Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

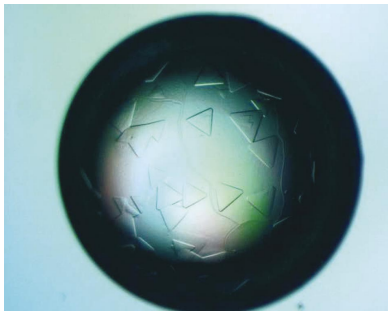
Crystallization and preliminary X-ray characterization of a lectin from *Cicer arietinum* (chickpea)

Uma V. Katre,[a] S. M. Gaikwad,[a] S. S. Bhagyawant,[b] U. D. Deshpande,[b] M. I. Khan[a] and C. G. Suresh[a]*

[a]Division of Biochemical Sciences, National Chemical Laboratory, Pune 411008, India, and [b]School of Life Sciences, S. R. T. M. University, Nanded 431606, India

Correspondence e-mail: suresh@ems.ncl.res.in

Received 1 October 2004
Accepted 6 December 2004
Online 24 December 2004



The lectin isolated from mature seeds of *Cicer arietinum* (CAL) agglutinates pronase-treated rabbit and human erythrocytes and its haemagglutination activity is inhibited by fetuin and desialated fetuin but not by simple monosaccharides or oligosaccharides. The purified lectin is a dimer of molecular weight 43 000 Da composed of two identical subunits (MW 21 500), as confirmed by SDS–PAGE. The lectin has been crystallized using the hanging-drop vapour-diffusion method at 295 K over a well solution containing 0.2 *M* sodium acetate, 0.1 *M* sodium phosphate buffer pH 6.5 and 14%(*w*/*v*) polyethylene glycol 8000. The triangular prism-shaped crystals belong to space group *R*3 and have unit-cell parameters $a = b = 81.2$, $c = 69.4$ Å. The diffraction data are 93.8% complete to 2.3 Å Bragg spacing with an $R_{\rm merge}$ of 0.103.

1. Introduction

Lectins are non-catalytic carbohydrate-binding proteins of non-immune origin that are widely distributed in microorganisms, plants and animals. They bind glycoconjugates or cell-surface carbohydrates and cause agglutination of erythrocytes and other cells. Lectins participate in many biological processes, including cell–cell recognition, host–pathogen interaction, serum glycoprotein turnover and innate immune responses (Sharon, 1993). The applications of lectins include use in structural characterization of complex carbohydrates; lectins are also useful as diagnostic tools and are employed in the protection of plants from pathogens *etc.* The crystal structures of several lectins from different sources and possessing varying specificities have been determined (Loris *et al.*, 1998). Most of these lectins are specific for simple sugars and their haemagglutination activity is inhibited by monosaccharides or oligosaccharides. In addition, lectins with complex specificity that are inhibited only by complex glycoproteins and not by simple sugars have also been reported (Guzmán-Partida *et al.*, 2004; Ueda *et al.*, 2002; Sato *et al.*, 2000; Singh *et al.*, 2004; Wright *et al.*, 1999). Recently, the structure of one such lectin from *Scilla campanulata* (SCAfet), specific for complex glycoproteins, has become available (PDB code 1dlp; Wright *et al.*, 2000).

Cicer arietinum, commonly known as the chickpea, is a pulse that is extensively grown and consumed in India. There is a previous report of the isolation of a lectin from chickpea (Kolberg *et al.*, 1983), but no further studies have been reported for this lectin. We have isolated and purified this *C. arietinum* lectin (CAL) possessing complex-sugar specificity from mature chickpea seeds. The molecular weight of the native protein as determined by gel filtration using HPLC is 43 000 Da. It has been identified as a homodimer of subunit molecular weight 21 500 Da by SDS–PAGE both in the presence and in the absence of β-mercaptoethanol. The lectin is basic in nature (pI 9.0) and is a glycoprotein containing 4.5% neutral sugars. Neither human (blood groups A, B and O) nor rabbit erythrocytes are agglutinated by this lectin; however, pronase-treated versions of both have been successfully agglutinated. None of the sugars, such as glucose, mannose or galactose, nor their derivatives, disaccharides, trisaccharides and tetrasaccharides have any effect on the agglutination activity of this lectin. The evidence for the complex specificity of CAL comes from the observation that the haemagglutination activity of 1 µg lectin can be inhibited using about 10 µg desialated fetuin. All

