Purification and characterization of a highly thermostable GlcNAc-binding lectin from *Collaea speciosa* seeds

Messias V. Oliveira [1], Vinicius J. S. Osterne [1,4], Claudia F. Lossio [1], Sonia Serna [2], Niels C. Reichardt [2,3], Kyria S. Nascimento [1], Els J. M. Van Damme [4,*], Benildo S. Cavada [1,*]

[1] BioMol-Lab, Campus do Pici, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil 60.440-970.

[2] Glycotechnology Laboratory, Center for Cooperative Research in Biomaterials (CIC biomaGUNE), Basque Research and Technology Alliance (BRTA), Paseo Miramón 182, 20014 San Sebastian, Spain.

[3] CIBER-BBN, Paseo Miramón 182, 20009 San Sebastian, Spain.

[4] Laboratory for Biochemistry and Glycobiology, Department of Biotechnology, Ghent University, 9000 Ghent, Belgium.

* Corresponding authors: Benildo S. Cavada (bscavada@ufc.br) / Els J.M. Van Damme (ElsJM.VanDamme@ugent.be)

ABSTRACT

Lectins from plants of the Diocleinae subtribe often exhibit specificity towards mannose/glucose and derived sugars, with some plants also displaying a second lectin specific to lactose/GalNAc. Here, we present a novel lectin from *Collaea speciosa*, named CsL, that displays specificity for GlcNAc/glucose. The lectin was extracted from *Collaea speciosa* seeds and purified by a single chromatographic step on a Sephadex G-50 matrix. In solution, the lectin appears as a dimeric protein composed of 25 kDa monomers. The protein is stable at pH 7-8 and dependent on divalent cations. CsL maintained its agglutination activity after heating to 90 ºC for one hour. Glycan array studies revealed that CsL binds to *N*-glycans with terminal GlcNAc residues, chitobiose and chitotriose moieties. The partial amino acid sequence of the lectin is similar to that of some lactose-specific lectins from the same subtribe. In contrast to other ConA-like lectins, CsL is not toxic to Artemia. Because of its remarkably different properties and specificity, this lectin could be the first member of a new group inside the Diocleinae lectins.

Keywords: Lectin; *Collaea speciosa*; Biological properties.

1. INTRODUCTION

Lectins constitute a select group of proteins found in all organisms. They are defined as proteins of nonimmune origin capable of interacting with carbohydrates in a specific and reversible way. Carbohydrates are one of the most abundant biomolecules in nature. Carbohydrate units can form glycans that can be attached to glycoproteins and glycolipids to play important roles in many biological processes. In particular, some proteins that interact with carbohydrates/glycans can decipher the information encoded in glycosylation patterns, more commonly known as the glycocode [1–3]. Although the carbohydrate-binding properties of lectins are the most studied, some plant lectins can also interact with small molecules that are predominantly hydrophobic [4]. For instance, some plant lectins bind to porphyrins and phthalocyanines, which are photosensitizers used in clinical imaging for cancer cell detection with particular application in photodynamic therapy (PDT) [5,6]. Studies with lectins from the Diocleinae subtribe reported that they possess a hydrophobic site capable of interacting with hydrophobic molecules, such as 2-p-toluidinonaphthalene-6-sulfonate, hydrophobic sugar derivatives, phytohormones and auxin (e.g., indole-3-acetic acid or IAA, a naturally occurring auxin) [7].

The best-studied plant lectins are those from the *Leguminosae* family, specifically the Papilionoideae subfamily and the Diocleinae subtribe from which Concanavalin A is considered the model lectin. The Diocleinae subtribe contains 13 different genera of plants, including *Canavalia, Dioclea, Cymbosema, Camptosema, Cratylia, Galactia, Cleobulia, Acropsychanthus, Neorudolphia, Lackeya, Luzonia, Collaea*, and *Rhodopis*. Judging from the number of publications, the first five genera have been studied extensively, but most of the other genera have been largely neglected. Several legume lectins have been extracted, purified, and characterized from plants of the Diocleinae subtribe. These

lectins often exhibit specificity towards mannose/glucose and related carbohydrates. Some plants even present a second lectin specific to lactose/GalNAc. Several biological applications have been reported for these lectins. For instance, ConA-like lectins display inflammatory, nociceptive, antiproliferative, and antimicrobial activities [8,9]. Although these lectins resemble each other insofar as their molecular structures and properties, their biological activities can be very different [10].

Collaea DC. is a tropical and subtropical genus, exclusive to South America. At present, the floristic survey and classification of species within this genus are still imprecise. Records indicate up to seven *Collaea* species distributed in Peru, Bolivia, Brazil, Paraguay, Argentina and Uruguay [11,12]. One of these species, *Collaea speciosa*, the subject of this study, is mostly found in southern Brazil. Flowering and fruiting occur between October and February and, more rarely, in May, July and August. The fruit is a linear pod 5 x 2 cm, varying in coat and color. Seeds are 3 to 4 mm, flattened laterally, and often yellowish. Seeds are dispersed by gravity. *Collaea* is classified as a shrub, and it can grow up to 3 m in height, branches erect, to slightly curved, at the apex. The leaves are trifoliate, and the axillary inflorescences are terminal with the corolla generally showy and red (Figure S1) [12–14]. The plant grows mainly in closed and rocky fields, shaded places, forest edges, or even in gallery forests, but also on road sides and degraded environments. It is well adapted to cold and resistant to frost [15].

In this work, we report on the purification of a lectin from seeds of C. speciosa (*Leguminosae*, Papilionoideae subfamily, Diocleinae subtribe), further referred to as *C. speciosa* lectin, or CsL. The physicochemical properties and carbohydrate-binding specificity of CsL were studied in detail, and its partial amino acid sequence was determined.

2. MATERIALS AND METHODS

2.1. Reagents and plant material

Mature seeds from *C. speciosa* were obtained from plants grown at the Federal University of Ceará, Brazil. Reagents of ultrapure grade were obtained from Merck (St. Louis, MO) and GE Healthcare (Little Chalfont, UK).

2.2. Protein extraction

Seeds of *C. speciosa* were peeled, ground until a fine flour in an electric grinder, and the soluble proteins were extracted by mixing the powder with 150 mM NaCl (1:10 w/v) under constant stirring for 4 h at room temperature. The crude extract was clarified by centrifugation at 10,000 x *g* for 20 min at 4 ºC, followed by filtration through filter paper. Protein content was quantified by applying the Bradford method using bovine serum albumin (BSA) as a standard [16]. Quantification results were expressed as milligram (mg) protein per milliliter (mL) of extract (mgP/mL).

2.3. Hemagglutinating activity (HA), sugar-inhibition assays, and CsL purification

Lectin activity was detected by hemagglutination assays, following the protocol described in Cavada et al. [17]. Aliquots of 50 µL for each sample were 2-fold diluted in a 96-well plate containing 50 µL saline (150 mM NaCl, 5 mM $CaCl_2$, 5 mM $MnCl_2$), after which 50 µL 3% erythrocytes (rabbit and human - A, B, and O cells), both native and treated with proteolytic enzymes, were added. A well containing only saline solution and erythrocytes was employed as a control. The plates were incubated at 37 °C for 30 minutes and then at 25 °C for an additional 30 minutes. Agglutination was visualized macroscopically, and the hemagglutination titer was calculated as the reciprocal of the

highest dilution presenting agglutination per mL (Hemagglutination Units/mL, H.U./mL). Specific activity was calculated as Hemagglutination Units (H.U.) per milligram of protein (H.U./mgP).

Sugar-inhibition assays were conducted using D-glucose, D-galactose, L-rhamnose, D-ribose, D-mannose, α-lactose, α-methyl-D-mannoside-, L-fucose, deoxyribose, N-acetylglucosamine (GlcNAc), and ovalbumin prepared in 150 mM NaCl at 100 mM concentration for the sugars and 2 mg/mL for the glycoprotein. For this, the sugars and glycoproteins were dissolved in NaCl 150 mM in a concentration of 100 mM, and then serial dilutions (1/2, 1/4, 1/8 and so on) of the carbohydrate and glycoprotein solutions were added to a 96-well plate, each well containing 25 µl 150 mM NaCl. Subsequently, 25 µl of the clarified extract at 4 HU/mL were added to each well. The plate was incubated for 30 minutes at 37 °C and then another 30 minutes at 25 °C. Finally, 50 µl of 3% rabbit erythrocytes were added to each well. After a 30 min incubation at 37 ºC, hemagglutination was evaluated macroscopically, and the minimal concentration of carbohydrate or glycoprotein capable of inhibiting hemagglutination, i.e., minimal inhibitory concentration (MIC), was determined. MIC was expressed as the reciprocal of the highest dilution that completely inhibited hemagglutination after 1 h. Saline and sugars/glycoproteins-only were employed as controls.

Taking into account the results of the sugar-inhibition assays, the purification of CsL was performed following the protocol described for most ConA-like lectins [8]. The extract was subjected to affinity chromatography on a Sephadex® G-50 matrix, pre-equilibrated with 150 mM NaCl containing 5 mM $CaCl_2$ and 5 mM $MnCl_2$. The extract was kept in contact with the matrix for 4 hours to increase the final yield of lectin. Afterwards, the non-retained material was washed with the extraction solution, followed by elution of the bound fraction with 100 mM glycine, pH 2.6, containing 150 mM NaCl. Fractions of 1 mL

were collected and monitored by spectrophotometry at 280 nm. The retained protein was dialyzed against distilled water and freeze-dried.

