Comparison of the amino acid sequences of lectins from the seeds of \textit{Dioclea reflexa} and \textit{Canavalia ensiformis}

A. KUKU $^1$*, C. BALDUINI $^2$, A. COBIANCHI $^3$ and A. ABODERIN $^1$

$^1$ Department of Biochemistry, Obafemi Awolowo University, Ile-Ife, Nigeria.
$^2$ Dipartimento Di Biochimica, Universita degli studi di Pavia, Via Bassi 21, 27100 Pavia, Italia.
$^3$ Centro Grandi Strumenti, Universita degli studi di Pavia, Via Bassi 21, 27100 Pavia, Italia.

* Corresponding author, E-mail: akuku@oauife.edu.ng, adenikekuku@yahoo.com

\textbf{ABSTRACT}

The amino acid sequence of a glucose/mannose specific lectin from the seeds of \textit{Dioclea reflexa} (Dioclea \textit{reflexa} agglutinin II, DRA-II) was determined by sequential Edman analyses of the intact subunit, and of the peptides derived from the protein by enzymatic digestion with trypsin. This sequence was found to be very similar to those of the lectins from other \textit{Dioclea} species, \textit{Dioclea grandiflora} and \textit{Dioclea lehmanii}, and also the lectin from \textit{Canavalia ensiformis} (Con A). Comparison of these amino acid sequences showed that a high degree of homology exists among these proteins.

\textbf{INTRODUCTION}

Lectins are proteins or glycoproteins of non-immunoglobin nature, capable of specific recognition and reversible binding to carbohydrate moieties of complex glycoconjugates without altering the covalent structure of any of the recognized glycosyl ligands (D’adamo, 1999). These lectins bind to sugar moieties in cell walls or membranes and thereby change the physiology of the membrane to cause agglutination, mitosis or other biochemical changes in the cell.

Lectins are a group of intensely investigated proteins because of their many exploitable biological activities. These include: antiproliferative, antitumor, immunomodulatory (Wang et al., 2000; Yu et al., 2001, Singh-Bains et al., 2005), antifungal (Ye et al., 2001), antiviral (Ooi et al., 2000), and anti-insecticidal (Machuka & Oladapo, 2000) activities. They have been isolated from a diverse number of organisms encompassing plants, animals and microorganisms. (Knibbs et al., 1991; Pusztai, 1991; Wang et al., 1996; Inamori & Saito, 1999). Plant lectins, which are multivalent cell-agglutinating proteins, are widely used as powerful tools in cell-surface and complex carbohydrate research, as well as for the purification and characterization of glycoconjugates (Sharon and Lis, 1990; Goldstein and Poretz, 1986) due to their unique ability to bind specific sugars or sugar containing macromolecules (Goldstein and Hayes, 1978). Concanavalin A (Con A), the lectin from the seeds of \textit{Canavalia ensiformis} is one of the most investigated lectins (Cunningham et al., 1975; Terada et al., 1992) and belongs to the general group of Mannose/Glucose binding lectins (Kristiansen, 1974). Previous studies on the genus \textit{Dioclea} have also shown that the lectins from seeds of \textit{Dioclea grandiflora} (Ainouz et al., 1987; Perez et al., 1991) and \textit{Dioclea lehmanii} (Richardson et al., 1984) are glucose/mannose specific and are very similar in properties and amino acid sequences to the
Comparison of the amino acid sequences of lectins from these Dioclea species (Dioclea reflexa, Dioclea grandiflora and Dioclea lehmanii) and that of lectins from Canavalia ensiformis (Con A) has shown that a high degree of homology exists among these proteins. In addition, these lectins exhibit a circular permutation in their amino acid sequences when compared with those of lectins from seeds of other tribes of Leguminosae. This behaviour has been attributed to a transpeptidation event during post-translational processing of a precursor protein (Bowles et al., 1986).

In this paper, we report the amino acid sequence of DRA-II, one of the two lectins found in the seed of Dioclea reflexa, and further compare the sequence with those of lectins from two other Dioclea species and Canavalia ensiformis.

**MATERIALS AND METHODS**

**Materials**

Potassium dihydrogen orthophosphate, Glycine, Glutaraldehyde, Ammonium persulphate, N,N,N',N'-tetramethyl ethylene-diamine (TEMED), Trypsin, different monosaccharide sugars and cellobiose are all products of BDH Chemical Limited, Poole, England. All other reagents were of analytical grade and were purchased from Sigma Chemical Company, St.Louis, Mo, USA.

The seeds of Dioclea reflexa were collected from Ile-Ife in Nigeria and authenticated by the Department of Botany of Obafemi Awolowo University, Ile-Ife, Nigeria.

