepharose-4B-lactose affinity chromatography column (Fig. 1). Purification was monitored by hemagglutination activity and SDS-PAGE.

CRLII showed hemagglutinating activity toward native and enzyme-treated rabbit erythrocytes. The minimal concentration of purified protein that agglutinated a 2% rabbit non-treated erythrocyte suspension was $<2\ \mu g.mL^{-1}$, a value similar to that for other ConA-like lectins. The glycan-recognizing specificity of CRLII was investigated through

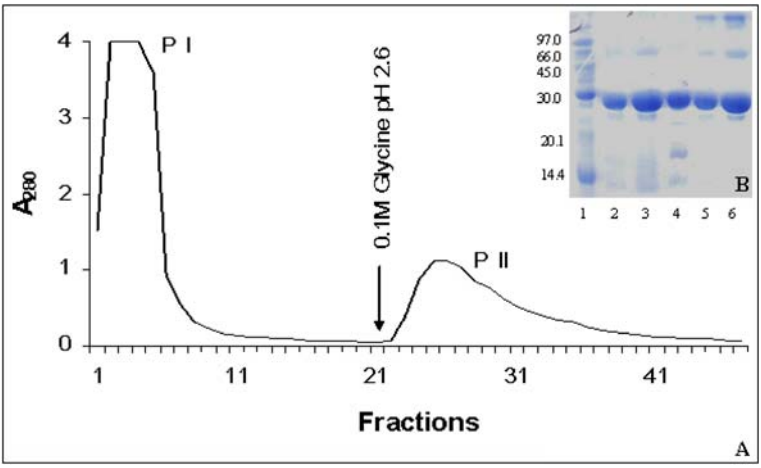

Fig. 1 Affinity chromatography and SDS electrophoresis of the purified *Cymbosema roseum* lactose-specific lectin (CRLII). (*A*) Sepharose-4B-lactose chromatogram; the peak labeled PI is the unbound sample eluted with the equilibration buffer. PII represents a peak with hemagglutination activity and was eluted with 0.1 M glycine pH 2.6. (*B*) SDS-PAGE showing protein markers (*lane 1*), reduced CRLII α chain 10 μg/mL (*lane 2*), and 20 μg/mL (*lane 3*), α, β, and γ chains of nonreduced lectin from ConBr (*lane 4*) and nonreduced CRLII α chain (10 μg/mL—*lane 5*; 20 μg/mL—*lane 6*). The high molecular bands present in *lanes 2–6* is the pre-pro-protein with a MW of approximately 65 kDa

hemagglutination-inhibition assays using several different carbohydrates and glycoproteins (Table 1).

The saccharides showed a substantial inhibitory activity. N-acetyl-α-D-galactosamine and N-acetyl-α-D-lactosamine were the two most potent of the sacchardies, at minimum inhibitory concentrations of 1.17 and 2.34 mM, respectively (Table 1). Desialylated bovine lactotransferrin, inhibited CRLII at 2.4 $\mu g.mL^{-1}$ and the O-glycoproteins from mucin origin (bovine submaxillary mucin, porcine stomach mucin and desialylated ovine submaxillary mucin) showed the most potent inhibition capacity at a minimum concentration of 1.2 $\mu g.mL^{-1}$ (Table 1). The porcine stomach mucin contains O-linked carbohydrate structures sharing a core 1 Galβ1–3GalNAc disaccharide. These structures can be substituted by N-acetyllactosamine branches terminated with fucose α1–2-Gal, GalNAcα1–3[Fuc α1–2]Gal, or GlcNAcα1–4-Gal at their nonreducing ends. Tn (GalNAcα1-Ser/ Thr) and T (Galβ (1–3) GalNAcα1-Ser/Thr) antigens are also present in the porcine stomach mucin [22, 23, 24].

