Acta Cryst. (1998). D**54**, 1446–1449

Crystallization of two related lectins from the legume plant *Dolichos biflorus*

Minh-Hoa Dao-Thi,[a] Thomas W. Hamelryck,[a]* Julie Bouckaert,[a] Fritjof Körber,[b] Valentina Burkow,[c] Freddy Poortmans,[d] Marilynn Etzler,[e] Gerard Strecker,[f] Lode Wyns[a] and Remy Loris[a] *at* [a]*Laboratorium voor Ultrastructuur, Vlaams Interuniversitair Instituut voor Biotechnologie, Vrije Universiteit Brussel, Paardenstraat 65, B-1640 Sint-Genesius-Rode, Belgium,* [b]*Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, England,* [c]*Department of Chemistry, University of Tromso, N900 Tromso, Norway,* [d]*Vlaamse Instelling voor Technologisch Onderzoek – VITO, Boeretang 200, B-2400 Mol, Belgium,* [e]*Section of Molecular and Cellular Biology, University of California, Davis, CA 95616, USA, and* [f]*Université des Sciences et Technologies de Lille, Bâtiment C9, 59655 Villeneuve D'Ascq CEDEX, France. E-mail: thamelry@vub.ac.be*

(*Received 8 January 1998; accepted 23 March 1998*)

Abstract

The seed lectin DBL and the related stem and leaves lectin DB58 of the tropical legume *Dolichos biflorus* were crystallized, as well as complexes of DBL with adenine and with GalNAc(α1-3)[Fuc(α1-2)]Gal. The different crystal forms of DBL diffract to about 2.8 Å, while DB58 crystals diffract to 3.3 Å.

1. Introduction

The tropical legume *Dolichos biflorus* contains several lectins (Etzler, 1996), of which the seed lectin DBL (Etzler & Kabat, 1970) and the stem and leaves lectin DB58 (Schnell & Etzler, 1988) are best characterized. These two proteins share 87% sequence identity (Harada *et al.*, 1990), bind to trisaccharide GalNAc(α1-3)[Fuc(α1-2)]Gal containing blood-group A substance (Etzler, 1994) and bind adenine and adenine-derived plant hormones (Gegg *et al.*, 1992; Gegg & Etzler, 1994). Despite these similarities, they also show some important differences, which make them an interesting model system. DBL is a heterotetramer containing two long (subunit I) and two short subunits (subunit II), the short subunits being generated by post-translational C-terminal truncation (Young *et al.*, 1995). Only long-subunit aggregates are thought capable of binding carbohydrates, and their activity is enhanced by the presence of the short subunits (Etzler *et al.*, 1981). Haemagglutination is inhibited by GalNAc, but not by Gal. DB58, on the other hand, is a dimer in solution. It also consists of a long and a short subunit, making the dimer monovalent for carbohydrate binding. In contrast to DBL, DB58 does not seem to bind any simple monosaccharide (Etzler, 1994). In this work, we report the crystallization of DBL and DB58 as well as the complexes of DBL with adenine and with GalNAc(α1-3)[Fuc(α1-2)]Gal.

2. Material and methods

Recombinant DBL subunit I and recombinant DBL subunits I + II were expressed in *E. coli* as described by Chao *et al.* (1994). Recombinant DBL, natural DBL and DB58 were purified using a blood-group A + H substance-Sepharose affinity column. The trisaccharide GalNAc(α1-3)[Fuc(α1-2)]Gal was isolated from human urine (Strecker *et al.*, 1989). zCrystallization conditions were screened using the hanging-drop method. The refined conditions that led to the best crystals are given in Tables 1 and 2. Crystals were mounted in borosilicate capillaries and X-ray data were collected at room temperature using a MAR image plate. Data for crystal form I of uncomplexed DBL and for DB58 were collected at station 7.2 of the Daresbury synchrotron. All data were processed and scaled with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The details of the processing are also given in Tables 1 and 2.