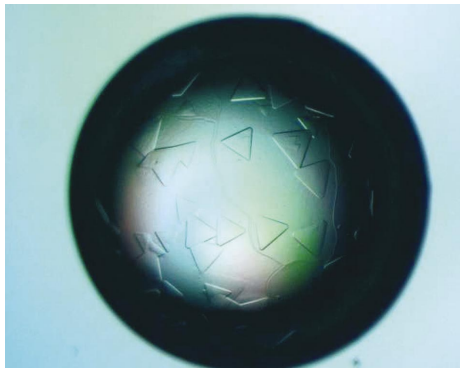


Figure 1
Crystals of *C. arietinum* lectin (CAL) grown in 0.1 *M* sodium phosphate buffer pH 6.5, 0.2 *M* sodium acetate and 14%(*w*/*v*) polyethylene glycol 8000 using the hanging-drop vapour-diffusion method.

the above observations for CAL grossly resemble those reported by Kolberg *et al.* (1983) for the lectin isolated by them.

2. Materials and methods

2.1. Purification of *C. arietinum* lectin

All purification steps were performed at 277 K unless otherwise stated. To start with, 100 g dry seeds of *C. arietinum* cv. BDN 9-3 were finely powdered and soaked in 500 ml 10 m*M* Tris–HCl, 150 m*M* NaCl pH 7.2. The suspension was stirred for 16 h and filtered through a muslin cloth. The filtrate was centrifuged at 10 000*g* for 20 min. The supernatant was precipitated with 80% saturated ammonium sulfate (AS) and centrifuged at 10 000*g* for 20 min. The precipitate was dissolved in 20 m*M* Tris–HCl buffer pH 7.2 and dialysed extensively against the same buffer. The dialysate was centrifuged at 10 000*g* for 10 min and the clear supernatant was loaded onto a DEAE-cellulose column pre-equilibrated with 20 m*M* Tris–HCl buffer pH 7.2. The column was washed with the same buffer until the wash showed no absorbance at 280 nm. The lectin eluted with the unadsorbed portion. Fractions (2 ml) of the wash with OD greater than 0.2 and possessing haemagglutination activity were pooled together, concentrated and dialysed against 20 m*M* acetate buffer pH 5.0. This sample was loaded onto an SP-Sephadex column (4 × 20 cm) pre-equilibrated with 20 m*M* acetate buffer pH 5.0 and washed with the same buffer. Finally, the bound protein was eluted with acetate buffer pH 5.0 containing NaCl in a step gradient of 0.1–0.5 *M* NaCl in steps of 0.1 *M*. The fractions with OD greater than 0.2 and possessing haemagglutination activity were pooled and dialysed against de-ionized water. The final yield of the lectin was 150 mg from 100 g dried seeds, with a specific activity of 5×10^4 U mg^{-1}. The protein solution was concentrated to 15 mg ml^{-1} in deionized water prior to setting up crystallization experiments.

2.2. Circular-dichroism experiment

The CD spectrum of CAL was recorded on a Jasco J-500A spectropolarimeter using a sample cell of path length 1 mm. The protein solution used had a concentration of 25 µ*M* and was prepared in 20 m*M* phosphate buffer pH 7.2. The spectra were collected with a response time of 4 s and a scan speed of 100 nm s^{-1} in the wavelength range 200–250 nm. Each data point is an average of six accumulations. The software package *K2D* (Andrade *et al.*, 1993) was used to predict the secondary structure from the UV-CD spectral values.

Table 1
Diffraction data statistics for CAL.

Values in parentheses are for the last shell.

Temperature (K)	295
X-ray source	Cu *Kα*
Wavelength (Å)	1.542
Resolution limits (Å)	30.0–2.30 (2.38–2.30)
No. measured reflections	126672
No. unique reflections	7121
Completeness of data (%)	93.8 (97.9)
R factor (%)	10.3 (18.9)
Average $I/\sigma(I)$	8.8 (4.3)
Mosaicity (°)	0.3
Space group	*R*3
Unit-cell parameters (Å)	$a = b = 81.2$, $c = 69.4$
Unit-cell volume (Å^3)	396613
Matthews coefficient (Å^3 Da^{-1})	2.2
No. molecules per unit cell (*Z*)	9
Solvent content (%)	44.2

2.3. Crystallization

Initial crystallization trials were conducted at 295 K with the sparse-matrix screen Crystal Screen I (Hampton Research) employing the hanging-drop vapour-diffusion method. The hanging drops, which were set up on siliconized cover slips over 0.5 ml reservoir solutions in multi-well trays, contained 1 µl protein solution mixed with 1 µl reservoir solution. Rhombohedral crystals grew within 6 d from conditions Nos. 28, 40 and 46 of Crystal Screen. The crystals grown from condition No. 28 [0.2 *M* sodium acetate, 0.1 *M* sodium cacodylate pH 6.5, 30%(*w*/*v*) polyethylene glycol (PEG) 8000] were found to be suitable for X-ray diffraction data collection. The quality of the crystals was further improved by refining the above condition, particularly with regard to buffer choice and PEG concentration. Finally, a solution containing 0.2 *M* sodium acetate, 0.1 *M* sodium phosphate buffer pH 6.5 and 14%(*w*/*v*) PEG 8000 yielded the best diffraction-quality crystals grown at 295 K.