2.4. Physicochemical characterization

Thermostability and pH stability of CsL, as well as divalent cation dependence of CsL activity, were examined using hemagglutination assays. A 1 mg/ml lectin solution in 150 mM NaCl was used in all assays. For thermostability analysis, the lectin solution was heated for one hour at different temperatures, ranging from 20 to 100 °C, with 10 °C intervals. Immediately afterwards, hemagglutination assays with rabbit erythrocytes were performed. CsL activity was investigated at different pH. Therefore, 1 mL of lectin solution was dialyzed against 100 mL of different buffers, including 100 mM glycine-HCl, pH 3.0; 100 mM sodium citrate, pH 4.0; 100mM sodium acetate, pH 5.0; 100 mM sodium citrate, pH 6.0; 20 mM sodium phosphate, pH 7.0; 100 mM Tris HCl, pH 8.0; and 100 mM glycine-NaOH, pH 9.0 or 10.0, all containing 150 mM NaCl. After dialysis for one hour, hemagglutination assays were performed. To verify whether CsL is dependent on divalent cations, the lectin solution was dialyzed against 150 mM NaCl containing 100 mM EDTA for 24 h, followed by an additional dialysis step against 150 mM NaCl to remove excess EDTA. Hemagglutination assays were then carried out in the presence and absence of $CaCl_2$ and $MnCl_2$, both at 10 mM [17].

Hemagglutinating activity was also evaluated after incubation of CsL with proteolytic enzymes (papain and trypsin). For this, a 1 mg/mL solution of CsL was mixed with each protease (both at 1 mg/mL in 150 mM NaCl), as described in the protocols for each enzyme. Hemagglutination assays were performed after 2 hours of digestion.

2.5. Molecular weight calculation

The apparent molecular weight of CsL was determined using size exclusion liquid chromatography coupled to an ultra-performance system (UPLC). Commercial kits (GE Healthcare) with low and high molecular weight proteins were used to make the calibration curve: ribonuclease A (13,700 Da), carbonic anhydrase (29,000 Da), ovalbumin (44,000 Da), and conalbumin (75,000 Da). The applied matrix was BioSuite ™ 250, 5 µm HR SEC (7.8 x 300 mm) with its volume (V_0) determined by Blue Dextran 2000. The elution volume (V_e) of all samples was calculated from the start of the injection to the center of peak intensity. Column volume (V_c) was calculated as $V_c = \pi \times r^2 \times h$. The calibration curve ($K_{av} = -14.43 \ln(x) + 23.339$) was obtained based on calculating the partition constant (K_{av}) versus the logarithm of the molecular weight of the proteins. The partition constant was calculated as $K_{av} = V_e - V_0 / V_c - V_0$. After obtaining the partition constant, the calibration curve equation with $R^2 = 95\%$ was applied, and the molecular weight was calculated.

2.6. Electrophoresis and PAS staining

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to the method described by Laemmli [18] with some modifications. The acrylamide concentration for the separation gel and stacking gel was 12 % and 4 %, respectively. Lyophilized protein fractions were solubilized at 4 mg/mL in sample buffer consisting of 62.5 mM Tris-HCl, pH 6.8, containing 2% SDS and 12.5% glycerol, followed by heating at 100 °C for 5 min. For the electrophoretic run, 15 µl of each sample were loaded. The presence of disulfide bridges was evaluated by addition of 2% 2-mercaptoethanol to the protein samples.

SDS-PAGE was performed on a Mini-PROTEAN II mini-gel system (Bio-Rad; Milan, Italy) with an electrical current of 25 mA. The running buffer consisted of 25 mM Tris-HCl, 192 mM Glycine and 0.1% SDS, pH 8.8. After electrophoresis, the separation gel was stained with Coomassie Brilliant Blue G-250. The molecular marker MM-V849A (Promega) with twelve polypeptides, ranging from 10 to 225 kDa, was used to estimate the apparent molecular weight of CsL. SDS-PAGE was followed by periodic acid Schiff (PAS) staining to show any presence of glycans in the lectin's structure [19].

2.7 Amino acid sequence analysis

To characterize the primary structure of CsL, electrophoresis was first performed in denaturing conditions on a 12% polyacrylamide gel (SDS-PAGE). Afterwards, the 25 and 50 kDa bands were excised from the gel, bleached with 100 mM ammonium bicarbonate in acetonitrile (1:1 v/v), and dried in acetonitrile. For each molecular weight range, enzymatic digestion with trypsin, pepsin or chymotrypsin (Promega, Madison, WI, USA) in 50 mM ammonium bicarbonate was performed, and the obtained peptides were extracted in 5% formic acid in 100% acetonitrile (1:2 v/v) and dried in a SpeedVac [20].

The dried peptides were solubilized in 0.1% trifluoroacetic acid (TFA) and separated on a BEH300 C18 column (100 μm × 100 mm) using the nanoAcquity ™ system (Waters) connected to the nanoelectrospray mass spectrometer (SYNAPT HDMS System). Elution was performed with an acetonitrile gradient (10 to 85%) containing 0.1% formic acid at 600 μL/min. The configuration of the mass spectrometer was as follows: positive mode, source temperature of 80 °C, and capillary voltage of 3.5 kV. The calibration was performed with fragments of the double protonated ion [Glu1] - fibrinopeptide B (Sigma-Aldrich) (m/z 785.84). Afterwards, the double or triple charge precursor ions were selected and fragmented

by collision-induced dissociation (CID) using argon as the collision gas and ramping collision energy, varying according to the charge state of the selected precursor ion. The data were processed and analyzed using ProteinLynx (Waters), and the search parameter was the peptide fragmentation pattern. The CID spectra were interpreted manually using the Peptide Sequencing tool of MassLynx 4.1 (Waters) software [17].

The obtained CsL sequence was submitted to BLASTp against the non-redundant protein database [21], and a multiple alignment was performed by combining ClustalOmega [22] and ESPript3 [23].

2.8 Glycan array analyses

The fine carbohydrate-binding specificity of CsL was determined by glycan array screening. Synthetic glycan microarrays were prepared as previously described [24]. The representation of each tested glycan is depicted in Figure S2. CsL was fluorescently labeled with Alexa Fluor™ 555 succinimidyl ester (NHS) (Fisher Scientific) following the manufacturer's instructions. Excess of labeling reagent was removed by employing an Amicon 10 kDa filter. Labeled CsL was diluted in 25 mM Tris-HCl buffer with 150 mM NaCl, pH 7.5 (TSM), containing 4 mM $CaCl_2$ to A_{280nm} = 0.2 and incubated on the glycan microarray in the dark at room temperature for one hour. The glycan array was washed with TSM containing 0.01% Tween 20. The slide was dried by centrifugation in a slide spinner, and fluorescence was analyzed on an Agilent G265BA microarray scanner (PerkinElmer). Average values of fluorescence (mean $\pm$ SD from four spots with local background value removed) are reported as histograms. Results are shown as a graph of relative fluorescence units in which a higher fluorescence indicates a higher binding specificity.

2.9 Toxicity of CsL for *Artemia* sp. nauplii

Seawater collected at Paracuru Beach, Ceará, Brazil, was sterilized and used according to saline water quality standards, Class I (Section III, Article 18, Item 1 J of Resolution 357/05, Conselho Nacional do Meio Ambiente (CONAMA)) [25]. Artemia cysts were incubated under lighting and constant aeration. Incubation was performed in glass flasks with 1 g of cysts per liter of seawater. After 48 hours, aeration was interrupted, and only highly motile Artemia nauplii were collected for trials [17].

The lethality assay was performed in a 24-well Limbro plate, each well containing a final volume of 2 mL. CsL solutions (3.125 - 200 µg/mL) solubilized in seawater were incubated with 10 nauplii in each well. The experiment was carried out in triplicate, and the results were compared with the negative control group (without lectin) and with the positive control group (with ConBr). After 24 and 48 hours, the death toll was counted. For this test, toxicity was represented as the concentration that kills 50% of nauplii.

3. RESULTS AND DISCUSSION

3.1 Extraction and purification of CsL

The presence of a lectin in the crude extract of *C. speciosa* seeds was detected by hemagglutination of rabbit erythrocytes. For all types of rabbit blood cells, either nontreated or treated with papain and trypsin, the hemagglutination titer was similar, while no agglutination activity was observed with human erythrocytes. CsL was purified using a single chromatography step (Figure 1A). The single protein peak eluted from the Sephadex G-50 matrix displayed a high purification factor and yield, as shown in Table 1. This study reports the first lectin purified and characterized from a plant of the genus *Collaea*. Previously,

Cavada et al. [8] had standardized the extraction and purification protocol for other ConA-like lectins.

Some mannose/glucose-specific lectins from the Phaseoleae tribe have been purified by applying very similar protocols. The lectins from *Canavalia* species, such those from *C. grandiflora* [27], *C. virosa* (ConV) [28], *C. gladiata* [29], *C. oxyphylla* (CoxyL) [30], *C. maritima* [31], *C. bonariensis* (CaBo) [32], *C. boliviana* [33] and *C. villosa* [34], were purified by affinity chromatography. In fact, the vast number of lectins already reported in the Diocleinae tribe could be purified using a purification scheme consisting of only a few purification steps, such as binding to a Sephadex G-50 matrix. These data confirm the high specificity of these lectins for their ligands. Lectins have been purified from several species within the genus *Dioclea*. Most of these lectins show affinity for D-mannose/D-glucose and have been purified following procedures similar to those described above [8].

However, not all lectins from the Diocleinae subtribe can be purified using the same protocols. For example, purification of the mannose/glucose-specific lectins from *Cratylia mollis* (Iso1, Iso 2 and Iso 3) [35,36], *Cymbosema roseum* (CRL I and CRL II) [37,38] and lectins from *Camptosema pedicellatum* [39], as well as the lactose-specific lectin from *Camptosema ellipticum* [40], all required multiple chromatographic steps, including ion exchange chromatography and gel filtration, in addition to the affinity chromatography. Other lectins of the Diocleinae subtribe, such as the lectins purified from the genus *Galactia*, can also be obtained using simple purification protocols consisting only of affinity chromatography. However, these lectins are specific to lactose/GalNAc [41–43].

3.2 Lectin characterization

Like most lectins from the Diocleinae subtribe, CsL shows affinity for glucose and its derivatives. Inhibition assays also demonstrated the inhibition of agglutination by ovalbumin, most likely a result of the interaction between CsL and the complex glycans present on ovalbumin. CsL also exhibited a strong affinity for GlcNAc with MIC of approximately 0.2 mM (Table 2). Since the specificity of CsL for GlcNAc is very different from that of other lectins purified from species of the same subtribe, CsL could represent a third group of Diocleinae lectins, quite apart from mannose-specific and lactose-specific lectins.