3. Results and discussion

3.1. *Uncomplexed DBL*

Initially, crystallization conditions were screened for the natural as well as the two recombinant forms. The first crystals we obtained came from preparations of the recombinant subunit I. However, it is known that pure subunit I, when stored for several weeks at room temperature, degrades to an equal mixture of subunits I and II. Gel electrophoretic analysis of some of the crystals we obtained indicated that this happened with the protein in the crystals.

Three crystal forms of the uncomplexed lectin were grown (Table 1). The first form is orthorhombic, space group $P2_12_12_1$ (a = 81.65, b = 116.46, c = 225.39 Å), and contains two tetramers in its asymmetric unit. This crystal form diffracts to a resolution of about 3.3 Å on a rotating-anode source and to 2.75 Å using synchrotron radiation (see Fig. 1). Isomorphous crystals were also obtained in the presence of adenine.

The second crystal form belongs to the lower symmetry space group $P2_1$ and diffracts to 3.0 Å resolution on a rotating-anode source. The crystals contain a single tetramer in their asymmetric unit and have unit-cell dimensions a = 79.18, b = 101.21, c = 57.53 Å, β = 91.90°. Often, a small number of weak reflections can be observed that belong to the superlattice a = 96.39, b = 101.21, c = 99.49 Å, β = 108.08°. Although the relative intensity of these reflections varies from crystal to crystal, they are generally very weak and not observable beyond 4.0 Å. Both form I and form II crystals have a similar plate-shaped morphology and are difficult to distinguish by eye.

Subsequently, small crystals were also grown of the recombinant co-expressed subunits I + II. These crystals are isomorphous to the monoclinic form II crystals described above, but diffract to only about 4.5 Å resolution. Despite the poor quality of these crystals, this result is nevertheless of importance, as it shows us that the post-translational trimming of the C-terminus of half the subunits does not change the quaternary structure of the protein.

We have so far failed to produce single diffracting crystals of the natural (glycosylated) seed lectin.

ISSN 0907-4449

Table 1. *Crystal forms of uncomplexed DBL and DB58*

	DBL subunit I form I	DBL subunit I form II	DBL subunit I + II	DB58
Protein solution	5 mg ml^{-1} protein	5 mg ml^{-1} protein	5 mg ml^{-1} protein	5–7 mg ml^{-1} protein
Precipitant solution	100 m*M* sodium cacodylate pH 6.5, 8–12%(*w*/*v*) PEG 8000, 20 m*M* $MgAc_2$	100 m*M* sodium cacodylate pH 6.5, 10%(*w*/*v*) PEG 8000, 20 m*M* $MgCl_2$	valign="top"100 m*M* HEPES pH 7.5, 20%(*w*/*v*) PEG 20000, 10%(*v*/*v*) 2-propanol	100 m*M* sodium citrate pH 5.6, 8–12%(*w*/*v*) PEG 6000, 0.2–0.4 *M* NaCl
Space group	$P2_12_12_1$	$P2_1$	$P2_1$	$P2_12_12_1$
Unit cell				
a (Å)	81.65	79.18	79	101.46
b (Å)	116.46	101.21	101	130.73
c (Å)	225.39	57.53	58	136.88
β (°)		91.90	92	
Contents of unit cell	2 tetramers	1 tetramer	—	3 dimers
Resolution (Å)	2.75†	3.0	<4.5	3.3†
R_{merge}	0.096†	0.108	—	0.135†
Completeness	89.8%†	93.5%	—	80.6%†
Number of reflections measured	116838†	60647	—	22754†
Number of unique reflections	49305†	16790	—	47794†

† Measured at station 7.2 of the Daresbury Synchrotron Source.

3.2. *DBL complexed with adenine*

Many, but not all, legume lectins bind adenine and adenine-related plant hormones with an affinity 3–4 orders of magnitude higher than that for monosaccharides (Maliarik *et al.*, 1987; Roberts & Goldstein, 1983; Puri & Surolia, 1994; Gegg *et al.*, 1992; Gegg & Etzler, 1994). The binding site for these compounds is distinct from the carbohydrate-binding site. Because of the binding stoichiometry of two ligands per tetramer or one ligand per dimer, these binding sites are thought to reside on a symmetric position of a subunit interface. As a consequence, this activity is probably dependent on the specific quaternary structure of the lectin, and a crude model has been suggested on the basis of the crystal structure of PHA-L (Hamelryck *et al.*, 1996).