Human blood cells from the ABO system were obtained from healthy donors in the Department of Biochemistry, Obafemi Awolowo University, Ile-Ife.

**Preparation of Phosphate buffered saline (PBS), pH 7.2**

KH$_2$PO$_4$ (6.8 g) and NaCl (11.70 g) were mixed and dissolved in 1.5 litre of distilled water. The mixture was adjusted to pH 7.2 with 1M KOH and the volume was made up to 2 litres with distilled water.

**Preparation of crude seed extracts**

Mature seeds were removed from their husks and ground to a fine powdery form. One gram of seed powder was suspended in 10ml of PBS, pH 7.2 and the mixture stirred overnight at 4 °C after which it was centrifuged at 5000 rpm for 15 min. The supernatant which constitutes the lectin preparation (crude extract) was collected, filtered and stored at -20 °C until used.

**Affinity chromatography**

Sephadex G-150 (15 g) was swollen in distilled water for three days (sodium azide was added to prevent fungal growth). This was packed into a column (30 x 2.5cm) and equilibrated with PBS, pH 7.2. Ten millilitres of the crude extract was layered on the column, and after washing the sample into the gel bed, the column was connected to a buffer reservoir containing PBS. Fractions of 3 – 4 ml were collected until the absorbance at 280 nm was zero. The column was then washed with an additional 200 ml of the buffer, after which 0.2 M Glucose in PBS was applied to elute any bound lectin.

Fractions in the protein peaks were assayed for haemagglutinating activity as described by Bing et al. (1967). In a U-shaped microtitre plate arranged in rows of wells, 1:2 dilutions of the sample (each fraction) in PBS buffer were mixed with 50µl of 2% suspension of erythrocytes (the erythrocytes of human blood group A,B and O were fixed with 1% glutaraldehyde). The plate was left undisturbed for 30 min to 1 hr at room temperature in order to allow for agglutination of the erythrocytes to take place. The titre of the lectin was expressed as the reciprocal of the highest dilution showing visible agglutination of erythrocytes. The fractions eluted with glucose-buffer solution were pooled and dialysed exhaustively against PBS before haemagglutination assay was carried out. This bound lectin peak was labeled DRA–II.

The protein peaks containing haemagglutinating activity were pooled and dialysed against PBS, pH 7.2 at 4 °C. The unbound lectin peak (labeled DRA–I) was purified further by concentrating the pooled fractions and rechromatographing on the Sephadex G-150 column with PBS, pH 7.2.

**Amino acid analysis**

The amino acid composition of the native protein, DRA-II, was obtained after
hydrolysis in vacuo at 110 °C in 6N HCl for 24 hrs. The concentration of free amino acids were determined on a Hewlett-Packard Amino Quant Series II Amino Acid Analyser by precolumn derivatization technique employing both ortho-phthalaldehyde (OPA) and 9-fluorenylmethylchloroformate (FMOC) chemistry reaction according to the manufacturer’s protocol. Detection was performed by measuring UV absorbance at 338 and 262 nm respectively.

N-terminal group analysis
The N-terminal group of the native protein was determined by the automated Edman degradation method using the Hewlett Packard G 100 A protein sequencer system. The released amino acids were estimated after derivatization with phenylisothiocyanate.

Digestion of the protein with trypsin
One milligram of protein was digested with enzyme (ratio of enzyme to protein 1:25) according to the method of Aiken et al. (1989). This generated tryptic peptides. The peptides were purified by reverse-phase HPLC (µBond pack C 18 column) using a 50min linear gradient of 0% - 80% acetonitrile containing 0.05% trifluoroacetic acid. Elution was performed at room temperature at a flow rate of 1ml/min.

Sequence determination
Sequence determination of purified peptides and the intact protein was carried out using the adsorptive biphasic column technology using the HPG 1000A Protein sequencer (Hewlett-Packard, Palo Alto, CA). The peptides are numbered on the basis of progressive peptide sequencing starting from the N-terminal end.

Sequence analysis
The amino acid sequence of DRA II was compared and aligned with those of lectins from two other species of Dioclea and the lectin from Canavalia ensiformis using the hierarchical clustering procedure of Corpet (1988).

RESULTS AND DISCUSSION
The isolation and purification of the lectins in the saline extract of the seeds of Dioeclea reflexa are as summarized in the legends to Figure 1. Two lectins, DRA-I, which flows out of the column, and DRA-II which is adsorbed on the gel matrix and elutes upon the addition of glucose to the eluting buffer were obtained. DRA-I was found to be specific for N-Acetylglucosamine while DRA-II showed specificity for D-Mannose/D-Glucose as determined by Hapten-Inhibition studies using different monosaccharides (Goldstein and Hayes, 1978).