Bovine submaxillary mucin is a glycoprotein bearing at least 16 different carbohydrate moieties [25, 26], 85% are acidic O-linked oligosaccharide chains, including a high density of sialyl Tn antigens and sialyl core 3 saccharide sequences. The neutral O-linked glycans of bovine submaxillary mucin include GalNAcα1–3[Fuc α1–2]Gal, N-acetyllactosamine branches terminated with fucose α1–2-Gal and the core 3 determinants [25, 26]. In addition, Tn antigen accounts for >75% of the carbohydrate chains of the desialylated ovine submaxillary mucin. Thus, the carbohydrate affinity data indicate that CRLII preferably binds to GalNAc- and Gal- substituted with a neutral sugar through 1–3, 1–4, or 1–2 linkages. The desialylated bovine lactotransferrin was 2-fold less inhibitory than O-glycoproteins.

The molecular basis of the binding properties of lactose-binding legume lectins are available only for three structures: ECL (PDB code 1GZC) [27], EcorL (PDB code 1AX1)

Table 1 Inhibition of haemagglutation by saccharides and glycoproteins of the lactose-specific lectin CRLII from *Cymbosema roseum.*

	Minimal Inhibition Concentration
Saccharides	mM
L-Fucose	>75
D-Glucose	37.5
D-Mannose	37.5
N-Acetyl D-glucosamine	18.75
D-Galactose	9.37
D-Lactose	2.34
N-Acetyl-D-lactosamine	2.34
N-Acetyl D-galactosamine	1.17
N-Glycoproteins	μg/mL
Human serotransferrin	>2500
Bovine lactotransferrin	78.1
Desialylated human serotransferrin	4.8
Desialylated bovine lactotransferrin	2.4
O-Glycoproteins	μg/mL
Bovine fetuin	>2500
Ovine submaxillary mucin	4.8
Bovine asialofetuin	2.4
Bovine submaxillary mucin	1.2
Porcine stomach mucin	1.2
Desialylated ovine submaxillary mucin	1.2

[28] and PNA (PDB code 1CR7) [29]. Crystallographic studies on CRLII would therefore contribute to understanding the specificity of the sugar-binding features in lectins.

Protein Sequence and Biochemical Characterization

The apparent molecular mass of CRLII was determined by electrophoresis and the lectin shows only a single 25 kDa electrophoresis band even in reduced conditions (Fig. 1). This is the same mass found in other ConA-like lectins like *Canavalia brasiliensis* lectin (ConBR). The gel does not show the three polypeptide chains with a molecular masses of 25.7, 12.8, and 8 kDa commonly observed in lectins from Papilionoideae subfamily, such as *C. brasiliensis* (see Fig. 1) and *Canavalia ensiformis* [30]. ConA is composed of two chains (β and γ). These chains are cleavage products of pre-pro-protein. The active protein is a fused final product (α chain) in an inversed order of these two smaller chains [31], although some lectins such as DLL II [4], *Vatairea macrocarpa* lectin (VML) [32], and *Robinia pseudoacacia* lectin (Rob) [33] have been purified and identified as originating from the same lectin gene, similar to the lectins previously purified from the same leguminoseae species. This fact can explain the high similarity found between unprocessed protein sequences and the chain precursor of ConA-like lectins. CRLII has an α chain sequence without the inversion of β and γ chains but shows an amino acid sequence similar to the gene products of CRLI and DLL I before their posttranslational modifications.

CRLII partial protein sequence was obtained from overlaps of digested peptides sequenced by tandem mass spectrometry (Q-ToF), in which 28 peptides were sequenced resulting in a total of 189 amino acids residues and a coverage of approximately 80%

(Table 2). This preliminary polypeptide chain presents a molecular mass of 20,528 Da, leaving a nonsequenced mass of approximately 5,000 Da.