Four crystal forms were observed. Only one crystal of the first crystal form has ever been observed. The crystal initially diffracted to 2.4 Å on a home source and belongs to space group $C222$ or $C222_1$ (a = 107.90, b = 132.34, c = 176.96 Å). Unfortunately, only 20° of data could be collected before radiation damage became severe.

The second crystal form of the DBL–adenine complex can be grown reproducibly by the hanging-drop method. The resulting crystals looked like ragged layered plates, but turned out to be single crystals when exposed to X-rays. A full data set to 2.75 Å could be collected from a single crystal. At low resolution (>4.5 Å) about three quarters of the reflections are very weak or absent, consistent with a smaller $P2_1$ cell (a = 59.19, b = 102.23, c = 80.41 Å, β = 91.52°) with only a single molecule in the asymmetric unit. At higher resolution, the larger cell (a = 99.30, b = 102.51, c = 196.93 Å, β = 101.43°) becomes apparent, suggesting a super-lattice structure (see Fig. 2). Indeed, the native Patterson contains a set of strong non-origin peaks consistent with the presence of four tetramers related by (almost) pure translational non-crystallographic symmetry.

Only one crystal of each of the third and fourth crystal forms was grown. The crystal belonging to the third crystal form diffracted anisotropically, up to a resolution of 2.4 Å in the c^* direction and 2.8 Å perpendicular to c^*, and belonged to space group $I4_122$ (a = b = 78.93, c = 260.50 Å). The data are therefore complete up to 2.8 Å and have an 18.8% completeness between 2.8 and 2.4 Å.

The fourth crystal form was grown under quite similar crystallization conditions, but with larger cell dimensions (a = 85.7, b = 113.1, c = 118.5 Å). The crystal belongs to space group $P2_12_12_1$ or $P2_12_12$.

3.3. *DBL complexed with the blood-group A + H trisaccharide GalNAc(α1-3)[Fuc(α1-2)]Gal*

For the Gal/GalNAc specificity group, structural information comes from the complexes of peanut agglutinin and *Erythrina corallodendron* lectin with lactose (Banerjee *et al.*, 1996; Shaanan *et al.*, 1991) and from soybean agglutinin with a synthetic bivalent oligosaccharide (Dessen *et al.*, 1995). With the exception of a peanut agglutinin T-antigen–disaccharide complex (Ravishankar *et al.*, 1997), no complex of a member of this group with an oligosaccharide that binds with high affinity has previouly been described. DBL further differs from these three lectins in that it only binds GalNAc and not Gal. No legume lectin–GalNAc complex is presently available. Therefore, we decided to crystallize DBL in the presence of the blood-group A + H trisaccharide GalNAc(α1-3)[Fuc(α1-2)]Gal.

Only two crystals were grown of this complex. Both were very thin plates. The first crystal survived only about 1 h of data collection, after which the PEG-Me in the surrounding mother liquor crystallized and made the diffraction images unuseable. The second crystal survived a full data collection and data analysis showed it to belong to space group $C2$ (a = 96.73, b = 108.00, c = 80.96 Å, β = 124.67°) with a dimer in its asymmetric unit. This crystal had a very high mosaicity (about 1.3°), but because of the relatively small cell, useable data could be collected to a resolution of 2.8 Å.

3.4. *DB58*

The variation in quaternary structure in the legume-lectin family is well documented but poorly understood. DB58 has 87% shared sequence identity with the seed-lectin DBL, but forms dimers instead of tetramers. A detailed comparison of

the structures of DBL and DB58 therefore may prove to be valuable for understanding quaternary structure formation in the legume-lectin family. DB58 is also unusual as it is one of the two dimeric legume lectins that are known to possess the high-affinity adenine-binding site (Gegg *et al.*, 1992; Gegg & Etzler, 1994; Puri & Surolia, 1994). Adenine binding is thought to be dependent on the quaternary structure of the lectin, with the binding site located at a subunit interface. A putative location for this binding site was suggested from affinity labelling and the structures of the tetrameric soybean agglutinin and PHA-L (Hamelryck *et al.*, 1996). It is, however, difficult to fit DB58 into this model and it is expected that the crystal structure of DB58 will shed light on this matter.