2.4. Data collection and processing

A single crystal chosen from the drop was mounted in a thin-walled glass capillary of diameter 1.0 mm and X-ray diffraction data were collected at 295 K on an R-AXIS IV^{++} image plate using Cu *Kα* radiation generated by a Rigaku rotating-anode X-ray generator operated at 50 kV and 100 mA and equipped with a confocal mirror focusing system. The crystal-to-detector distance was kept at 150 mm and each frame was recorded with 0.5° crystal oscillation and 120 s exposure time. The diffraction data consisted of 180 images. The data were processed with *DENZO* and scaled using *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results and discussion

The secondary-structural components for this lectin estimated from CD experiments are 34% helix, 28% *β*-sheet and 38% random coil; the helix component is unexpected for a legume lectin. Thus, the structure of CAL seems to differ from the characteristic antiparallel *β*-sheet structure of legume lectins (Loris *et al.*, 1998; Chandra *et al.*, 2001) and, unlike other legume lectins, CAL lacks specificity for simple sugars or sugar derivatives (Reeke & Becker, 1988).

The crystals of CAL (Fig. 1) diffract to a resolution of 2.3 Å and belong to the rhombohedral space group *R*3, with unit-cell parameters $a = b = 81.2$, $c = 69.4$ Å. Data-collection and processing statistics are summarized in Table 1. A total of 126 672 measured reflections were merged into 7121 unique reflections with an $R_{\rm merge}$ of 10.3%. The merged data set was 93.8% complete to a Bragg spacing

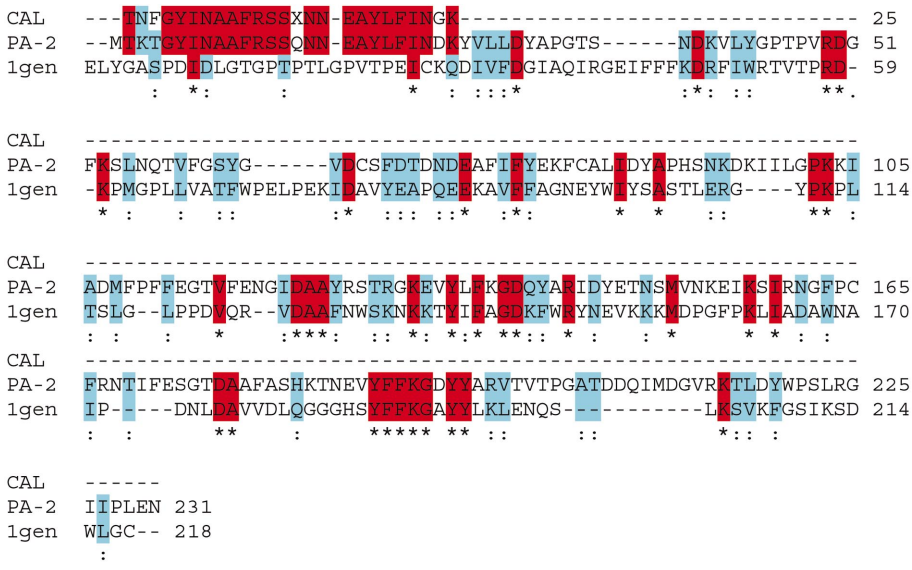


Figure 2
Alignment of *C. arietinum* lectin (CAL) N-terminal sequence with that of albumin 2 (PA-2) from *P. sativum* and also with the C-terminal domain of gelatinase A for which the structure is available in the Protein Data Bank (PDB code 1gen). (The identical amino acids are highlighted in red and marked with a star and the conservatively substituted amino acids are in blue and marked with a colon). The alignment was produced using *ClustalW* v.1.82 (Thompson *et al.*, 1994).

of 2.3 Å. No data beyond this resolution were included because further extending the resolution (to 2 Å) sharply increased $R_{\rm merge}$ and correspondingly reduced the signal-to-noise ratio. Attempts at low temperature (100 K) data collection with cryoprotectants such as 20–30%(*v*/*v*) PEG 200, PEG 400, glycerol or 2,4-methylpentanediol (MPD) did not succeed owing to high mosaicity and poor diffraction quality of the flash-cooled crystals. The number of molecules in the crystallographic asymmetric unit was estimated with the Matthews Probability Calculator, with the resolution as an additional input (Kantardjieff & Rupp, 2003). The highest probability (1.00) was obtained for one monomer in the asymmetric unit and the Matthews coefficient $V_{\rm M}$ (Matthews, 1968) was estimated to be 2.2 Å^3 Da^{-1}, corresponding to a solvent content of 44.2%.