Most ConA-like lectins are specific for mannose, glucose, and derived monosaccharides, without binding to galactose, fucose or sialic acid. It was reported that the hydroxyl group C-2 of mannose is not essential for binding, whereas C-3, C-4, and C-6 are important for carbohydrate interaction with the lectin [44,45]. For methyl α-D-mannopyranoside, the most specific monosaccharide for these mannose/glucose-binding lectins, the free binding energy varies from -4.4 to -5.62 kcal/mol [8,44,45]. Previous studies also demonstrated that ConA-like lectins can interact with glycoproteins, a property also reported for *Phaseolus vulgaris* erythroagglutinin (PHA-E) and soybean agglutinin (SBA), which showed interaction with arcelin-1, ovalbumin, orosomucoid α-glycoprotein, bovine lactotransferrin and human serotransferrin [46].

Differences in the carbohydrate-binding specificity of lectins are associated with small variations at sequence level. These small differences in lectin sequences have been linked to changes in biological activities, such as those involving proinflammatory processes [35,36], antibacterial activity [37], insecticidal activity [39], or antitumor activity [40].

The apparent molecular weight of CsL was estimated using molecular exclusion chromatography and SDS-PAGE. After electrophoresis under denaturing conditions, CsL

presented a profile consisting of two polypeptide bands, one more intense band around 25 kDa, representing the lectin monomer, and another less intense band, around 50 kDa, indicating the formation of a dimer (Figure 1BC). Under native conditions, CsL migrated with an approximate molecular weight of 56.23 kDa, indicating the formation of dimers (Figure 2). Most likely, the denaturation conditions used for the SDS-PAGE were insufficient to dissociate the dimers completely [47]. Alternatively, the largest band in the gel could represent the mature lectin polypeptide (50 kDa) which is partially cleaved into fragments of 25 kDa during post-translational processing, thus resulting in the double band on SDS-PAGE. Cleavage of lectin polypeptides can be observed in other Legume lectins, such as those from *Vatairea guianensis* and *V. macrocarpa* [48,49]. Future works aiming to study the molecular cloning and structural analysis of the lectin can shed light on lectin processing. According to the data presented in Table 3, Phaseoleae lectins usually present an electrophoretic profile with polypeptides ranging from 25 to 30 kDa. Even though CsL is a Diocleineae lectin, its SDS-PAGE profile makes it clear that CsL does not suffer the circular permutation processing common to several other lectins of this group [8,27,32,33].

When the hemagglutinating activity was analyzed at different pH, CsL displayed maximal agglutination activity at pH 7 and 8 (Figure 3A). CsL is thermostable when heated at temperatures between 20 and 90 ºC for one hour. However, when heated to 100 ºC, 50% of its activity was lost, and at 105 ºC, its activity disappeared almost completely (Figure 3B). Like most lectins purified from species belonging to the Phaseoleae tribe, lectin activity remains in the range of pH 5-8, with maximal activity in the range of pH 7-8. CsL and related lectins show a different thermostability. Legume lectins normally maintain their maximum activity up to 60 °C, after which activity is reduced because of protein denaturation at temperatures of approximately 70–80 °C [50]. At this point, a total loss of hemagglutinating

activity occurs (Table 3). CsL is more thermostable than other Diocleinae lectins, only losing its hemagglutinating activity after heating at 100 ºC. The combined high prevalence of cysteine residues in the partial sequence (approximately 16%) and the presence of residues with nonpolar aliphatic and aromatic R groups (approximately 60%) in the partial sequence of CsL can suggest a possible mechanism for its stability. Cysteine residues contribute to the formation of S-S bridges, and hydrophobic amino acids can form a compact three-dimensional structure of the lectin through hydrophobic interactions, which, in turn, can increase lectin thermostability.

The hemagglutinating activity of CsL was lost after dialysis against 100 mM EDTA. After addition of Ca^{2+} or Mn^{2+} ions, a partial restoration of the activity was observed, while full activity was recovered only after the addition of both divalent cations. These results show that CsL, similar to most Diocleinae lectins, is dependent on metal ions, such as Ca^{2+} and Mn^{2+} (Figure 3CD and Table 3) [7,51]. Judging from the absence of PAS staining CsL, does not have any covalently attached glycan in its structure in contrast to the positive control lectin (VML) (Figure 4) [52,53]. Finally, treatment of CsL with the proteolytic enzymes papain and trypsin completely abolished its hemagglutinating activity, confirming that the hemagglutination activity is indeed caused by the lectin [54,55].

3.3 Amino acid sequence

Sequence information for six peptides of CsL (25 kDa polypeptide) was obtained by tandem mass spectrometry, identifying a total of 82 amino acids (Table 4), corresponding to approximately 34% of the Camptosemin sequence used as a template. BLASTp against the nonredundant database of NCBI showed that CsL peptides are very different from those of other ConA-like lectins, but that they still share a very high sequence similarity with

Camptosemin from *Camptosema ellipticum* (Figures 5 and 6). Sequence alignments of the partial sequence for CsL and other lectin sequences suggest that CsL does not share a high degree of identity with other ConA-like lectins, except for Camptosemin (69.62% identity). CsL peptides show a fair degree of sequence identity with PNA (NCBI id: XP_025621342.1; 41.79%), *Medicago truncatula* lectin (GenBank: KEH33924.1; 41.56%), *Robinia pseudoacacia* lectin (GenBank: AAA80183.1; 39.47%), and *Trifolium medium* lectin (GenBank: MCH79900.1; 35.44%) (Figure 5) [56,57].

Given the important differences in lectin specificity, it is very surprising that the sequences for CsL and Camptosemin are highly similar. Studies reported that even small differences at sequence level can affect the structure and organization of the carbohydrate recognition domain and will, therefore, also alter its specificity [10,58]. However, the current data do not allow a deep analysis since this requires the complete lectin sequence and three-dimensional structure of CsL.

3.4 Glycan array screening

Nowadays, several oligosaccharides, *N*-glycans, *O*-glycans, *N*-glycan core oligosaccharides, glycolipid type glycans, and other glycans are available following chemoenzymatic synthetic routes [59,60]. A collection of 144 synthetic glycan structures of mammalian and nonmammalian origin was printed in 7 identical subarrays onto NHS-activated slides. Purified CsL lectin was fluorescently labeled and applied to the array for glycan-binding analysis. The binding profile for CsL is compiled in Figure 7 and shows a preference for certain *N*-glycans with terminal GlcNAc, chitobiose (GL81), and chitotriose (GL96). The trisaccharide GlcNAcβ1-2(GlcNAcβ1-4Man) (GL46) was one of the stronger binders on the array, and *N*-glycans displaying this unsubstituted epitope were also

recognized very efficiently (GL47, GL52, GL53). Core fucosylation does not affect binding (GL71, GL78), while the presence of terminal galactose in triantennary *N*-glycans abolished binding (GL61, GL62). Terminal GlcNAc substituted *N*-glycans lacking this trisaccharide epitope, such as biantennary G0 (GL48) and triantennary GL50, were not recognized. Nevertheless, biantennary *N*-glycan structures with terminal GalNAc and Galactose residues (GL29, GL37, GL59, GL63) are bound much weaker than those oligosaccharides with a terminal GlcNAcβ1-2(GlcNAcβ1-4Man) motif.

These results are in line with the hemagglutination inhibition assays that also revealed the affinity of CsL for GlcNAc. Consequently, CsL exhibits a carbohydrate-binding specificity that is very different from other Diocleinae lectins, most of which present specificity for high-mannose *N*-glycans or lactose/GalNAc [44,61].

3.5 Toxicity of CsL against Artemia nauplii

The lethality test using Artemia is a simple preliminary assay to determine toxicity of CsL against cells [17,62, 63,64]. When toxicity of CsL for Artemia was examined using lectin concentrations ranging between 3.125 and 200 μg/mL, it was found that this lectin is innocuous to *Artemia* sp. within 24 and 48 hours. In contrast, Arruda et al. [62] showed that several lectins belonging to the Diocleinae subtribe are toxic to Artemia with lethal concentrations as low as 54.48, 110.51, 146.55 μg/mL for ConBr, ConGF, and ConM, respectively. Based on these results, CsL can be considered nontoxic to *Artemia* sp.

4. CONCLUSIONS

A new lectin was purified from the seeds of *Collaea speciosa*, the first lectin isolated from a plant of the genus *Collaea*. Since the dimeric lectin CsL exhibits an

interesting affinity for GlcNAc and can maintain its full activity even when heated until 90 °C for 1 hour, the properties of CsL are quite different from other lectins purified from plant species belonging to the Diocleinae subtribe. In addition, the partial sequence of CsL does not present sequence similarity with any known ConA-like lectin, but rather demonstrates a high degree of sequence identity with Camptosemin, the lectin from *Camptosema ellipticum*, with specificity towards lactose and derivatives. In contrast to other ConA-like lectins, CsL is not toxic to *Artemia* sp. Although the complete amino acid sequence of CsL is required for a definitive analysis, the current data suggest that CsL may be the first member of a third group of lectins inside the Diocleinae subtribe, next to the large group of mannose/glucose-specific ConA-like lectins and the small group of lactose-specific lectins.

ACKNOWLEDGEMENTS

This study was supported by grants from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP). B.S.C. and K.S.N. are senior investigators of CNPq. N.C.R. and S.S. acknowledge funding from the Ministry of Science and Education (MINECO), Grant No. CTQ2017-90039-R and RTC-2017-6126-1, and the Maria de Maeztu Units of Excellence Program from the Spanish State Research Agency, Grant No. MDM-2017-0720. David Martin helped with the English editing of the manuscript. VJSO is grateful to FWO-Vlaanderen for the postdoctoral fellowship (12T4622N).

REFERENCES

[1] W.J. Peumans, E.J.M. Van Damme, Lectins as Plant Defense Proteins, Plant Physiol. 109 (1995) 347–352. https://doi.org/10.1104/pp.109.2.347.

[2] E.J.M. Van Damme, W.J. Peumans, A. Barre, P. Rougé, Plant Lectins: A Composite of Several Distinct Families of Structurally and Evolutionary Related Proteins with Diverse Biological Roles, CRC Crit. Rev. Plant Sci. 17 (1998) 575–692. https://doi.org/10.1080/07352689891304276.