The crystals of DB58 are bipyramidal in shape and when grown without seeding remain smaller than 0.05 mm in all dimensions. Larger crystals (up to 0.3 mm in all dimensions) could be grown by the use of microseeding. These crystals do not diffract on an in-house Cu $K\alpha$ source, but showed useable diffraction at station 7.2 of the Daresbury synchrotron to at least 3.3 Å. The space group is $P2_12_12_1$ (a = 101.46, b = 130.73, c = 136.88 Å) and the asymmetric unit contains three dimers.

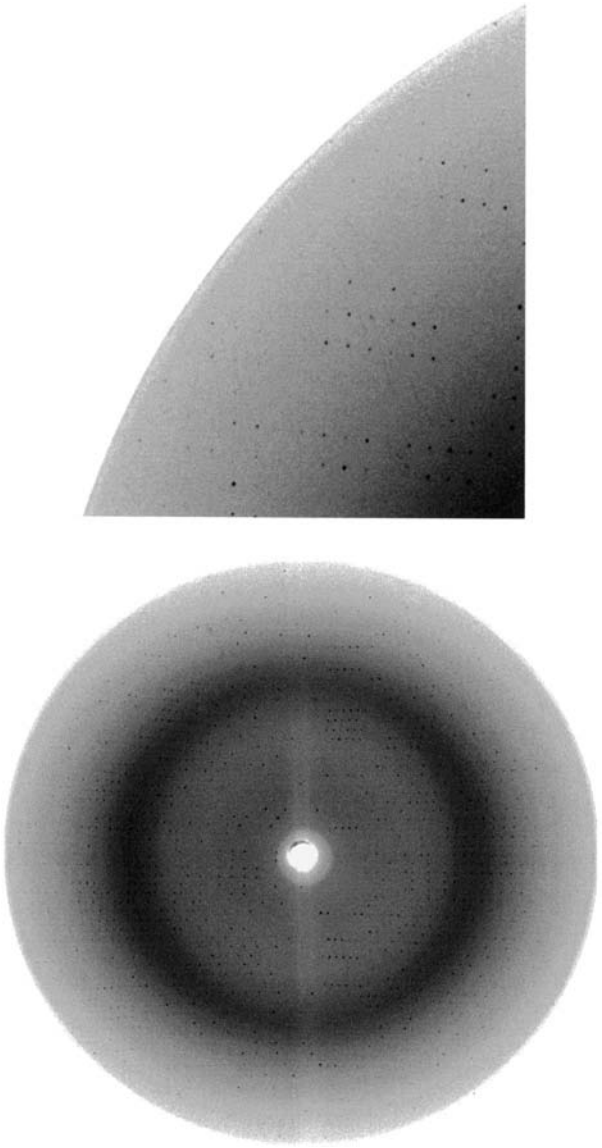

Fig. 1. Rotation image of recombinant subunit A of DBL taken at station 7.2 of the Daresbury synchrotron. An exposure time of 300 s was used together with a rotation of 0.7°. The crystal-to-detector distance was 235 mm and the wavelength 1.488 Å. Diffraction spots can be seen up to the edge of the detector, which corresponds to a resolution of 2.7 Å.