This sequence does not show similarity higher than 30% with other known mature ConA-like lectins such as *D. lehmanni* lectin I (DLL I) [4], ConBr [34], *Dioclea grandiflora* lectin (Dgran) [35], concanavalin A (ConA) [30] and *Phaseolus vulgaris* agglutinin (PHA) [36] (Fig. 2). However, CRLII partial sequence exhibits 47% similarity with the ConA-like precursor (CAA25787) [8, 37].

The phylogenetic analysis based on primary sequences shows a monophyletic group of proteins, and basically, three paraphyletic groups in it. The group containing CRLII is composed by lectins from *Papilionoideae* and *Mimmosoideae* subfamilies; the second group is prior composed by members from *Dioclea* genus; and the third by members from *Canavalia* genus (both belonging to *Papilionoideae* subfamily). The monophyletic group synapomorphy is characterized by the carbohydrate binding property and the post-

Table 2 Amino acid peptide sequences of CRLII obtained by tandem mass spectrometry and their correspondening m/z of analyzed ions.

Enzyme digestion	Peptide sequence	m/z (M + H)
D1 (D017–L025)	D L T L Q G S A L	458.50
D2 (D044–S054)	D H S V G R A L Y R S	630.15
D3 (D060–F069)	D E T T G K V A S F	527.05
D4 (D070–A079)	D A T F S F V S E A	536.55
D5 (D097–Q106)	D T Q P D R I G G Q	599.60
D6 (D111–V119)	D S Y N V I R T V	533.05
D7 (G125–V130)	G Y R N R V	381.90
D8 (D131–L139)	D P S A R H X S L	496.60
D9 (D187–S196)	D V X X N A V L P S	518.25
D10 (D215–W225)	D H V E T N T X X S W	655.35
T1 (T019–K029)	T L Q G S A L V S S K	545.10
T2 (G040–R049)	G K P V D H S V G R	525.55
T3 (A079–K090)	A P A I P M L F P S S K	629.25
T4 (G091–R102)	G E L N P P D T Q P D R	669.20
T5 (G107–R117)	G V V N D S Y N V I R	617.65
T6 (T118–R129)	T V A V E N D G Y R N R	696.70
T7 (V130–R135)	V D P S A R	321.85
T8 (H136–K146)	H X S L P K L K I A K	622.85
T9 (N147–K151)	N N G I K	272.30
T10 (W152–K158)	W N M Q T G K	432.67
T11 (V172–K184)	V G T A H X S Y N S V A K	673.83
T12 (A175–R185)	A H X S Y N S V A K R	621.75
T13 (Y186–K198)	Y D V X X N A V L P S K	663.95
T14 (N220–K230)	N T X X S W S F T S K	639.85
T15 (X233–N239)	X T N S D S N	373.95
E1 (S068–E078)	S F D A T F S F V S E	618.15
E2 (P080–E092)	P A I P M L F P S S K G E	686.80
E3 (A208–E218)	A T S G X Y K D H V E	608.70

The peptides were obtained by enzymatic digestions of (*D*) endoproteinase AspN from *Pseudomonas*, (*E*) endoproteinase Glu-C or (*T*) trypsin.

*The X residues can be either isoleucine or leucine. The numbers indicate the possible sequence position based on the alignments.

Fig. 2 Phylogenetic analysis based on legume lectins sequence alignment. The protein sequence alignment made among partial CRLII sequence with legume seed lectins: CRLI [15], Dgran [35], *Dioclea guianensis* lectin (Dgui) [37], *Dioclea violacea* lectin (DVL; Gallego del Sol, F., Chornet, J. J. C., & Cavada, B. S. PDBcode: 2GDF; to be published), *Dioclea lehmanii* lectin (DLLI) [4], *Canavalia gladiata* lectin (CGL) [38], ConA [30], ConBr [34], *Acacia farnesiana* lectin-like protein (AFAL) [39], PHAL [40], PHA [36], *Dolichos biflorus* lectin (DBL) [41], VML [32], Rob [33], and *Dolichos lablab* lectin (Dlab) [42]. The alignment and phylogram was generated using CLUSTALW web program

