Acta Cryst. (1998). D**54**, 844–847

Quaternary Structure of UEA-II, the Chitobiose Specific Lectin from Gorse

MINH-HOA DAO-THI,[a]* PIERRE RIZKALLAH,[b] LODE WYNS,[a] FREDDY POORTMANS[c] AND REMY LORIS[a]

[a]*Laboratorium voor Ultrastruktuur, Vlaams Interinuversitair Instituut voor Biotechnologie, Vrije Universiteit Brussel, Paardenstraat 65, B-1640 Sint-Genesius-Rode, Belgium,* [b]*Daresbury Laboratory, Warrington WA4 4AD, England, and* [c]*Vlaamse Instelling voor Technologisch Onderzoek, Boeretang 200, B-2400 Mol, Belgium. E-mail: rloris@vub.ac.be*

(*Received 27 October 1997; accepted 19 January 1998*)

Abstract

The chitobiose specific *Ulex europaeus* lectin II crystallizes in space group $P3_221$ with unit-cell dimensions $a = b = 105.54$, $c = 176.26$ Å. The asymmetric unit contains a complete lectin tetramer. The crystals were shown to diffract to 4.5 Å on a rotating-anode source and to 2.7 Å at the Daresbury synchrotron source. Molecular replacement and subsequent rigid-body refinement using data to 4.5 Å yielded a solution corresponding to a tetramer very similar to that of phytohemagglutinin-L and soybean agglutinin. The monomers in the *Ulex* lectin tetramer are rotated ~5° compared with the phytohemagglutinin-L and soybean agglutinin structures.

1. Abbreviations

UEA-I, lectin I from *Ulex europaeus*; UEA-II, lectin II from *Ulex europaeus*; PHA-L, phytohemagglutinin-L from *Phaseolus vulgaris*; SBA, soybean agglutinin; GS-IV, lectin IV from *Griffonia simplicifolia*; EcorL, lectin from *Erythrina corallodendron*; PEG, polyethylene glycol.

2. Introduction

The legume lectins are a large family of homologous carbohydrate-binding proteins that are found in the seeds of most legume plants (Van Driessche, 1988). Despite their strong similarity on the level of their amino-acid sequences and three-dimensional structures, their carbohydrate specificities vary widely. Because of their abundance in mature seeds and their broad range of carbohydrate specificities, they are used as model systems to study the nature of protein–carbohydrate interactions (Loris *et al.*, 1998).

The structures of several legume lectins have been determined to date, including lentil lectin, soybean agglutinin, concanavalin A, pea lectin, isolectins I and II from *Lathyrus occhrus*, lectin IV from *Griffonia simplicifolia*, the lectin from *Erythrina corallodendron*, phytohemagglutinin-L and peanut agglutinin (Loris *et al.*, 1993; Loris, Van Overberge *et al.*, 1994; Hardman *et al.*, 1982; Naismith *et al.*, 1993; Einspahr *et al.*, 1986; Bourne *et al.*, 1990; Delbaere *et al.*, 1993; Shaanan *et al.*, 1991; Dessen *et al.*, 1995; Banerjee *et al.*, 1994; Hamelryck *et al.*, 1996). These proteins possess very similar tertiary structures, dominated by two antiparallel β-sheets. Their quaternary structures, on the contrary, are not conserved and seem to be mediated not only by differences in amino-acid sequence at the possible combining sites, but also by the presence or absence of covalently linked carbohydrate.

Legume lectins can be classified into five different groups on the basis of their sugar specificity: glucose/mannose, galactose/*N*-acteyl-galactosamine, fucose, chitobiose and complex (*i.e.* no simple mono- or disaccharide is capable of inhibiting agglutination). At present, crystal structures of carbohydrate complexes are available for the glucose/mannose group of lectins (concanavalin A, lentil, pea and *Lathyrus* lectins), the galactose/*N*-acetyl-galactosamine group of lectins (peanut agglutinin, EcorL and soybean lectin) and one lectin with complex specificity (GS-IV), while the crystallization of one fucose specific lectin (UEA-I) has been reported (Vandonselaar & Delbaere, 1994).