4. Conclusions

In the past decade, the crystal structures of 13 different legume lectins, many in complex with a bound carbohydrate, have been determined (for a recent review, see Loris *et al.*, 1997). Despite this wealth of information, our knowledge of the structure–function relationship of this interesting group of proteins remains far from complete, and several important questions remain unsolved. What is the molecular basis of the variety of quaternary structures that are observed? How do legume lectins bind adenine and adenine-related plant hormones? What is the molecular basis of oligosaccharide specificity outside the Man/Glc specificity group? The DBL and DB58 crystals described in this paper will provide some important clues to answering these questions. The high shared sequence identity between both lectins will help in interpreting the molecular basis of the difference in quaternary structure. The complex with the blood-group A trisaccharide will provide a structure of a non-mannose specific lectin in complex with a relevant oligosaccharide. The interaction of DBL with

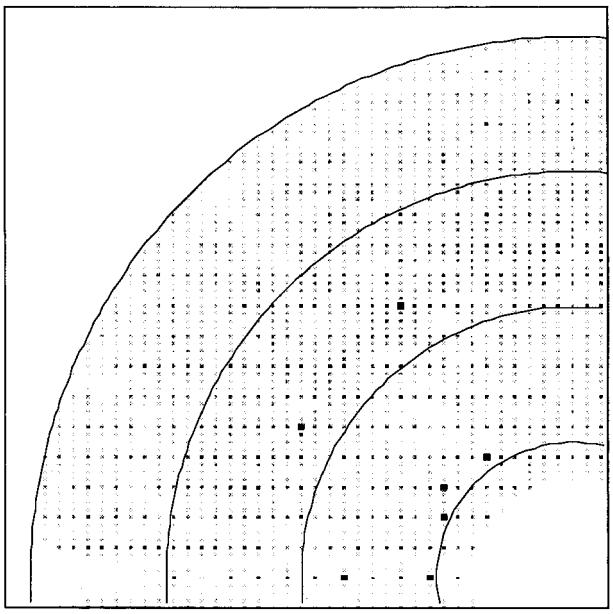

Fig. 2. Pseudo-precession image (0kl layer) generated with the *CCP*4 (Collaborative Computational Project, Number 4, 1994) program *HKLVIEW* from the 2.7 Å data set of DBL in complex with adenine. The circles correspond to 2.7, 3.6, 5.4 and 10.8 Å resolution. The presence of the superlattice is clearly visible: strong spots are associated with (h, k, 4n + k). At low resolution (<5 Å) the weak spots are essentially absent.

Table 2. *Crystal forms of complexed DBL*

	DBL subunit I + adenine form I	DBL subunit I + adenine form II	DBL subunit I + trisaccharide	DBL subunit I + II + adenine	DBL subunit I + II + adenine
Protein solution	5 mg ml^{-1} protein, 10 m*M* adenine	3.75 mg ml^{-1} protein saturated with adenine	2.8 mg ml^{-1} protein, 10 m*M* trisaccharide	5 mg ml^{-1} protein, 5 m*M* adenine	5 mg ml^{-1} protein, 5 m*M* adenine
Precipitant solution	100 m*M* sodium cacodylate pH 6.5, 20 m*M* MgAc, 8%(*w*/*v*) PEG 8000	100 m*M* sodium cacodylate pH 6.5, 20%(*w*/*v*) PEG-Me 5000, 0.1 *M* ammonium sulfate	100 m*M* sodium cacodylate pH 6.6, 15%(*w*/*v*) PEG-Me 5000	100 m*M* HEPES pH 7.5, 0.2 *M* $CaCl_2$, 28%(*w*/*v*) PEG 400	100 m*M* HEPES pH 7.5, 0.2 *M* $MgCl_2$, 30%(*w*/*v*) PEG 400
Space group	*C*222 or $C222_1$	$P2_1$	*C*2	$P2_12_12_1$ or $P2_12_12$	$I4_122$
Unit cell					
a (Å)	107.90	99.30	96.73	85.7	78.93
b (Å)	132.34	102.51	108.00	113.1	78.93
c (Å)	176.96	196.93	80.96	118.5	260.50
β (°)			101.43	124.67	
Contents of unit cell	1 tetramer	4 tetramers	1/2 tetramer	1 tetramer	1/4 tetramer
Resolution (Å)	2.4 (initially)	2.75	2.8	2.6	2.8–2.4†
R_{merge}	—	0.105	0.111	0.152	0.136
Completeness	—	98.2%	85.2%	97.8%	81.3%
Number of reflections measured	—	494178	39403	156181	148337
Number of unique reflections	—	105360	14300	30823	13545

† The crystal diffracted anisotropically up to 2.4 Å resolution in the *c** direction, and up to 2.8 Å resolution perpendicular to *c**.