translational process seems to be a plesiomorphy character between the group of CRLII and the others groups (Fig. 2; Gallego del Sol, F., Chornet, J. J. C., & Cavada, B. S. PDBcode: 2GDF; to be published) [4, 15, 32–42]. A sequence alignment of the sequenced peptides of CRLII against other legume lectin sequence displays a similarity of 53% and 44% with the galactose-specific lectins VML and Rob, respectively (Fig. 3).

```
CRLII ----------------------------------------------DLTLQGSALVSSK 13
VML   ------------------------------SEVVSFSFTKFNPNPKDIILQGDALVTSK 29
Rob   MASYKFKTQNSFLLLLSISFFFLLLLNKVNSTGSLSFSFPKFKHSQPDLIFQSDALVTSK 60
                                                    *: :*..***:**

CRLII ----------GKPVDHSVGRALYRSPIHIWDETTGKVASFDATFSFVSEAPAIPMLFPSS 63
VML   GKLQLTKVKDGKPVDHSLGRALYAAPIHIWDDSTDRVASFATSFSFVVEAPDESKTADGI 89
Rob   GVLQLTTVNDGRPVYDSIGRVLYAAPFQIWDSTTGNVASFVTSFSFIIKAPNEGKTADGL 120
               *:** .*:**.** :*::***.:*..**** ::***: :**       .

CRLII KGELNPP-DTQPDRIGG-QGVVNDS--YNVIRTVAVENDGYRN-RVDPSARHXSLPKLKI 118
VML   AFFLAPP-DTQPQKDGGFLGLFNDS--NKSIQTVAVEFDTFSN-TWDPSARHIGINVNSI 145
Rob   VFFLAPVGSTQPLKGGGLLGLFKDESYNKSNQIVAVEFDTFRNVAWDPNGIHMGIDVNSI 180
         * *  .*** : **  *:.:*.   :  : **** * : *   **.. * .:   .*

CRLII AKNNGIKWNMQTGK-------------VGTAHXSYNSVAKRYDVXXNAVLPSKV------ 159
VML   ESMKYVKWGWENGKVANVYISYEASTKTLTASLTYPSNATSYIVSANVDLKSALPEWVRV 205
Rob   QSVRTVRWDWANGEVANVFISYEASTKSLTASLVYPSLEKSFILSAIVDLKKVLPEWVRV 240
        . . ::*.  . *.            .:**   * *  . : :   : * . :

CRLII ---ATSGXYKDHVETNTXXSWSFTSKXTNSDSN------------- 189
VML   GFSATSGLSRDHVETHDVLDWSFTSTLQAPSDDSN----------- 240
Rob   GFTATTGLSEDYVQTNDVLSWSFESNLPGGNSVASVKNAGLSTYAA 286
         **:* .*:*:*:  .*** *.  ..:.
```

Fig. 3 Sequence alignment of the partial CRLII sequence obtained by mass spectroscopy with the seed lectins of VML and Rob. The peptides can be observed located in their likely position in the sequence. The X residues can be either isoleucine or leucine. The alignment was performed using CLUSTALW web program