[3] M. Ambrosi, N.R. Cameron, B.G. Davis, Lectins: tools for the molecular understanding of the glycocode, Org. Biomol. Chem. 3 (2005) 1593–1608. https://doi.org/10.1039/b414350g.

[4] S.S. Komath, M. Kavitha, M.J. Swamy, Beyond carbohydrate binding: new directions in plant lectin research, Org. Biomol. Chem. 4 (2006) 973–988. https://doi.org/10.1039/b515446d.

[5] V. Bogoeva, L. Petrova, J. Bouckaert, A. Yordanova, I. Ivanov, R. Vanderesse, C. Frochot, Dual function of lectins — new perspectives in targeted photodynamic therapy, J. Porphyr. Phthalocyanines. 23 (2019) 1241–1250. https://doi.org/10.1142/S1088424619300209.

[6] S. Naithani, S.S. Komath, A. Nonomura, G. Govindjee, Plant lectins and their many roles: Carbohydrate-binding and beyond, J. Plant Physiol. 266 (2021) 153531. https://doi.org/10.1016/j.jplph.2021.153531.

[7] B.S. Cavada, V.J.S. Osterne, C.F. Lossio, V.R. Pinto-Junior, M.V. Oliveira, M.T.L. Silva, R.B. Leal, K.S. Nascimento, One century of ConA and 40 years of ConBr

research: A structural review, Int. J. Biol. Macromol. 134 (2019) 901–911. https://doi.org/10.1016/j.ijbiomac.2019.05.100.

[8] B.S. Cavada, V.R. Pinto-Junior, V.J.S. Osterne, K.S. Nascimento, ConA-Like Lectins: High Similarity Proteins as Models to Study Structure/Biological Activities Relationships, Int. J. Mol. Sci. 20 (2018). https://doi.org/10.3390/ijms20010030.

[9] L.M. Melgarejo, N. Vega, G. Pérez, Isolation and characterization of novel lectins from *Canavalia ensiformis* DC and *Dioclea grandiflora* Mart. ex Benth. seeds, Braz. J. Plant. Physiol. 17 (2005) 315–324. https://doi.org/10.1590/S1677-04202005000300006.

[10] B. Cavada, T. Barbosa, S. Arruda, T. Grangeiro, M. Barral-Netto, Revisiting proteus: Do Minor Changes in Lectin Structure Matter in Biological Activity? Lessons from and Potential Biotechnological Uses of the Diocleinae Subtribe Lectins, Current Protein & Peptide Science. 2 (2001) 123–135. https://doi.org/10.2174/1389203013381152.

[11] G.P. Lewis, B. Schrire, B. Mackinder, M. Lock, Legumes of the World, Royal Botanic Gardens, Kew, 2005.

[12] G.B. Ceolin, S.T.S. Miotto, O gênero *Collaea* DC. (Leguminosae, Papilionoideae) na Região Sul do Brasil, Acta Botanica Brasilica. 23 (2009) 991–998. https://doi.org/10.1590/s0102-33062009000400009.

[13] M. de L. Abruzzi Aragao de Oliveira, Collaea riparia Abruzzi de Oliveira, new species of *Leguminosae* (Faboideae: Phaseoleae-Diocleineae) from South of Brazil, Iheringia

Ser. Bot. 69 (2014) 323–328.

[14] R. Queiroz, V. my C. Profile, *Fabaceae - Collaea speciosa* (Loisel.) DC, (n.d.). http://rubens-plantasdobrasil.blogspot.com/2011/11/collaea-speciosa-loisel-dc.html (accessed February 1, 2020).

[15] Iucn, IUCN, *Collaea speciosa*: Botanic Gardens Conservation International (BGCI) & IUCN SSC Global Tree Specialist Group, IUCN Red List of Threatened Species. (2018). https://doi.org/10.2305/iucn.uk.2019-2.rlts.t144311564a149058934.en.

[16] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem. 72 (1976) 248–254. https://doi.org/10.1006/abio.1976.9999.

[17] B.S. Cavada, A.U. Bari, V.R. Pinto-Junior, C.F. Lossio, M.T.L. Silva, L.A.G. Souza, M.V. Oliveira, C.H.D. Souza-Filho, S.E.G. Correia, A.P.M.S. Vital, L.D. Lima, V.J.S. Osterne, K.S. Nascimento, Purification and partial characterization of a new lectin from *Parkia panurensis* Benth. ex H.C. Hopkins seeds (*Leguminosae* family; Mimosoideae subfamily) and evaluation of its biological effects, Int. J. Biol. Macromol. (2019). https://doi.org/10.1016/j.ijbiomac.2019.10.102.

[18] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature. 227 (1970) 680–685. https://www.ncbi.nlm.nih.gov/pubmed/5432063.

[19] R.M. Zacharius, T.E. Zell, J.H. Morrison, J.J. Woodlock, Glycoprotein staining following electrophoresis on acrylamide gels, Anal. Biochem. 30 (1969) 148–152.

https://doi.org/10.1016/0003-2697(69)90383-2.

[20] A. Shevchenko, H. Tomas, J. Havlis, J.V. Olsen, M. Mann, In-gel digestion for mass spectrometric characterization of proteins and proteomes, Nat. Protoc. 1 (2006) 2856–2860. https://doi.org/10.1038/nprot.2006.468.

[21] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

[22] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol. 7 (2011) 539. https://doi.org/10.1038/msb.2011.75.

[23] X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server, Nucleic Acids Res. 42 (2014) W320–4. https://doi.org/10.1093/nar/gku316.

[24] K. Brzezicka, B. Echeverria, S. Serna, A. van Diepen, C.H. Hokke, N.-C. Reichardt, Synthesis and microarray-assisted binding studies of core xylose and fucose containing *N*-glycans, ACS Chem. Biol. 10 (2015) 1290–1302. https://doi.org/10.1021/cb501023u.

[25] Resolução CONAMA Nº 357/2005 - “Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes, e dá outras providências.,” (n.d.).

http://www2.mma.gov.br/port/conama/legiabre.cfm?codlegi=459 (accessed January 3, 2020).

[26] J.B. Sumner, The globulins of the jack bean, *Canavalia ensiformis* preliminary paper, J. Biol. Chem. 37 (1919) 137–142. http://www.jbc.org/content/37/1/137.short.

[27] V.M. Ceccatto, B.S. Cavada, E.P. Nunes, N.A.P. Nogueira, M.B. Grangeiro, F.B.M.B. Moreno, E.H. Teixeira, A.H. Sampaio, M.A.O. Alves, M.V. Ramos, J.J. Calvete, T.B. Grangeiro, Purification and partial characterization of a lectin from *Canavalia grandiflora* benth. seeds, Protein Pept. Lett. 9 (2002) 67–73. https://doi.org/10.2174/0929866023409002.

[28] V.J.S. Osterne, M.Q. Santiago, V.R. Pinto-Junior, J.B. Cajazeiras, J.L.A. Correia, C.C.F. Leitão, R.F. Carneiro, F.N. Pereira-Junior, M.A. Vasconcelos, B.A.M. Rocha, A.M.S. Assreuy, P.H.S.F. Bringel, C.S. Nagano, K.S. Nascimento, B.S. Cavada, Purification, partial characterization, and CNBr-sepharose immobilization of a vasorelaxant glucose/mannose lectin from *Canavalia virosa* seeds, Appl. Biochem. Biotechnol. 172 (2014) 3342–3353. https://doi.org/10.1007/s12010-014-0751-3.

[29] K. Kojima, H. Ogawa, N. Seno, I. Matsumoto, Purification and characterization of Canavalia gladiata agglutinin, Carbohydr. Res. 213 (1991) 275–282. https://doi.org/10.1016/s0008-6215(00)90614-1.

[30] M.Q. Santiago, C.C.F. Leitão, F.N. Pereira Jr, V.R. Pinto Jr, V.J.S. Osterne, C.F. Lossio, J.B. Cajazeiras, H.C. Silva, F.V.S. Arruda, L.P. Pereira, A.M.S. Assreuy, K.S. Nascimento, C.S. Nagano, B.S. Cavada, Purification, characterization and partial sequence of a pro-inflammatory lectin from seeds of *Canavalia oxyphylla* Standl. & L.

O. Williams, J. Mol. Recognit. 27 (2014) 117–123. https://doi.org/10.1002/jmr.2340.

[31] C.A. de A. Gadelha, F.B.M.B. Moreno, T. Santi-Gadelha, J.B. Cajazeiras, B.A.M. da Rocha, A.M.S. Assreuy, M.R. Lima Mota, N.V. Pinto, A.V. Passos Meireles, J.C. Borges, B.T. Freitas, F. Canduri, E.P. Souza, P. Delatorre, D.N. Criddle, W.F. de Azevedo, B.S. Cavada, Native crystal structure of a nitric oxide-releasing lectin from the seeds of *Canavalia maritima*, J. Struct. Biol. 152 (2005) 185–194. https://doi.org/10.1016/j.jsb.2005.07.012.

[32] M. Silva, S. Silva, K. Nascimento, C. Nagano, B. Cavada, Physico-chemical characterization and partial sequence of a lectin from *Canavalia bonariensis* Lindl seeds, BMC Proc. 8 (2014) P227. https://doi.org/10.1186/1753-6561-8-S4-P227.

[33] T.R. Moura, G.A. Bezerra, M.J.B. Bezerra, C.S. Teixera, E.H.S. Bezerra, R.G. Benevides, B.A.M. da Rocha, L.A.G. de Souza, P. Delatorre, C.S. Nagano, B.S. Cavada, Crystallization and preliminary X-ray diffraction analysis of the lectin from *Canavalia boliviana* Piper seeds, Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 65 (2009) 213–215. https://doi.org/10.1107/S1744309109000797.