GalNAc(α1-3)[Fuc(α1-2)]Gal has recently been studied by NMR and molecular modelling, which provide information complementary to that obtained by X-ray crystallography. Finally, we also crystallized the first complex of a legume lectin and adenine. This interaction is potentially relevant to the *in vivo* function of legume lectins, but is also only poorly characterized.

This work was supported by the Vlaams Interuniversitair Instituut voor Biotechnologie. RL and JB are postdoctoral fellows of the Fonds voor Wetenschappelijk Onderzoek (FWO). TWH received financial support from the Vlaams Instituut voor de Bevordering van het Wetenschappelijk-technologisch Onderzoek in de Industrie (IWT).

References

Banerjee, R., Das, K., Ravishankar, R., Suguna, K. & Surolia, A. (1996). *J. Mol. Biol.* **259**, 281–296.
Chao, Q., Casalongue, C., Quinn, J. M. & Etzler, M. E. (1994). *Arch. Biochem. Biophys.* **313**, 346–350.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dessen, A., Gupta, D., Sabesan, S., Brewer, F. & Sacchettini, J. C. (1995). *Biochemistry*, **34**, 4933–4942.
Etzler, M. E. (1994). *Glycoconjugate J.* **11**, 395–399.
Etzler, M. E. (1996). *Lectins, Biology–Biochemistry–Clinical Biochemistry*, edited by E. Van Driessche, P. Rougé, S. Beeckmans & T. C. Bøg-Hansen, pp. 3–9. Hellerup, Denmark: Textop.
Etzler, M. E., Gupta, S. & Borrebaeck, C. (1981). *J. Biol. Chem.* **256**, 2367–2370.
Etzler, M. E. & Kabat, E. A. (1970). *Biochemistry*, **9**, 869–877.
Gegg, C. V. & Etzler, M. E. (1994). *J. Biol. Chem.* **269**, 5687–5692.
Gegg, C. V., Roberts, D. D., Segel, I. H. & Etzler, M. E. (1992). *Biochemistry*, **31**, 6938–6942.
Hamelryck, T. W., Dao-Thi, M., Poortmans, F., Chrispeels, M. J., Wyns, L. & Loris, R. (1996). *J. Biol. Chem.* **271**, 20479–20485.
Harada, J. J., Spadoro-Tank, J., Maxwell, J. C., Schnell, D. J. & Etzler, M. E. (1990). *J. Biol. Chem.* **265**, 4997–5002.
Loris, R., Bouckaert, J., Hamelryck, T. & Wyns, L. (1998). *Biochim. Biophys. Acta*, **1383**, 9–36.
Maliarik, M. J., Roberts, D. D. & Goldstein, I. J. (1987). *Arch. Biochem. Biophys.* **255**, 194–200.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Puri, K. D. & Surolia, A. (1994). *J. Biol. Chem.* **269**, 30917–30926.
Ravishankar, R., Ravindran, M., Suguna, K., Surolia, A. & Vijayan, M. (1997). *Curr. Sci.* **72**, 855–861.
Roberts, D. D. & Goldstein, I. J. (1983). *J. Biol. Chem.* **258**, 13820–13824.
Schnell, D. J. & Etzler, M. E. (1988). *J. Biol. Chem.* **263**, 14648–14653.
Shaanan, B., Lis, H. & Sharon, N. (1991). *Science*, **254**, 862–865.
Strecker, G., Wieruszcski, J. M., Michalski, J. C. & Montreuil, J. (1989). *Glycoconjugate J.* **6**, 271–284.
Young, M. N., Watson, D. C., Yaguchi, M., Adar, R., Arango, R., Rodriguez-Arango, E., Sharon, N., Blay, P. K. S. & Thibault, P. (1995). *J. Biol. Chem.* **270**, 2563–2570.