Galactose-binding lectin from the seeds of champedak (Artocarpus integer): sequences of its subunits and interactions with human serum O-glycosylated glycoproteins

Onn Haji Hashim, University of Malaya
Galactose-binding lectin from the seeds of champedak (*Artocarpus integer*): sequences of its subunits and interactions with human serum O-glycosylated glycoproteins

Mariati Abdul Rahman, Saiful Anuar Karsani, Iekhsan Othman, Puteri Shafinaz Abdul Rahman, and Onn Haji Hashim*

Department of Biochemistry, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia

Received 24 June 2002

Abstract

Our group has previously reported the isolation, partial characterisation, and application of a Galβ1–3GalNAc- and IgA1-reactive lectin from the seeds of champedak (*Artocarpus integer*). In the present study, we have subjected the purified lectin to reverse-phase high performance liquid chromatography and sequenced its subunits. Determination of the N-terminal sequence of the first 47 residues of the large subunit demonstrated at least 95% homology to the N-terminal sequence of the α chains of a few other galactose-binding *Artocarpus* lectins. The two smaller subunits of the lectin, each comprised of 21 amino acid residues, demonstrated minor sequence variability. Their sequences were generally comparable to the β chains of the other galactose-binding *Artocarpus* lectins. When used to probe human serum glycopeptides that were separated by two-dimensional gel electrophoresis, the lectin demonstrated strong apparent interactions with glycopeptides of IgA1, hemopexin, α2-HS glycoprotein, α1-antichymotrypsin, and a few unknown glycoproteins. Immobilisation of the lectin to Sepharose generated an affinity column that may be used to isolate the O-glycosylated serum glycoproteins. © 2002 Elsevier Science (USA). All rights reserved.

Keywords: Lectin; *Artocarpus integer*; Sequence; Serum glycoproteins; O-Glycan

*Artocarpus integer*, commonly known as champedak is found in abundance throughout Malaysia. Its fruit is generally similar to jackfruit (*Artocarpus heterophyllus integrifolia*), but smaller in size, with less pith, more seeds, and softer pulp. The seeds of *A. integer* have been reported to have a galactose-binding lectin, which is similar to the jackfruit lectin, jacalin [1]. However, the structures and interactions of the champedak galactose-binding (CGB) lectins that were isolated from different cultivars were more consistent and uniform when compared to jacalins of distinctive origins [2,3]. The champedak seed lectin apparently recognised the T antigenic determinant Galβ1–3GalNAc in IgA1 and C1 inhibitor molecules [1,3]. Binding of the lectin to C1 inhibitor caused a spontaneous activation of the classical pathway of complement [4].

Because of its reactivity to the O-glycans at the hinge region of IgA1, the champedak seed lectin has been applied to study the carbohydrate structures of serum IgA1 from IgA nephropathy patients. The patients’ IgA1 demonstrated lower affinity towards the champedak seed lectin and its α-heavy chains were relatively more cationic as compared to that of normal controls [5]. The binding abnormality as well as the differential charge distribution of the IgA1 α-heavy chains was abrogated when the experiments were performed using IgA1-containing sera that were treated with neuraminidase, indicating involvement of sialic acid residues [6].

In the present study, we have subjected purified champedak seed lectin to reverse-phase high performance liquid chromatography (HPLC) and sequenced its separated subunits. The interactions of the lectin with denatured and intact human serum glycoproteins were investigated.
Materials and methods

**Extraction and purification of champedak galactose-binding lectin.** The extraction of the CGB lectin was performed as previously described [1,3]. The galactose-binding lectin was purified using an immobilised galactose-Sepharose 4B-affinity chromatography in PBS containing 0.15M NaCl, pH 7.2. Elution of the lectin was performed using 0.8 M of α-galactose in PBS. The effluent was dialyzed extensively against PBS and stored in aliquots at −20 °C. Purity and specificity of the lectin were confirmed using previously described biochemical and immunochemical analyses [1,3].