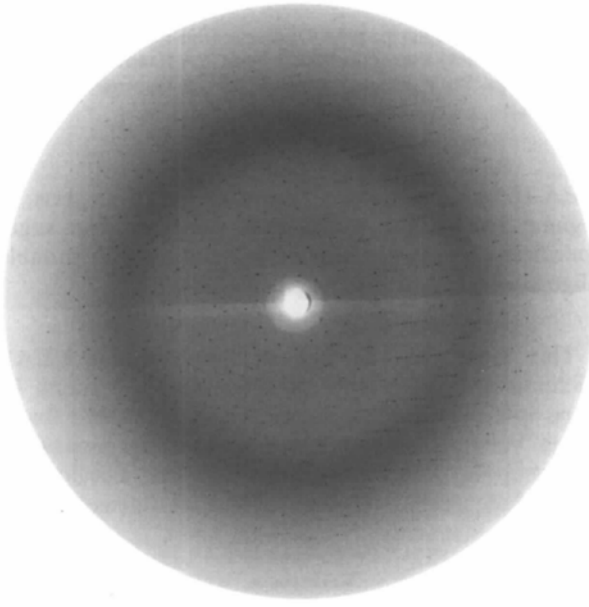
Here we report on the crystallization of UEA-II, the chitobiose specific lectin from *Ulex europaeus*. UEA-II was first purified in the early 1970's (Matsumoto & Osawa, 1970), but its amino-acid sequence was determined only 6 years ago (Konami *et al.*, 1991*a*). The lectin forms a homotetramer in solution and contains 22% covalently bound carbohydrate (Matsumoto & Osawa, 1970), mainly mannose, galactose and arabinose, and in this respect resembles the related lectins from *Cytisus sessifolius* and *Laburnum alpinum* (Konami *et al.*, 1983, 1991*b*). Although the initial studies identified chitobiose as the most potent inhibitory sugar (Matsumoto & Osawa, 1970), it was found that Fucα1-2Galβ1-4GlcNac and Fucα1-2Galβ1-4Glc are better ligands by a factor of 4, while GalNAcα1-3Galβ1-3GlcNac also shows some affinity (it being about half as potent as chitobiose) (Pereira *et al.*, 1979).

3. Materials and methods

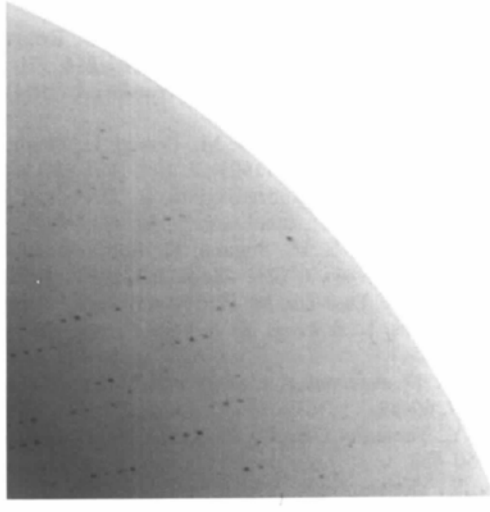
Purified *Ulex europaeus* lectin II was obtained from Sigma as highly purified lyophilized powder (Catalog No. L-6391, Lot 12 H4010) and used without further purification. The protein was dissolved in water, desalted and then brought to PBS buffer using a PD 10 Sepharose column. Crystallization conditions were screened using the hanging-drop method and small twinned needle-shaped crystals appeared after a few weeks in drops containing 5 µl of the 2.0–5.0 mg ml^{-1} protein solution mixed with an equal amount of 100 m*M* cacodylate or MES buffer (pH 6.5) containing 8–16%(*w*/*v*) PEG 6000 and equilibrated against the same precipitant solution. Larger single crystals, suitable for X-ray diffraction experiments were grown by micro-seeding with the initial twinned crystals, followed by washing and re-seeding of second generation single crystals as macro-seeds using the same conditions. These crystals grow to a size of 0.7 × 0.05 × 0.05 mm in a period of a few weeks. Data were collected on a small MAR image plate to 4.5 Å resolution, using 1.4° of rotation for each frame. In this way, 60 frames were collected that were integrated and merged with the *CCP*4 suite of programs for X-ray crystallography (Collaborative Computational Project, Number 4, 1994). Molecular replacement calculations and rigid-body refinement were performed with the program *AMoRe* (Navaza, 1994)

Table 1. *Data-collection statistics*

Space group	$P3_221$
Unit-cell dimensions (Å)	$a = b = 105.54$, $c = 176.26$
Resolution (Å)	50.0–4.5
Number of measured reflections	37394
Number of unique reflections	6752
Completeness (%)	99.2
R_{merge} (last shell)	0.098 (0.205)
$I\sigma(I)$ (last shell)	10.5 (5.4)



(*a*)



(*b*)

Fig. 1. (*a*) 1° rotation picture taken at station 7.2 of the Daresbury synchrotron. The crystal is oriented with its *c** axis approximately parallel to the camera axis around which the crystal was rotated during the exposure. (*b*) Enlargement of a portion of the diffraction pattern showing diffraction at the edge of the image (2.7 Å).