[34] C.F. Lossio, C.G. Moreira, R.M.F. Amorim, C.S. Nobre, M.T.L. Silva, C.C. Neto, V.R. Pinto-Junior, I.B. Silva, J. Campos, A.M.S. Assreuy, B.S. Cavada, K.S. Nascimento, Lectin from *Canavalia villosa* seeds: A glucose/mannose-specific protein and a new tool for inflammation studies, Int. J. Biol. Macromol. 105 (2017) 272–280. https://doi.org/10.1016/j.ijbiomac.2017.07.038.

[35] M.T. Correia, L.C. Coelho, Purification of a glucose/mannose specific lectin, isoform 1, from seeds of *Cratylia mollis* Mart. (Camaratu bean), Appl. Biochem. Biotechnol.

55 (1995) 261–273. https://doi.org/10.1007/bf02786865.

[36] P.M.G. Paiva, L.C.B.B. Coelho, Purification and partial characterization of two lectin isoforms from *Cratylia mollis* mart. (camaratu bean), Appl. Biochem. Biotechnol. 36 (1992) 113–118. https://doi.org/10.1007/BF02929691.

[37] B.S. Cavada, E.S. Marinho, E.P. Souza, R.G. Benevides, P. Delatorre, L.A.G. Souza, K.S. Nascimento, A.H. Sampaio, F.B.M.B. Moreno, J.K.R. Rustiguel, F. Canduri, W.F. de Azevedo Jr, H. Debray, Purification, partial characterization and preliminary X-ray diffraction analysis of a mannose-specific lectin from *Cymbosema roseum* seeds, Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 62 (2006) 235–237. https://doi.org/10.1107/S174430910600371X.

[38] B.A.M. Rocha, F.B.M.B. Moreno, P. Delatorre, E.P. Souza, E.S. Marinho, R.G. Benevides, J.K.R. Rustiguel, L.A.G. Souza, C.S. Nagano, H. Debray, A.H. Sampaio, W.F. de Azevedo Jr, B.S. Cavada, Purification, characterization, and preliminary X-ray diffraction analysis of a lactose-specific lectin from *Cymbosema roseum* seeds, Appl. Biochem. Biotechnol. 152 (2009) 383–393. https://doi.org/10.1007/s12010-008-8334-9.

[39] C. Souza Teixeira, H.C. da Silva, T.R. de Moura, F.N. Pereira-Júnior, K.S. do Nascimento, C.S. Nagano, A.H. Sampaio, P. Delatorre, B.A.M. Rocha, B.S. Cavada, Crystal structure of the lectin of *Camptosema pedicellatum*: implications of a conservative substitution at the hydrophobic subsite, J. Biochem. 152 (2012) 87–98. https://doi.org/10.1093/jb/mvs047.

[40] F.A.H. Batista, L.S. Goto, W. Garcia, D.I. de Moraes, M. de Oliveira Neto, I.

Polikarpov, M.R. Cominetti, H.S. Selistre-de-Araújo, L.M. Beltramini, A.P.U. Araújo, Camptosemin, a tetrameric lectin of *Camptosema ellipticum*: structural and functional analysis, Eur. Biophys. J. 39 (2010) 1193–1205. https://doi.org/10.1007/s00249-009-0571-5.

[41] M. Almanza, N. Vega, G. Pérez, Isolating and characterising a lectin from *Galactia lindenii* seeds that recognises blood group H determinants, Arch. Biochem. Biophys. 429 (2004) 180–190. https://doi.org/10.1016/j.abb.2004.06.010.

[42] J. Le Pendu, G. Gérard, F. Lambert, R. Mollicone, A new anti-H lectin from the seeds of Galactia tenuiflora, Glycoconj. J. (1986).

[43] N. Fukuda, A. Yoshimaru, T. Hidaka, H. Ohta, K. Yamamoto, H. Yomo, Isolation and Characterization of N-Acetylgalactosamine-specific Lectin from *Galactia tashiroi* Seeds, Biosci. Biotechnol. Biochem. 58 (1994) 423–424. https://doi.org/10.1271/bbb.58.423.

[44] K.A. Maupin, D. Liden, B.B. Haab, The fine specificity of mannose-binding and galactose-binding lectins revealed using outlier motif analysis of glycan array data, Glycobiology. 22 (2012) 160–169. https://doi.org/10.1093/glycob/cwr128.

[45] T.K. Dam, B.S. Cavada, T.B. Grangeiro, C.F. Santos, F.A.M. de Sousa, S. Oscarson, C. Fred Brewer, 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, Journal of Biological Chemistry. 273 (1998) 12082–12088. https://doi.org/10.1074/jbc.273.20.12082.

[46] M.V. Ramos, B.S. Cavada, A.-M. Mazard, P. Rougé, Interaction of Diocleinae lectins with glycoproteins based in surface plasmon resonance, Mem. Inst. Oswaldo Cruz. 97 (2002) 275–279. https://doi.org/10.1590/s0074-02762002000200025.

[47] C. Achilli, A. Ciana, G. Minetti, Oxidation of cysteine-rich proteins during gel electrophoresis, J Biol Methods. 5 (2018) e104. https://doi.org/10.14440/jbm.2018.275.

[48] J.J. Calvete, C.F. Santos, T.B. Grangeiro, M. Nimtz, B.S. Cavada, Primary structure and posttranslation processing of *Vatairea macrocarpa* lectin, J. Protein Chem. 17 (1998) 545-547. 10.1007/BF00392207.

[49] H.C. Silva, C.S. Nagano, L.A.G. Souza, K.S. Nascimento, R. Isidro, P. Delatorre, B.A.M. Rocha, A.H. Sampaio, A.M.S. Assreuy, A.F. Pires, L.E.A. Damasceno, G.F.O. Marques-Domingos, B.S. Cavada, Purification and primary structure determination of a galactose-specific lectin from *Vatairea guianensis* Aublet seeds that exhibits vasorelaxant effect, Proc. Biochem. 47 (2012) 2347-2355. https://doi.org/10.1016/j.procbio.2012.09.014.10.1083/jcb.102.4.1284.

[50] K.S. Nascimento, M.T.L. Silva, M.V. Oliveira, C.F. Lossio, V.R. Pinto-Junior, V.J.S. Osterne, B.S. Cavada, Dalbergieae lectins: A review of lectins from species of a primitive Papilionoideae (leguminous) tribe, Int. J. Biol. Macromol. 144 (2020) 509–526. https://doi.org/10.1016/j.ijbiomac.2019.12.117.

[51] B.S. Cavada, V.J.S. Osterne, V.R. Pinto-Junior, K.S. Nascimento, ConBr, the Lectin from Canavalia brasiliensis Mart. Seeds: Forty Years of Research, Curr. Protein Pept. Sci. 20 (2019) 600–613. https://doi.org/10.2174/1389203720666190104123210.

[52] B.S. Cavada, C.F. Santos, T.B. Grangeiro, E.P. Nunes, P.V. Sales, R.L. Ramos, F.A. De Sousa, C.V. Crisostomo, J.J. Calvete, Purification and characterization of a lectin from seeds of *Vatairea macrocarpa* Duke, Phytochemistry. 49 (1998) 675–680. https://www.ncbi.nlm.nih.gov/pubmed/9779593.

[53] B.S. Cavada, Lectinas de *Canavalia brasiliensis* Mart. Isolamento, caracterização parcial e comportamento durante a germinação, [Dissertação]. Fortaleza: Universidade Federal do Ceará, 1980.

[54] M. Kaiser, R. Huber, M. Ehrmann, Proteolysis, in: S. Maloy, K. Hughes (Eds.), Brenner's Encyclopedia of Genetics (Second Edition), Academic Press, San Diego, 2013: pp. 501–503. https://doi.org/10.1016/B978-0-12-374984-0.01229-8.

[55] C. López-Otín, J.S. Bond, Proteases: multifunctional enzymes in life and disease, J. Biol. Chem. 283 (2008) 30433–30437. https://doi.org/10.1074/jbc.R800035200.

[56] S. Bliven, A. Prlić, Circular permutation in proteins, PLoS Comput. Biol. 8 (2012) e1002445. https://doi.org/10.1371/journal.pcbi.1002445.

[57] D.M. Carrington, A. Auffret, D.E. Hanke, Polypeptide ligation occurs during post-translational modification of concanavalin A, Nature. 313 (1985) 64–67. https://doi.org/10.1038/313064a0.

[58] M.V. Ramos, T.B. Grangeiro, B.S. Cavada, I. Shepherd, R.O. de M. Lopes, A.H. Sampaio, Carbohydrate/glycan-binding specificity of legume lectins in respect to their proposed biological functions, Braz. Arch. Biol. Technol. 43 (2000) 0–0. https://doi.org/10.1590/S1516-89132000000400001.

[59] S. Serna, J. Etxebarria, N. Ruiz, M. Martin-Lomas, N.-C. Reichardt, Construction of N-glycan microarrays by using modular synthesis and on-chip nanoscale enzymatic glycosylation, Chemistry. 16 (2010) 13163–13175. https://doi.org/10.1002/chem.201001295.

[60] B. Echeverria, S. Serna, S. Achilli, C. Vivès, J. Pham, M. Thépaut, C.H. Hokke, F. Fieschi, N.-C. Reichardt, Chemoenzymatic Synthesis of N-glycan Positional Isomers and Evidence for Branch Selective Binding by Monoclonal Antibodies and Human C-type Lectin Receptors, ACS Chem. Biol. 13 (2018) 2269–2279. https://doi.org/10.1021/acschembio.8b00431.

[61] B.S. Cavada, M.T.L. Silva, V.J.S. Osterne, V.R. Pinto-Junior, C.F. Lossio, J.C. Madeira, M.G. Pereira, R.B. Leal, W.P. Ferreira, K.S. Nascimento, A.M.S. Assreuy, Exploring the carbohydrate-binding ability of *Canavalia bonariensis* lectin in inflammation models, J. Mol. Recognit. (2020) e2870. https://doi.org/10.1002/jmr.2870.

[62] F.V.S. Arruda, A.A. Melo, M.A. Vasconcelos, R.F. Carneiro, I.L. Barroso-Neto, S.R. Silva, F.N. Pereira-Junior, C.S. Nagano, K.S. Nascimento, E.H. Teixeira, S. Saker-Sampaio, B. Sousa Cavada, A.H. Sampaio, Toxicity and binding profile of lectins from the Genus *Canavalia* on brine shrimp, Biomed Res. Int. 2013 (2013) 154542. https://doi.org/10.1155/2013/154542.