N-terminal sequencing of the CAL preparation has confirmed the previously reported sequence of the first 25 amino acids of the peptide chain (Kolberg *et al.*, 1983); a *BLAST* search (Marchler-Bauer *et al.*, 2003) based on the above partial sequence against a non-redundant database discloses a match at 90% identity (Fig. 2) with the N-terminal sequence of a major seed albumin (PA-2) from *Pisum sativum* [National Center for Biotechnology Information (NCBI) accession No. P08688; Higgins *et al.*, 1987]. Since no three-dimensional structure of this plant albumin is available, its full sequence has been used for comparison with the sequences of structures deposited in the Protein Data Bank (PDB). The *BLAST* search, now using the PA-2 sequence, had 28% identity (Fig. 2) with the C-terminal domain of gelatinase A (PDB code 1gen; Libson *et al.*, 1995). Thus, the presently available partial sequence of CAL suggests that the structure of C-terminal domain of gelatinase A, although only distantly related to CAL, might be used for structure solution by the molecular-replacement method. Regretably, molecular-replacement calculations with this gelatinase A structure as the model did not yield a solution. The above indirect attempt at structure solution has been attempted only because no full sequence of CAL is presently available.

Our efforts are now to grow crystals of heavy-atom derivatives and thus to determine the structure of *C. arietinum* lectin through MIR phasing. Simultaneously, attempts are being made to determine the full amino-acid sequence of CAL through biochemical methods or gene sequencing.

UVK is a Junior Research Fellow of the Council of Scientific and Industrial Research, New Delhi. The authors thank Ms Sushma Nagpal of the National Institute of Immunology, Delhi for determining the N-terminal sequence of CAL.

References

Andrade, M. A., Chacón, P., Merelo, J. J. & Morán, F. (1993). *Protein Eng.* **6**, 383–390.
Chandra, N. R., Prabu, M. M., Suguna, K. & Vijayan, M. (2001). *Protein Eng.* **14**, 857–866.
Guzmán-Partida, A. M., Robles-Burgueño, M. R., Ortega-Nieblas, M. & Vázquez-Moreno, I. (2004). *Biochimie*, **86**, 335–342.
Higgins, T. J. V., Beach, L. R., Spencer, D., Chandler, P. M., Randall, P. J., Blagrove, R. J., Kortt, A. A. & Guthrie, R. E. (1987). *Plant Mol. Biol.* **8**, 37–45.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
Kolberg, J., Michaelsen, T. E. & Sletten, K. (1983). *Hoppe-Seyler's Z. Physiol. Chem.* **364**, 655–664.
Libson, A. M., Gittis, A. G., Collier, I. E., Marmer, B. L., Goldberg, G. I. & Lattman, E. E. (1995). *Nature Struct. Biol.* **2**, 938–942.
Loris, R., Hamelryck, T., Boukaert, J. & Wyns, L. (1998). *Biochim. Biophys. Acta*, **1383**, 9–36.
Marchler-Bauer, A. *et al.* (2003). *Nucleic Acids Res.* **31**, 383–387.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Reeke, G. N. Jr & Becker, J. W. (1988). *Curr. Top. Microbiol. Immunol.* **139**, 35–58.
Sato, Y., Murakami, M., Miyazawa, K. & Hori, K. (2000). *Comput. Biochem. Physiol. B*, **125**, 169–177.
Sharon, N. (1993). *Trends Biochem. Sci.* **18**, 221–226.
Singh, J., Singh, J. & Kamboj, S. S. (2004). *Biochem. Biophys. Res. Commun.* **318**, 1057–1065.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Ueda, H., Matsumoto, H., Takahashi, N. & Ogawa, H. (2002). *J. Biol. Chem.* **277**, 24916–24925.
Wright, L. M., Reynolds, C. D., Rizkallah, P. J., Allen, A. K., Van Damme, E. J. M., Donovan, M. J. & Peumans, W. J. (2000). *FEBS Lett.* **468**, 19–22.
Wright, L. M., Van Damme, E. J. M., Barre, A., Allen, A. K., Van Leuven, F., Reynolds, C. D., Rouge, P. & Peumans, W. J. (1999). *Biochem. J.* **340**, 299–308.