4. Results and discussion

The crystals are trigonal, space group $P3_221$, with unit-cell dimensions $a = b = 105.54$, $c = 176.26$ Å, $\alpha = \beta = 90$, $\gamma = 120°$. Assuming a tetramer in the asymmetric unit gives a Matthews coefficient of 2.68 Å^3 Da^{-1} and a solvent content of 45%, well within the values found for other protein crystals. A 4.5 Å data set could be collected using a conventional X-ray source. Preliminary tests at station 7.2 of the Daresbury synchrotron indicated that the crystals diffract to about 2.7 Å, (Fig. 1). High-resolution data collection and structure refinement, therefore, will have to await the availability of synchrotron beamtime (Table 1).

Legume lectins are known to adopt a wide variety of different dimers and tetramers. Therefore, no initial assumption on the quaternary structure of UEA-II could be made. Thus, molecular replacement was tried with the different legume lectin monomers, dimers and tetramers. Since of all the legume lectins with known crystal structures, PHA-L shows the highest percentage of sequence identity, the six known dimers and tetramers were constructed using PHA-L monomers. The clearest rotation-function solution was obtained when using the PHA-L type of tetramer as search model. The same solution was also be obtained when using the canonical

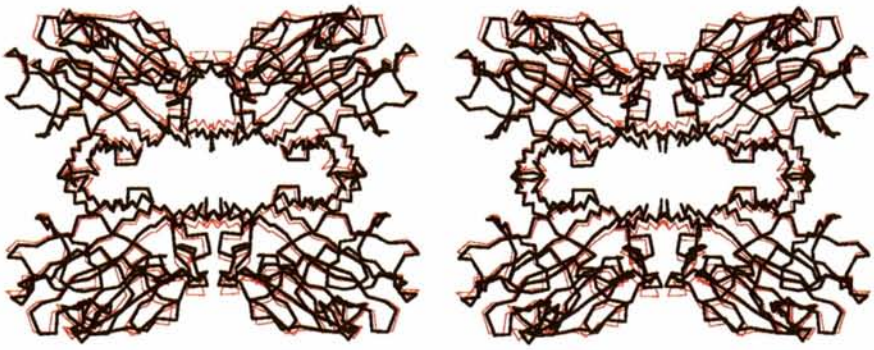


Fig. 2. Stereoview of a superposition of a Cα trace of the UEA-II tetramer (thick black lines) upon the PHA-L tetramer (thin red lines) illustrating the 5° rotation of the UEA-II subunits compared to PHA-L. The three molecular twofold axes of the PHA-L and UEA-II tetramers coincide and are oriented horizontally, vertically and perpendicular to the plane of the page. The 5° rotation is around the axis perpendicular to the plane of the page. The figure was prepared using *MOLSCRIPT* (Kraulis, 1991).

dimer (such as lentil lectin) as the search model, but not with the GS-IV or EcorL types of dimers, nor with the concanavalin A or peanut agglutinin type of tetramers. The translation search was, therefore, continued using only the PHA-L type of tetramer and established the true space group as $P3_221$.

Thus, UEA-II is the third legume lectin that forms the PHA-L type of tetramer. The two other observed tetramer types, the concanavalin A type and the peanut agglutinin type, on the contrary have only been observed once. This suggests that the PHA-L type of tetramer might be the 'default' or 'canonical' legume lectin tetramer. This tetramer has point group 222 symmetry and is a 'dimer of dimers'. Furthermore, the interface in this tetramer has been suggested to form the binding site for adenine and adenine-derived plant hormones (Hamelryck *et al.*, 1996). Therefore, it will be of interest to determine whether UEA-II is also capable of interacting with these ligands with high affinity. Refinement of the UEA-II structure at high resolution using synchrotron data will undoubtedly provide further clues about the molecular basis for quaternary structure formation. The UEA-II tetramer, although very similar, is not entirely identical to the PHA-L and SBA tetramers (Fig. 2). A small but significant rotation of the subunits of about 5° was found compared to PHA-L and SBA after rigid-body refinement of the individual subunits. This led to a reduction of the *R* factor from 0.472 to 0.404 at 4.5 Å resolution and a corresponding improvement of the correlation coefficient from 0.39 to 0.58.