The purity of the proteins was established by both non-denaturing and denaturing polyacrylamide gel electrophoresis on 7.5% acrylamide gels (figure not shown).

Dioeclea reflexa resembles Dolichos lab lab (Mo et al., 1990) in containing two different kinds of lectins each with its own sugar specificity. Our preliminary observation (Adeyemi et al., 1994) and work suggests that DRA-I which preferentially binds N-Acetylglucosamine, is structurally a different protein from DRA-II.

In this report, DRA-II is the lectin of interest because it belongs to the group of lectins that include Concanavalin A, lentil lectin as well as lectins from the seeds of many papilionaceae plants for which Mannose and Glucose are the inhibitory sugars. Its amino acid sequence has been deduced from sequencing of the intact protein and also of the peptides arising from the degradation of the intact protein with trypsin (Figure 2 & 3).

The similarities in the amino acid composition of lectins from three species of the genus Dioeclea is summarized in Table 1, and at the level of the primary structure of the proteins in Figure 4. This figure also contains the comparative data from Con A.

Analysis of the data using the hierarchical clustering procedure of Corpet (1988), suggests that the degree of homology between DRA-II and the lectins from D. lehmanii, D. grandiflora and Con A is, respectively 86.6%, 94% and 75.5%. This degree of homology, buttressed by the invariance of the residues which form the binding site for carbohydrate, viz.: Tyr12, Asn14,Asp36,Leu89,Tyr100,Asp208, and Arg228, suggests that DRA-II must also be subject to the same post-translational circular permutation in its amino acid sequence as is found in Con A.

It should be observed that although there is an equally complete conversion of
residues whose side chains serve as ligands in Con A for Mn\(^{2+}\) (Glu\(^8\), Asp\(^{19}\), and His\(^{29}\)) on the one hand, and for Ca\(^{2+}\) (Asp\(^{10}\), Tyr\(^{12}\), Asn\(^{14}\) and Asp\(^{19}\)) on the other hand, we do not have any evidence so far that DRA-II requires these metal ions for sugar binding.

ACKNOWLEDGEMENTS

A. Kuku would like to acknowledge the support of the TRIL Programme of the Abdus Salam International Centre for Theoretical Physics, Trieste, Italy for financial support through a visiting research grant that made some aspect of this work possible.

Table 1: Amino acid composition of DRA-II, *D. lehmanii* and *D. grandiflora*.

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>DRA-II Sequence</th>
<th><em>D. lehmanii</em> Sequence</th>
<th><em>D. grandiflora</em> Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASX</td>
<td>34</td>
<td>33</td>
<td>34</td>
</tr>
<tr>
<td>GLX</td>
<td>11</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>SER</td>
<td>29</td>
<td>29</td>
<td>33</td>
</tr>
<tr>
<td>HIS</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>GLY</td>
<td>17</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>THR</td>
<td>20</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>ALA</td>
<td>16</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>ARG</td>
<td>9</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>TYR</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>CYS</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>VAL</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>MET</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PHE</td>
<td>11</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>ILE</td>
<td>15</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>LEU</td>
<td>18</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>LYS</td>
<td>12</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>PRO</td>
<td>10</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>TRP</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

(Sources: Perez et al., 1991; Richardson et al., 1984)

Figure 1: Affinity chromatography of *Dioclea reflexa* seed extract on Sephadex G-150 column. Eluent was PBS, pH 7.2. Protein bound to the column was eluted with 0.2 M Glucose in PBS. Column size: 30 x 2.5cm; Fraction size: 4.0ml; Flow rate: 16ml/hr.

- $$A_{280}$$ x dilution factor
- Agglutination Titre (AT)
Figure 2: Reverse phase HPLC separation of the tryptic digest of DRA-II. The peptides have been numbered on the basis of progressive sequence starting from the N-terminal end.

Figure 3: Amino Acid sequence of *Dioclea reflexa* Agglutinin DRA-II. The tryptic peptides (-----) are indicated as are also the non-enzymatically derived 44-residue peptide fragment (—) the structure of which conveniently provided an overlap for the tryptic peptides T12, T13 and T14. The N-terminal sequence is also derived from the intact protein (→).

Figure 4: Comparison of the Amino Acid sequence of Lectins from 1) *Dioclea reflexa* (DRA-II), 2) *Dioclea grandiflora*, 3) *Dioclea lehmanni* and 4) *Concanavalis ensiformis*. (Sequence homology among lectins from *Dioclea* spp. and Con A.).
REFERENCES


Sharon N, Lis H. 1990. Legume lectins: a large family of homologous proteins FASEB. J., 4: 3198-3208