**Reverse-phase high performance liquid chromatography.** CGB lectin was solubilised in 6 M guanidine chloride and incubated at 60 °C for 2 h. The sample was subjected to pyridylethylation. Trifluoroacetic acid (TFA) was added to a final concentration of 25% (v/v) and the treated sample was applied into a reverse-phase HPLC column (C18; Genesis; 25 cm) and eluted with a gradient of acetonitrile in the presence of 0.1% TFA. For untreated champedak lectin, sample was separated using the same column and elution procedure but without prior guanidine chloride and pyridylethylation treatments.

**N-terminal sequencing of the subunits of champedak galactose-binding lectin.** Peak fractions were N-terminally sequenced by Edman degradation using an automated PE Biosystems Procise 491 protein sequencer. The resulting N-terminal sequences were subjected to a BLAST search program [7] to determine sequence homology with other known proteins in the database.

**Horseradish peroxidase conjugation.** Conjugation of the CGB lectin to horseradish peroxidase (HRP) was performed by incubation of the purified lectin with periodate-activated HRP in sodium carbonate buffer pH 9.5 for 2 h at room temperature. The solution was stirred for another 2 h at 4 °C upon addition of sodium borohydrate to reduce the remaining free enzymes. The conjugated lectin was kept at 4 °C in the presence of 60% (v/v) glycerol.

**Two-dimensional gel electrophoresis.** Two-dimensional gel electrophoresis (2DE) was performed using the Multiphor II Electrophoresis System based on the recommended protocols by Amersham Pharmacia Biotech (Uppsala, Sweden). Samples were diluted in buffer containing 8 M urea, 0.5% (v/v) Triton X-100, 2% (v/v) IPG buffer 4–7, and 12 Mm diithiothreitol (DTT). Pre-cast immobilised dry strips, pH 4–7 (Amersham Pharmacia Biotech, Uppsala, Sweden) were rehydrated with 8 M urea, 0.5% (v/v) IPG buffer 4–7, 0.5% (v/v) Triton X-100, and 12 Mm DTT. Isoelectric focusing was performed in three steps by varying the voltage as recommended by the manufacturer. For the second dimension, the strips were initially equilibrated in 50 mM Tris-HCl, pH 8.8 buffer, containing 6 M urea, 2% SDS (w/v), 30% glycerol (v/v), and 90 mg DTT, for 10 min. A second equilibration was performed for another 10 min in a similar solution but containing 0.45 g of iodoacetamide instead of DTT. The strips were placed on an 8–18% SDS–PAGE gradient gel. Electrophoresis was performed at a constant current of 40 mA per gel.

**Western blotting.** The gel was transferred onto a nitrocellulose membrane (0.45 μM) using a NovaBlot Kit of the Multiphor II Electrophoresis System (Amersham Pharmacia, Uppsala, Sweden). Blotting was performed at 0.8 mA/cm² for 2 h. The membrane was incubated in Tris-buffered saline (TBS, pH 7.5) in the presence of 0.1% (v/v) Tween 20, and 3% (w/v) gelatin. It was washed three times with TBS with 0.1% Tween 20 (pH 7.5) and subsequently incubated with HRP-conjugated champedak lectin in the same buffer overnight at 4 °C. The membrane was developed using 3, 3’-diaminobenzidine (Bio-Rad, Hercules, USA). Reaction was stopped by washing the membrane with distilled water.

**Glycoproteins and enzyme treatment.** Hemin-reactive serum glycoproteins were isolated by using immobilised-hemin-affinity column (Sigma Chemical, St. Louis, USA), performed as described by Tsutsui and Mueller [8]. Purified human colostral IgA1, α-antichymotrypsin and α-HS glycoprotein were purchased from Sigma Chemical, St. Louis, USA.

Removal of the O-glycans of IgA1 and serum glycoproteins was performed by sequential treatment with neuraminidase (Vibrio cholerae; Sigma Chemical, St. Louis, USA) and O-glycosidase (Diplococcus pneumoniae; Roche Diagnostic GmbH, Mannheim, Germany) according to the supplier’s instructions and previously described procedures [9].

**Polyacrylamide gel electrophoresis and protein determination.** Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) was performed using 8–18% gradient or 12.5% homogenous polyacrylamide gels according to the discontinuous Tris–glycine system described by Laemmli [10] under reducing condition. SDS–PAGE gels were developed by silver staining according to the method of Heukeshoven and Dernick [11]. Protein concentration was determined according to the method of Lowry as modified by Peterson [12] using bovine albumin as standard.