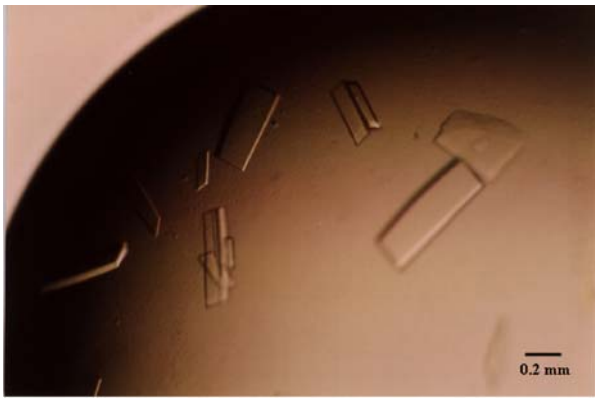

Fig. 4 Crystals of the lactose-specific lectin CRLII from *Cymbosema roseum*

Protein Crystallization and X-ray Data Collection

Microcrystals were obtained using neutral buffers containing different molecular-weight PEGs. After optimization, diffraction-suitable crystals were obtained from drops containing 8% (ethylene glycol), 0.1 M TRIS-HCl, pH 8.4 to 8.6 and 10% to 11% PEG 8,000. CRLII crystals grew within a month to maximum dimensions of approximately 0.8×0.4×0.1 mm (Fig. 4). The CRLII crystals are monoclinic, and belong to the $P2_1$ space group with unit cell parameters of a=49.4, b=89.6, and c=100.8 Å. CRLII crystal data were scaled in a resolution range of 48.7–2.5 Å. Statistics of the data collection are in Table 3. Assuming the presence of four molecules of 26 kDa for each monomer in the asymmetric unit, the calculated V_M is 2.1 $Å^3$ Da^{-1} [43], indicating a solvent content of 40.2%. The crystal structure will be solved by molecular replacement (MR) and MAD experiments also are been developed to overcome the difficult to find good search models to MR.

Conclusion

The crystallized lectin CRLII displays a high sequence similarity with galactose-binding plant lectins which is confirmed by its mucine glycoproteins affinity. This is an important difference when compared with CRLI. These lectins also differ at posttranslational process,

Table 3 Statistics of X-ray data collected for the CRLII crystal.

X-ray wavelength (Å)	1.427
Space group	$P2_1$
Unit-cell parameters (Å)	a=49.4
	b=89.3
	c=100.5
Resolution limits (Å)	48.7–2.5
Asymmetric unit content	4 molecules
Total reflections measured	280,204
Unique reflections measured	28,311
Completeness (%)	96.3 (96.3)
R_{merge} (%)	11.7 (48.2)
$\langle I/\sigma(I)\rangle$	5.4 (1.4)

Values in parenthesis are for the highest resolution shell.

which produces distinct chains in CRLI but does not seem to occur in CRLII, and phylogenetic analysis shows paraphyletic groups in legume lectins. The structural doubts about these aforementioned aspects are the purpose that led to crystallize these proteins and solve their structures, a work that is still in progress using molecular replacement methods to both structures and MAD phasing strategies to CRLII.

Acknowledgments We thank Dr. Juan J. Calvete from Instituto de Biomedicina de Valencia (Spain) for his scientific support and criticism, which improved this work substantially. Mass Spectrometry Laboratory and MX1 beamline/Brazilian Synchrotron Light Laboratory (LNLS), Campinas, Brazil. BSC, AHS, and WFA are senior investigators of CNPq. We also thank Dr. A. Leyva for English language editing of the manuscript.