There is no apparent relationship between sequence identity and the nature of the quaternary structure. Lectins with a quite high sequence identity such as EcorL and lentil lectin (70% amino-acid identity) form a different type of dimer, while lentil lectin and concanavalin A with only 39% sequence identity form the same type of dimer. Quaternary structure is also not correlated with carbohydrate specificity. Furthermore, UEA-II belongs to the chitobiose specific class of legume lectins, for which data on their fine sugar specificity are scarce. Co-crystallizations with suitable oligosaccharides are in progress.

This work was funded in part by the VLAB Biotechnology project of the Flemish government. R. Loris is a postdoctoral fellow of the FWO. We thank Maria Vanderveken for excellent technical assistance. Professor R. Palmer (Birbeck College, London) is acknowledged for providing facilities for data collection.

References

Banerjee, R., Mande, S. C., Ganesh, V., Das, K., Dhanaraj, V., Mahanta, S. K., Suguna, K., Surolia, A. & Vijiyan, M. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 227–231.

Bourne, Y., Abergel, C., Cambillau, C., Frey, M., Rougé, P. & Fontecilla-Camps, J.-C. (1990). *J. Mol. Biol.* **214**, 571–584.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Delbaere, L. T. J., Vandonselaar, M., Prasad, L., Quail, J. W., Wilson, K. S. & Dauter, Z. (1993). *J. Mol. Biol.* **230**, 950–965.

Dessen, A., Gupta, D., Subramaniam, S., Brewer, C. F. & Sachettini, J. C. (1995). *Biochemistry*, **34**, 4933–4942.

Einspahr, H., Parks, E. H., Suguna, K., Subramanian, E. & Suddath, F. L. (1986). *J. Biol. Chem.* **261**, 16518–16527.

Hamelryck, T. W., Dao-Thi, M.-H., Poortmans, F., Chrispeels, M. J. V, Wyns, L. & Loris, R. (1996). *J. Biol. Chem.* **271**, 20479–20485.

Hardman, K. D., Agarwal, R. C. & Freiser, M. J. (1982). *J. Mol. Biol.* **157**, 69–89.

Konami, Y., Yamamoto, K. & Osawa, T. (1991*a*). *J. Biochem.* **109**, 650–658.

Konami, Y., Yamamoto, K. & Osawa, T. (1991*b*). *Biol. Chem. Hoppe-Seyler*, **372**, 103–111.

Konami, Y., Yamamoto, K., Tsuji, T., Matsumoto, I. & Osawa, T. (1983). *Hoppe-Seyler's Z. Physiol. Chem.* **364**, 397–405.

Kraulis, P. J (1991). *J. Appl. Cryst.* **24**, 946–950.

Loris, R., Hamelryck, T., Bouckaert, J. & Wyns, L. (1998). *Biochim. Biophys. Acta*, **1383**, 9–36.

Loris, R., Steyaert, J., Maes, D., Lisgarten, J., Pickersgill, R. & Wyns, L. (1993). *Biochemistry*, **32**, 8772–8781.

Loris R., Van Overberge, D., Dao-Thi, M.-H., Poortmans, F., Maene, N. & Wyns, L. (1994). *Proteins Struct. Funct. Genet.* **20**, 330–346.

Matsumoto, I. & Osawa, T. (1970). *Arch. Biochem. Biophys.* **140**, 585–491.

Naismith, J. H., Habash, J., Harrop, S., Helliwell, J. R., Hunter, W. N., Wan, T. C. M., Weisgerber, S., Kalb (Gilboa), A. J. & Yariv, J. (1993). *Acta Cryst.* D**49**, 561–571.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Pereira, M. E. A., Gruezo, F. & Kabat, E. A. (1979). *Arch. Biochem. Biophys.* **194**, 511–525.

Shaanan, B., Lis, H. & Sharon, N. (1991) *Science*, **254**, 862–866.

Vandonselaar, M. & Delbaere, L. T. J. (1994). *J. Mol. Biol.* **243**, 345–346.

Van Driessche, E. (1988). *Advances in Lectin Research*, Vol. 1, edited by H. Franz, pp. 73–134. Berlin/Heidelberg: Springer Verlag.