[63] I.M. Vasconcelos, B.S. Cavada, R.D.E.A. Moreira, J.T.A.D.E. Oliveira, Purification and partial characterization of a lectin from the seeds of *Dioclea guianensis*, J. Food Biochem. 15 (1991) 137–154. https://doi.org/10.1111/j.1745-4514.1991.tb00150.x.

[64] V.R. Pinto-Junior, J.L.A. Correia, R.I. Pereira, F.N. Pereira-Junior, M.Q. Santiago, V.J.S. Osterne, J.C. Madeira, J.B. Cajazeiras, C.S. Nagano, P. Delatorre, A.M.S. Assreuy, K.S. Nascimento, B.S. Cavada, Purification and molecular characterization of a novel mannose-specific lectin from *Dioclea reflexa* hook seeds with inflammatory activity, J. Mol. Recognit. 29 (2016) 134–141. https://doi.org/10.1002/jmr.2512.

FIGURES LEGENDS

Figure 1. Purification of *Collaea speciosa* lectin. A) Affinity chromatography on Sephadex G-50 (2 cm x 9.6 cm). Fractions (non-retained and retained proteins) were collected and were detected spectrophotometrically at 280 nm. B) SDS-PAGE under denaturing conditions. Lanes 1 and 6 show molecular markers TrueColor High Range S2600. 2. Crude extract (4

mg/mL); 3. Non-retained protein fraction (4 mg/mL); 4. CsL (4 mg/mL) under nonreducing conditions; 5. CsL (4 mg/mL) after treatment with β-mercaptoethanol. C) Standard curve of proteins present in the molecular weight marker proteins.

Figure 2. High-performance size-exclusion liquid chromatography (UPLC). A) 7.8 × 300 mm column, ACQUITY, UPLC Waters. CsL (4 mg/mL) was applied. The proteins were analyzed in 50 mM sodium phosphate dibasic, pH 7.2, buffer containing 150 mM NaCl. Fractions of approximately 1 mL were collected at a flow rate of 0.5 mL/min and were detected spectrophotometrically at 280 nm. B) Calibration curve: ribonuclease A (13,700 Da), carbonic anhydrase (29,000 Da), ovalbumin (44,000 Da) and conalbumin (75,000 Da).

Figure 3. Percentage of hemagglutinating activity (H.A.) remaining after incubation of CsL at different pH conditions (A) and temperatures (B). Panels C and D show the influence of EDTA and divalent cations on hemagglutinating activity. Control group (only saline with erythrocytes without lectin). Analyses were done using untreated rabbit red blood cells.

Figure 4. Periodic acid Schiff (PAS) staining to detect glycoproteins. Lanes 1 and 7: molecular marker TrueColor High Range S2600. 2 and 8: crude extract. 3 and 9: non-retained fraction on Sephadex G-50. 4 and 10: CsL lectin. 5 and 11: ConBr lectin (positive control). 6 and 12: VML lectin (negative control). Lanes 1 to 6 were stained with Coomassie Brilliant Blue G250. Lanes 7 to 12 were stained using the PAS method.

Figure 5. Multiple alignments of the partial sequence of CsL with complete lectin sequences from *Camptosema ellipticum* (Camptosemin), *Trifolium medium*, *Medicago truncatula*, *Robinia pseudoacacia*, and *Arachis hypogaea* (PNA).

Figure 6. Partial sequence alignment of CsL and Camptosemin showing a high degree of similarity between both sequences.

Figure 7. Glycan array screening of CsL. Each histogram represents the mean RFU (relative fluorescence unit) values for four spots with SD (standard deviation) of the mean. Only select glycan structures are represented. All carbohydrate printed structures and complete array results are shown in Figures S2 and S3.

Figure 1

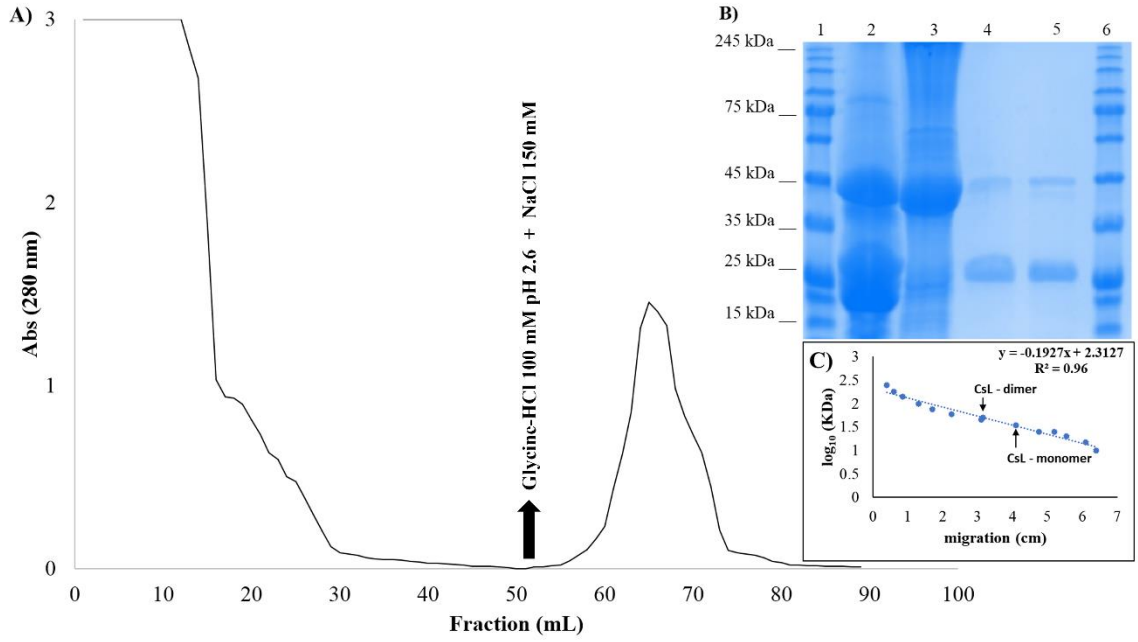

A)
Abs (280 nm)
Fraction (mL)
Glycine-HCl 100 mM pH 2.6 + NaCl 150 mM
B)
245 kDa
75 kDa
45 kDa
35 kDa
25 kDa
15 kDa
C)
y = -0.1927x + 2.3127
R² = 0.96
CsL - dimer
CsL - monomer
log10 (KDa)
migration (cm)

Figure 2

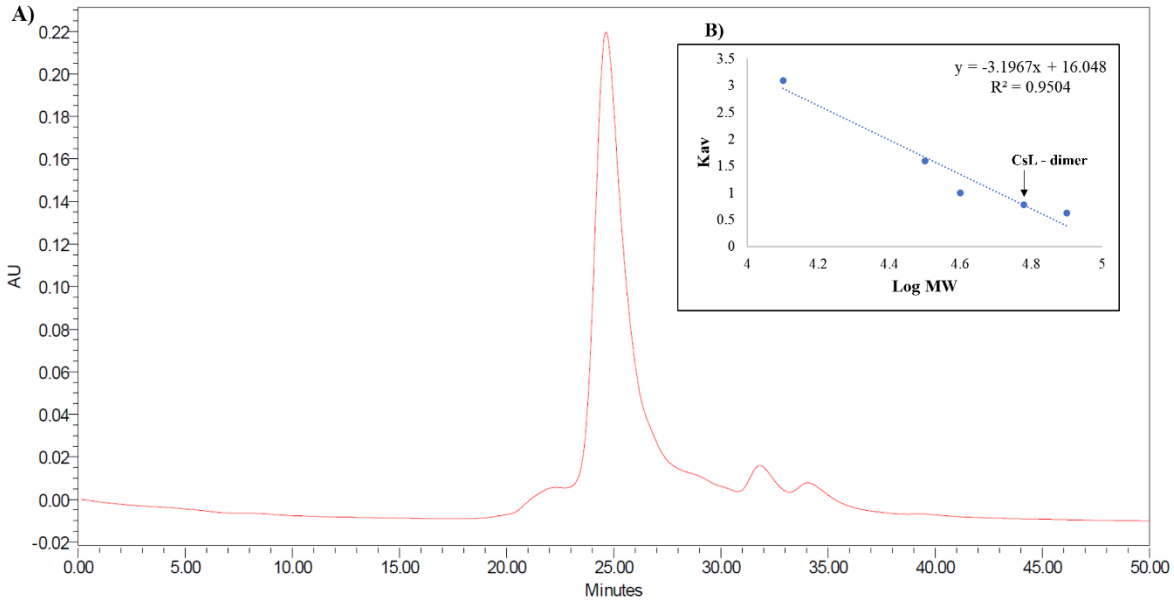

A)
AU
0.22
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
0.00
5.00
10.00
15.00
20.00
25.00
30.00
35.00
40.00
45.00
50.00
Minutes
B)
y = -3.1967x + 16.048
R² = 0.9504
Kav
3.5
3
2.5
2
1.5
1
0.5
0
4
4.2
4.4
4.6
4.8
5
Log MW
CsL - dimer