**Immobilised lectin-affinity chromatography.** The CGB lectin was conjugated to CNBr-activated Sepharose 4B according to the manufacturer’s instructions (Sigma Chemical, St. Louis, USA). Chromatography was performed by application of 0.8 ml of normal human serum into a 2 ml immobilised lectin column at room temperature. Unbound fractions were washed with PBS until the absorbance readings of the fractions reached baseline. Fractions of 1 ml were collected. The bound fractions were eluted with 0.1 M melibiose in PBS, dialysed, and subjected to 2DE.

Results

**Lectin subunits and their sequences.**

Champedak seed galactose-binding lectin was purified using immobilised galactose–Sepharose-affinity column chromatography as previously described [1,3]. The bound lectin was eluted with 0.8 M galactose solution. When the purified champedak seed lectin was subjected to reverse-phase HPLC, it was separated into two major (α and β) and one minor (β1) peaks, both with and without prior guanidinium chloride treatment (Figs. 1A and B). Whilst elutions of the α and β2 peaks were obtained at 100% acetonitrile, the β1 peak was eluted at the range of 93–97% acetonitrile. The β1 and β2 subunits apparently shared close amino acid sequence homology, comprising of 21 amino acid residues with only four residue differences (Table 1). The sequence of the α subunit was entirely different from the β1 and β2 subunits. Determination of the N-terminal sequence of its 47 amino acid residues demonstrated close sequence homology with the α chains of other galactose-binding Artocarpus lectins (Table 2).

**Interaction of the champedak lectin with reduced human serum glycoproteins.**

Separation of normal human serum proteins by 2DE and detection by silver staining generated a protein expression map consisting of hundreds of spots (Fig. 2A). Only selective few sets of these spots were detected when the 2DE-separated proteins were transferred onto a nitrocellulose membrane and exposed to the HRP-conju-
gated Galβ1–3GalNac-reactive champedak seed lectin (Fig. 2B). However, interaction with the enzyme-conjugated lectin was not observed when the experiment was repeated using human serum that was treated with neuraminidase and O-glycosidase. Using the SWISS-2DPAGE standard human plasma protein map as a reference [13] subsequently identified some of the lectin-detected spots. The data clearly indicate that the CGB lectin strongly interacted with IgA1 heavy chain, hemopexin, α1-antichymotrypsin, α2-HS glycoprotein, and a few unidentified protein spots.

To confirm the identities of the lectin-detected known glycopeptides, their purified samples were obtained and subjected to one-dimensional SDS–PAGE and Western blotting. Since hemopexin was not available commercially, hemin-reactive serum glycoproteins were isolated and used instead. Exposure of the nitrocellulose membrane with HRP-conjugated CGB lectin solution confirms the interactions of the lectin with the α heavy chain of IgA1, α2-HS glycoprotein, α1-antichymotrypsin, and the hemin-reactive serum glycoproteins (Figs. 3A and B).

Interactions of lectin with intact human serum glycoproteins

The interactions of CGB lectin with native intact human serum glycoproteins were subsequently analysed by subjecting normal human serum samples to immobilised champedak lectin-affinity chromatography. Elution of bound glycoproteins was performed using 0.1 M melibiose. Pooled bound fractions were dialysed, concentrated, and subjected to 2DE and silver staining to identify their glycoprotein constituents. The polypeptide components that were stained included that of the heavy
## Table 1
Sequences of β chains of galactose-binding lectins from different species of *Artocarpus* lectins

<table>
<thead>
<tr>
<th>Lectin Chain</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. integer</em> a</td>
<td>(1) N E Q S G K S Q T M I M G P X G A Q M S T 25</td>
</tr>
<tr>
<td>(\beta_2)</td>
<td>(1) N E Q S G I S Q T V I G G P X G A Q V S T 25</td>
</tr>
<tr>
<td><em>A. tonkinensis</em> b</td>
<td>(1) K Q R S G E S Q N I I V G S W G A K V S 20</td>
</tr>
<tr>
<td><em>A. integriolia</em> c</td>
<td>(1) D E Q S G I S Q T V I V G P W G A Q V S 20</td>
</tr>
<tr>
<td><em>A. integriolia</em> d</td>
<td>(1) D E Q S G I S Q T V I V G P W G A K V S 20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>(\text{X} \ldots)</td>
</tr>
</tbody>
</table>

X: unidentified amino acid.

a This study.
b Mahanta et al. [19].
c Ruffet et al. [16].
d Kabir et al. [17].