References

1. Peumans, W. J., & Van Damme, E. J. (1995). *Plant Physiology*, *109*, 347–352. doi:10.1104/pp.109.2.347.
2. Sharon, N., & Lis, H. (2003). Lectins. Dordresh: Kluwer Academic.
3. Gowda, L. R., Savithri, H. S., & Rao, D. R. (1994). *The Journal of Biological Chemistry*, *269*, 18789–18793.
4. Perez, G., Hernandez, M., & Mora, E. (1990). *Phytochemistry*, *29*, 1745–1749. doi:10.1016/0031-9422(90)85007-3.
5. Perez, G. (1998). *The International Journal of Biochemistry & Cell Biology*, *30*, 843–853. doi:10.1016/S1357-2725(98)00045-4.
6. Hamelryck, T. W., Moore, J. G., Chrispeels, M. J., Loris, R., & Wyns, L. (2000). *Journal of Molecular Biology*, *299*, 875–883. doi:10.1006/jmbi.2000.3785.
7. Gadelha, C. A. A., Moreno, F. B. M. B., Santi-Gadelha, T., Cajazeiras, J. B., Rocha, B. A. M., Assreuy, A. M. S., et al. (2005). *Journal of Structural Biology*, *152*, 185–194. doi:10.1016/j.jsb.2005.07.012.
8. Carrington, D. M., Auffret, A., & Hanke, D. E. (1985). *Nature*, *313*, 64–67. doi:10.1038/313064a0.
9. Moraes, S. M., Cavada, B. S., Moreira, R. A., Roque-Barreira, M. C., Santos-de-Oliveira, R., Pinto, V. P., et al. (1996). *Brazilian Journal of Medical and Biological Research*, *29*, 977–985.
10. Saper, M. A., Lis, H., Sharon, N., & Shaanan, B. (1987). *Journal of Molecular Biology*, *193*, 823–824. doi:10.1016/0022-2836(87)90363-9.
11. Iglesias, J. L., Lis, H., & Sharon, N. (1982). *European Journal of Biochemistry*, *123*, 247–252. doi:10.1111/j.1432-1033.1982.tb19760.x.
12. Konozy, E. H., Mulay, R., Faca, V., Ward, R. J., Greene, L. J., Roque-Barriera, M. C., et al. (2002). *Biochimie*, *84*, 1035–1043. doi:10.1016/S0300-9084(02)00003-2.
13. Konozy, E. H., Bernardes, E. S., Rosa, C., Faca, V., Greene, L. J., & Warda, R. J. (2003). *Archives of Biochemistry and Biophysics*, *410*, 222–229. doi:10.1016/S0003-9861(02)00695-1.
14. Silva, J. A., Damico, D. C., Baldasso, P. A., Mattioli, M. A., Winck, F. V., Fraceto, L. F., et al. (2007). *The Protein Journal*, *26*, 193–201. doi:10.1007/s10930-006-9061-0.
15. Cavada, B. S., Marinho, E. S., Souza, E. P., Benevides, R. G., Delatorre, P., Souza, L. A., et al. (2006). *Acta Crystallographica*, *F62*, 235–237.
16. Laemmli, U. K. (1970). *Nature*, *227*, 680–685. doi:10.1038/227680a0.
17. Faria, R. A., Andrade-Neto, M., Pinto, L. S., Castellon, R. R., Calvette, J. J., & Cavada, B. S. (2004). *Archivos Latinoamericanos de Nutricion*, *54*, 349–353.
18. Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). *Nucleic Acids Research*, *22*, 4673–4680. doi:10.1093/nar/22.22.4673.
19. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). *Nucleic Acids Research*, *25*, 3389–3402. doi:10.1093/nar/25.17.3389.
20. Leslie, A. G. W. (1992). Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, 26.
21. Collaborative Computational Project. (1994). Number 4. *Acta Crystallographica. Section D, Biological Crystallography*, *50*, 760–763. doi:10.1107/S0907444994003112.
22. Van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Kochetkov, K., Arbatsky, N. P., & Derevitskaya, V. A. (1982). *European Journal of Biochemistry*, *127*, 21–29. doi:10.1111/j.1432-1033.1982.tb06832.x.