Figure 3

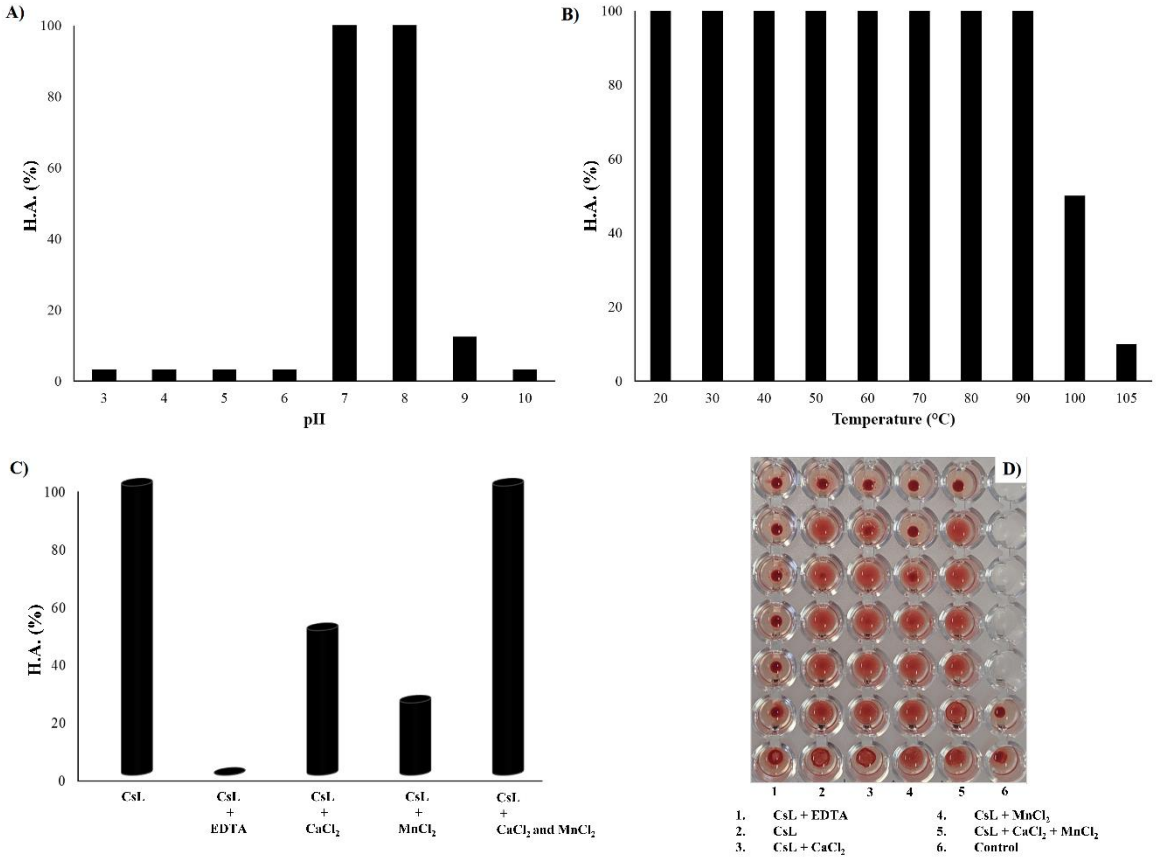

A)
H.A. (%)
pII
B)
H.A. (%)
Temperature (°C)
C)
H.A. (%)
CsL
CsL + EDTA
CsL + CaCl2
CsL + MnCl2
CsL + CaCl2 and MnCl2
D)
1. CsL + EDTA
2. CsL
3. CsL + CaCl2
4. CsL + MnCl2
5. CsL + CaCl2 + MnCl2
6. Control

Figure 4

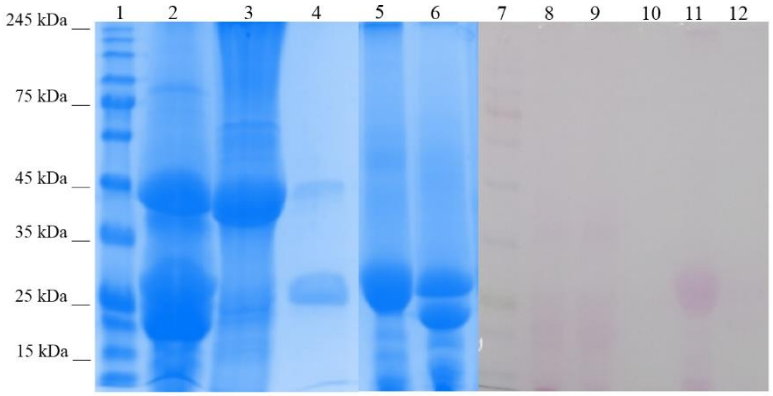


Figure 5

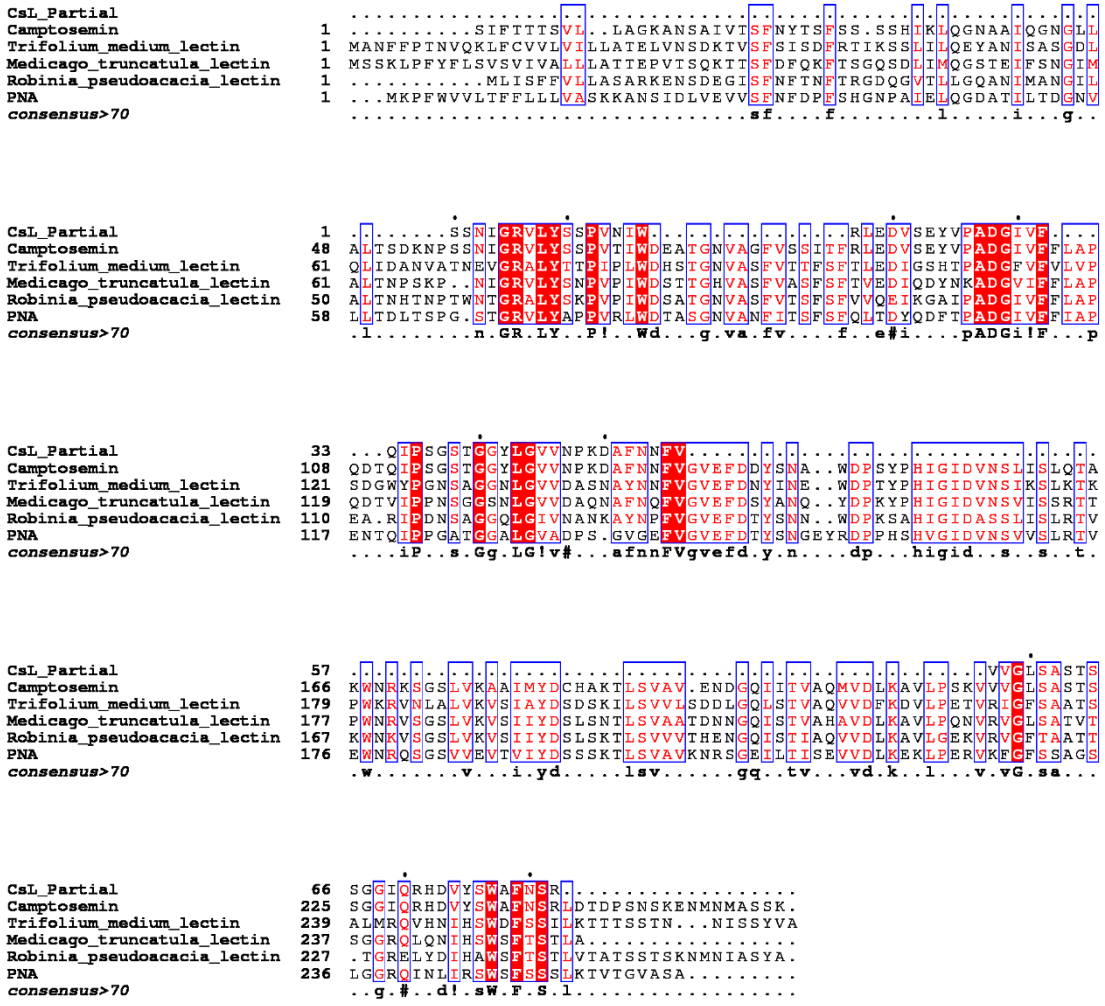

CsL_Partial
Camptosemin
Trifolium_medium_lectin
Medicago_truncatula_lectin
Robinia_pseudoacacia_lectin
PNA
consensus>70

Figure 6

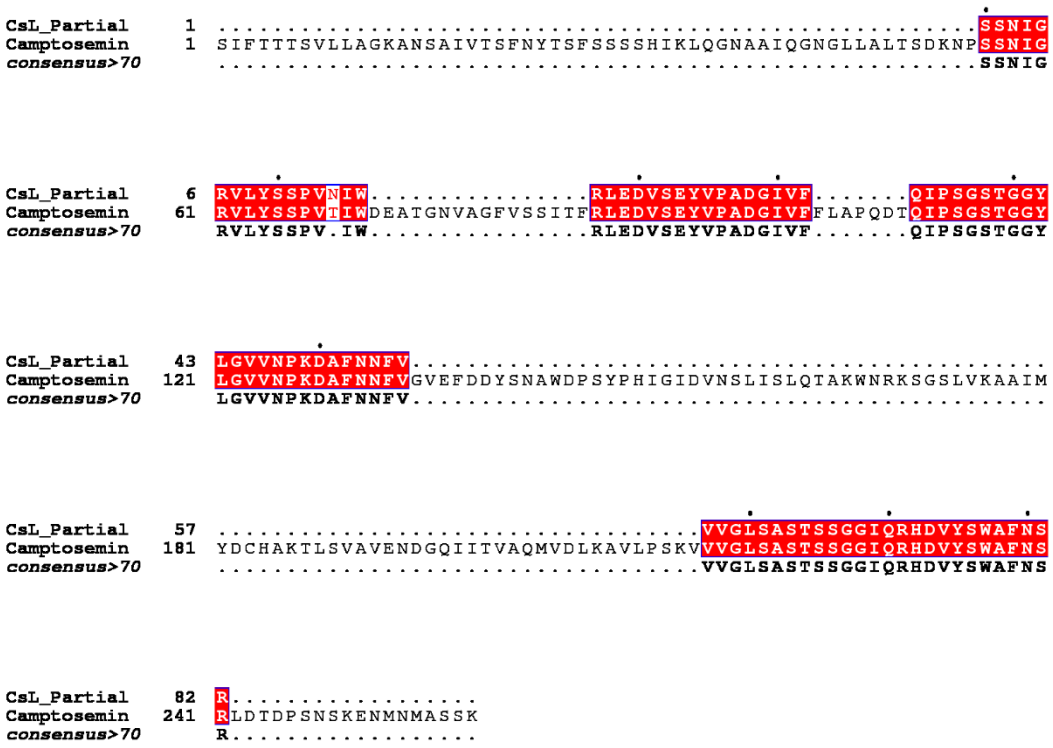
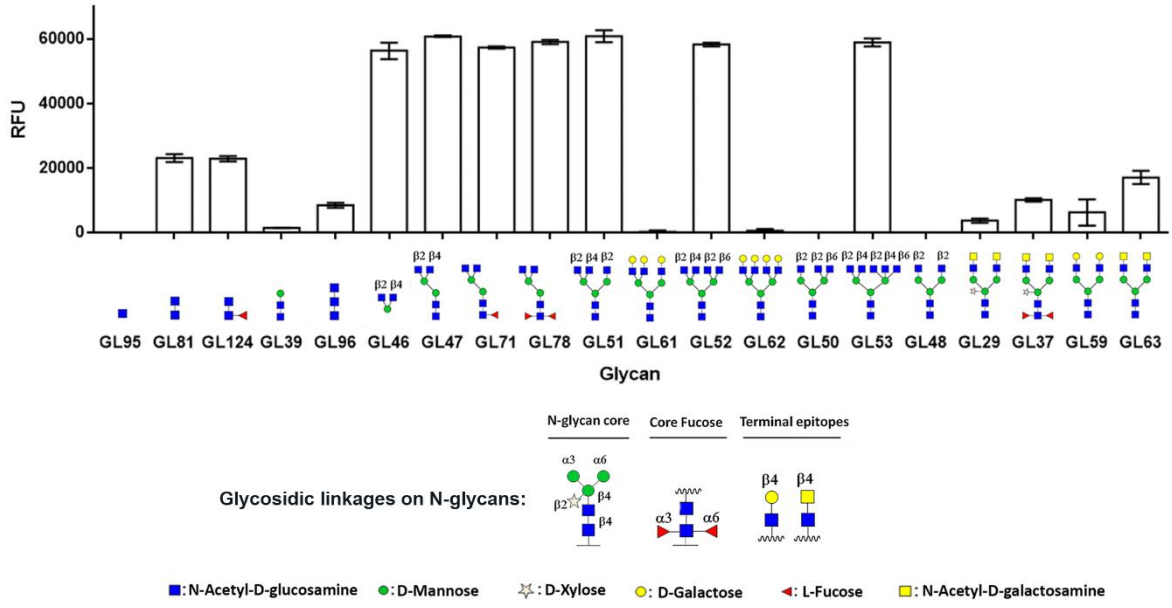