## Table 2
N-terminal sequences of α chains of CGB lectin and jacalin

<table>
<thead>
<tr>
<th>Lectin</th>
<th>Sequence of the α chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residues 1–25</td>
<td></td>
</tr>
<tr>
<td>CGB lectina</td>
<td>G K A F D D G A F T G I R E I N L S Y N K E T A I 25</td>
</tr>
<tr>
<td>Jocalinb</td>
<td>I G K A F D D G A F T G I R E I N L S Y N K E T A I 25</td>
</tr>
<tr>
<td>Jocalinc</td>
<td>I G K A F D D G A F T G I R E I N L S Y N K E T A I 25</td>
</tr>
<tr>
<td>Jocalind</td>
<td>I G K A F D D G A F T G I R E I N L . . .</td>
</tr>
<tr>
<td>Jocalin ((\alpha)d)</td>
<td>I G K A F D D G A F T G I R E I N L S Y N K E T A I 25</td>
</tr>
<tr>
<td>Residues 26–47</td>
<td></td>
</tr>
<tr>
<td>CGB lectina</td>
<td>G D F Q V V Y D L N G S P Y V G Q N H K X F 47</td>
</tr>
<tr>
<td>Jocalinb</td>
<td>G D F Q V V Y D L N G S P Y V G Q N H K S F 47</td>
</tr>
<tr>
<td>Jocalinc</td>
<td>G D F Q V V Y D L N G S P Y V G Q N H T S F 47</td>
</tr>
<tr>
<td>Jocalin ((\alpha)d)</td>
<td>G D F Q . . .</td>
</tr>
</tbody>
</table>

X: unidentified amino acid.

a This study.
b Mahanta et al. [19].
c Ruffet et al. [16].
d Kabir et al. [17].
and light chains of IgA1, \( \alpha_2 \)-HS glycoprotein, \( \alpha_1 \)-antichymotrypsin, hemopexin, and a few unidentified glycopeptides (Fig. 4). The immunoglobulin light chains were actually polypeptide components of IgA1 that had not directly interacted with the lectin-affinity column. When a similar column was applied with purified human IgA1, the immunoglobulin was clearly retained and only released upon elution with 0.1M melibiose. IgA1 that was subjected to neuraminidase and O-glycosidase treatments was apparently not retained by the lectin-affinity column (Fig. 5).

**Discussion**

Champedak galactose-binding lectin, which binds to the Gal\( \beta_1 \)-3GalNAc moieties of glycoproteins, was earlier reported to comprise of two types of non-cova-

lently linked subunits with \( M_r \) of approximately 13,000 and 16,000 [1,3]. In the present study, the lectin was apparently resolved into two major (\( \alpha \) and \( \beta_1 \)) and one minor (\( \beta_2 \)) peaks when its purified preparation was subjected to reverse-phase HPLC. Purity of the peaks was obvious as it was possible to sequence their N-terminal amino acid residues. The first 47 residues of the broad \( \alpha \) subunit peak demonstrated at least 95% sequence homology to the N-terminal sequence of the \( \alpha \) subunits of the *A. integrifolia* lectin, jacalin [14–17].

The two other \( \beta_1 \) and \( \beta_2 \) peak fractions, each having 21 amino acid residues, demonstrated close homology to the \( \beta \) chains of galactose-binding lectins from *Artocarpus tonkinensis*, and *A. integrifolia*, when sequenced [14,18,19]. The champedak lectin \( \beta \) chains differed in only four amino acid residues at positions 6, 10, 12, and 19. Whilst the champedak \( \beta_1 \) chain shared more than 70% sequence homology with the \( \beta \) chains of the other *Artocarpus* seed galactose-binding lectins, the champe-

dak \( \beta_2 \) chain demonstrated at least 80% homology. The sequences of both the \( \beta_1 \) and \( \beta_2 \) subunits of the CGB
lectin appeared to lack cysteine, indicating the absence of intra/inter chain disulphide linkage. The unknown residue X was not cysteine as the lectin was reduced and alkylated prior to separation by reverse-phase HPLC.