23. Zenteno, E., Vasquez, R., Cordoba, F., Wieruszeski, J. M., Montreuil, J., & Debray, H. (1995). *Glycoconjugate Journal*, *12*, 699–706. doi:10.1007/BF00731267.
24. Karlsson, N. G., Nordman, H., Karlsson, H., Carlstedt, I., & Hansson, G. C. (1997). *The Biochemical Journal*, *326*, 911–917.
25. Savage, A. V., Donohue, J. J., Koeleman, C. A., & Van Den Eijnden, D. H. (1990). *European Journal of Biochemistry*, *193*, 837–843. doi:10.1111/j.1432-1033.1990.tb19407.x.
26. Chai, W., Hounsell, E. F., Cashmore, G. C., Rosankiewicz, J. R., Feeney, J., & Lawson, A. M. (1992). *European Journal of Biochemistry*, *207*, 973–980. doi:10.1111/j.1432-1033.1992.tb17132.x.
27. Svensson, C., Teneberg, S., Nilsson, C. L., Kjellberg, A., Schwarz, F. P., Sharon, N., et al. (2002). *Journal of Molecular Biology*, *321*, 69. doi:10.1016/S0022-2836(02)00554-5.
28. Elgavish, S., & Shaanan, B. (1998). *Journal of Molecular Biology*, *277*, 917–932. doi:10.1006/jmbi.1998.1664.
29. Ravishankar, R., Thomas, C. J., Suguna, K., Surolia, A., & Vijayan, M. (2001). *Proteins*, *43*, 260–270. doi:10.1002/prot.1037.
30. Hague, D. R. (1975). *Plant Physiology*, *55*, 636–642.
31. Cunningham, B. A., Hemperly, J. J., Hopp, T. P., & Gerald, M. (1979). *Proceedings of the National Academy of Sciences of the United States of America*, *76*, 3218–3222. doi:10.1073/pnas.76.7.3218.
32. Calvete, J. J., Santos, C. F., Mann, K., Grangeiro, T. B., Nimtz, M., Urbanke, C., et al. (1998). *FEBS Letters*, *425*, 286–292. doi:10.1016/S0014-5793(98)00243-9.
33. Yoshida, K., & Tazaki, K. (1999). *Australian Journal of Plant Physiology*, *26*, 495–502.
34. Sanz-Aparicio, J., Hermoso, J., Grangeiro, T. B., Calvete, J. J., & Cavada, B. S. B. (1997). *FEBS Letters*, *405*, 114–118. doi:10.1016/S0014-5793(97)00137-3.
35. Naismith, J. H., & Field, R. A. (1996). *The Journal of Biological Chemistry*, *271*(2), 972–976. doi:10.1074/jbc.271.2.972.
36. Hamelryck, T. W., Dao-Thi, M. H., Poortmans, F., Chrispeels, M. J., Wyns, L., & Loris, R. J. (1996). *Biological Chemistry*, *271*, 20479–20485. doi:10.1074/jbc.271.34.20479.
37. Calvete, J. J., Thole, H. H., Raida, M., Urbanke, C., Romero, A., Grangeiro, T. B., et al. (1999). *Biochimica et Biophysica Acta*, *1430*(2), 367–375.
38. Delatorre, P., Rocha, B. A., Souza, E. P., Oliveira, T. M., Bezerra, G. A., Moreno, F. B., et al. (2007). *BMC Structural Biology*, *7*, 52. doi:10.1186/1472-6807-7-52.
39. Santi-Gadelha, T., Rocha, B. A. M., Gadelha, C. A. A., Toyama, M. H., Pinto, V. P., Nagano, C. S. et al. (2008). *Applied Biochemistry and Biotechnology*, *150*, 97–111
40. Lioi, L., Sparvoli, F., Galasso, I., Lanave, C., & Bollini, R. (2003). *Theoretical and Applied Genetics*, *107*, 814–822. doi:10.1007/s00122-003-1343-8.
41. Hamelryck, T. W., Loris, R., Bouckaert, J., Dao-Thi, M. H., Strecker, G., Imberty, A., et al. (2000). *Journal of Molecular Biology*, *286*, 1161–1177. doi:10.1006/jmbi.1998.2534.
42. Colucci, G., Moore, J. G., Feldman, M., & Chrispeels, M. J. (1999). *Proceedings of the National Academy of Sciences of the United States of America*, *96*(2), 646–650. doi:10.1073/pnas.96.2.646.
43. Matthews, B. W. (1968). *Journal of Molecular Biology*, *33*, 491–497. doi:10.1016/0022-2836(68)90205-2.