CsL_Partial 1 ..... SSNIG
Camptosemin 1 SIFTTTSVLLAGKANSAIVTSFNYTSFSSSSHIKLQGNAAIQGNGLLALTSDKNP SSNIG
consensus>70 ..... SSNIG
CsL_Partial 6 RVLYSSPVNIW ..... RLEDVSEYVPADGIVF ..... QIPSGSTGGY
Camptosemin 61 RVLYSSPVTIW DEATGNVAGFVSSITF RLEDVSEYVPADGIVF FLAPQDT QIPSGSTGGY
consensus>70 RVLYSSPV.IW ..... RLEDVSEYVPADGIVF ..... QIPSGSTGGY
CsL_Partial 43 LGVVNPKDAFNNFV .....
Camptosemin 121 LGVVNPKDAFNNFV GVEFDDYSNAWDPSYPHIGIDVNSLISLQTAKWNRKSGSLVKAAIM
consensus>70 LGVVNPKDAFNNFV .....
CsL_Partial 57 ..... VVGLSASTSSGGIQRHDVYSWAFNS
Camptosemin 181 YDCHAKTLSVAVENDGQIITVAQMVDLKAVLPSKV VVGLSASTSSGGIQRHDVYSWAFNS
consensus>70 ..... VVGLSASTSSGGIQRHDVYSWAFNS
CsL_Partial 82 R .....
Camptosemin 241 RLDTDPSNSKENMNMASSK
consensus>70 R .....

Figure 7


RFU
60000
40000
20000
0
GL95 GL81 GL124 GL39 GL96 GL46 GL47 GL71 GL78 GL51 GL61 GL52 GL62 GL50 GL53 GL48 GL29 GL37 GL59 GL63
Glycan
N-glycan core
Core Fucose
Terminal epitopes
Glycosidic linkages on N-glycans:
α3 α6 β2 β4 β4
α3 α6
β4 β4
: N-Acetyl-D-glucosamine
: D-Mannose
: D-Xylose
: D-Galactose
: L-Fucose
: N-Acetyl-D-galactosamine

A)

B)

Figure S1. Figures of plants and seeds of *Collaea speciosa*.

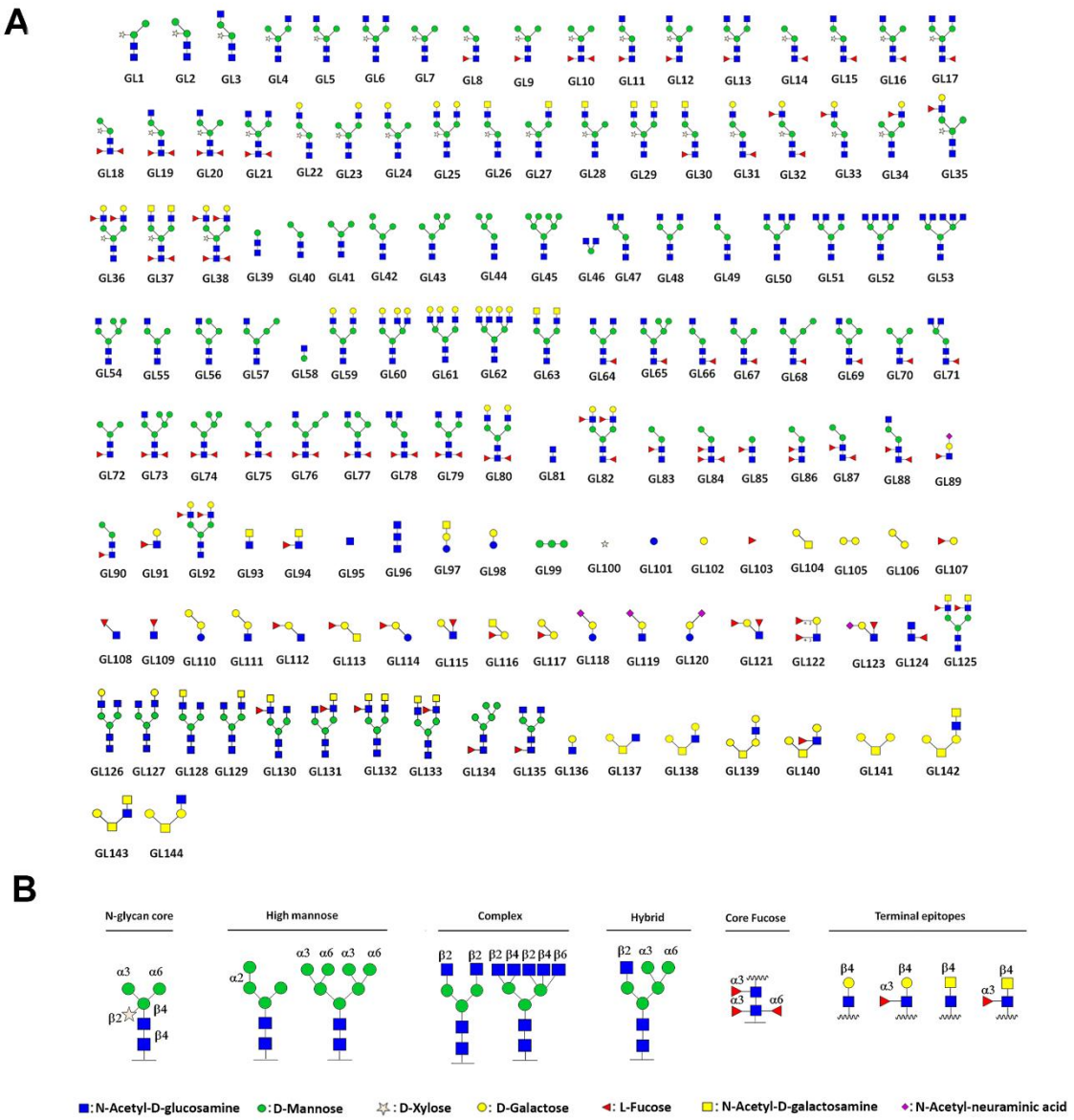


Figure S2. Glycan microarray consisting of 144 different glycan structures; A) Glycan structures included on the microarrays, B) Nature of the glycosidic linkages of the N-glycan structures on the microarrays.

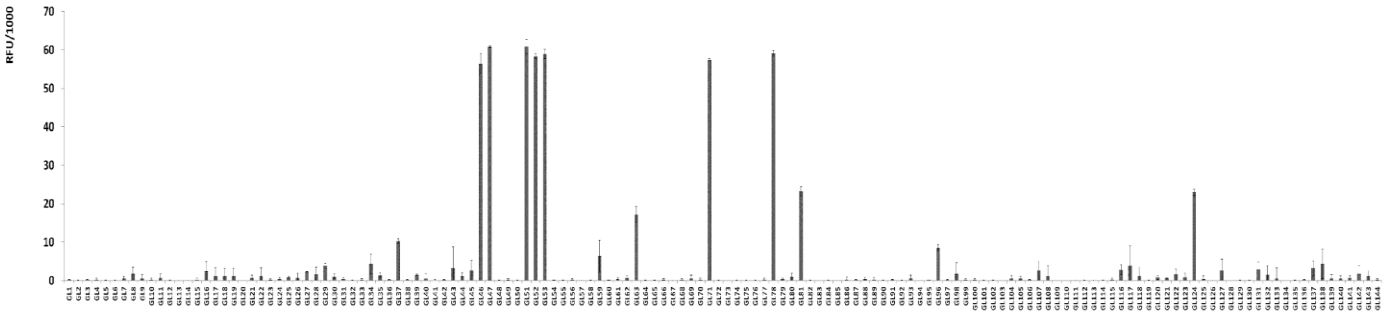


Figure S3. Glycan microarray screening of CSL. Each histogram represents the mean RFU (relative fluorescence unit) values for four spots with the SD (standard deviation) of the mean.