When HRP-conjugated CGB lectin was allowed to interact with reduced human serum glycopeptides that were separated by 2DE and transferred onto a nitrocellulose membrane, strong apparent interactions were observed with IgA α chain, α1-antichymotrypsin, α2-HS glycoprotein, hemopexin, and a few unidentified glycopeptide spots. Since the former glycoproteins were identified by visual comparison with the SWISS-2DPAGE standard human plasma protein map [13], further confirmation was resorted by allowing the HRP-conjugated lectin to interact with purified IgA1, α1-antichymotrypsin, and α2-HS glycoprotein that were subjected to one-dimensional SDS–PAGE and Western blotting. The lack of purified hemopexin was compensated by inclusion of hemin-interactive serum protein fractions. Our data confirmed the interactions of the champedak lectin with all the human glycoproteins that were tested.

The interactions of the CGB lectin with IgA1, α2-HS glycoprotein, hemopexin, and α1-antichymotrypsin occurred via their O-glycosylated component chains. This is evident from the abrogation of the lectin binding when the O-glycans of the serum glycoproteins were enzymatically removed. In addition, our previous studies have demonstrated the inhibitory effects of carbohydrates with structures that constitute the O-glycan moieties [1]. Thus, the interaction of HRP-conjugated champedak lectin with IgA1 is restricted to its O-glycosylated α heavy chains [20]. However, its binding to α2-HS glycoprotein may have occurred on all component chains as both the α and β subunits of the glycoprotein were O-glycosylated [21]. Since hemopexin was reported to have only a single O-glycan structure, the ability of the champedak lectin to interact with human serum hemopexin demonstrated its potential use as an efficient indicator for the presence of O-glycan(s) in glycoproteins [22]. Unlike IgA1, α2-HS glycoprotein, and hemopexin, which are known to possess at least a single Galβ1–3GalNAc O-glycan moiety, α1-antichymotrypsin was previously reported to contain exclusively of the N-linked oligosaccharide structures [23]. The interaction of the CGB lectin with α1-antichymotrypsin suggests that the serum glycoprotein may also contain at least one O-linked Galβ1–3GalNAc structure.

In addition to the identified glycoproteins, our data also demonstrated the interactions of the champedak lectin with several unidentified glycoproteins. Like α1-antichymotrypsin, these glycopeptides are likely to be O-glycosylated. Some of these spots appeared in the immunoglobulin light chain region of the serum protein expression map. Thus, they may represent the minute proportion of the immunoglobulin light chains, which may have been O-glycosylated at their $F_{ab}$ regions because of the variability of their amino acid sequences. The other unknown glycopeptides appeared to have $pI$ distribution patterns like that of hemopexin and haptoglobin.

Immobilisation of the CGB lectin to Sepharose generated an affinity column chromatography capable of isolating the O-glycosylated glycoproteins. Whilst highly purified IgA1 bound strongly to the affinity column, non-O-glycosylated IgA1 was not retained and allowed to flow through. Subjecting normal human serum to the
affinity separation process was demonstrated to retain IgA1, α1-antichymotrypsin, α2-HS glycoprotein, and hemopexin. Some of the unidentified spots that were earlier detected were not present when the immobilised lectin retained fractions were analysed by 2DE and silver staining. In the presence of the various other highly reactive O-glycosylated serum glycoproteins, the interactions of the champedak lectin to the unknown glycoproteins may not be strong enough for their retention by the immobilised lectin-affinity column. Alternatively, the lack of lectin interaction with the unidentified glycoproteins also suggests that the O-glycan moieties may not be well exposed when the serum glycoproteins were in their native tertiary conformation.

Acknowledgments

This work was supported by IRPA Research Grants 06-02-03-0563 and 06-02-03-1005 awarded to Dr. O.H. Hashim by the Ministry of Science, Technology, and the Environment of Malaysia.

References