Volume 268, number 1, 281–286 FEBS 08675 July 1990

The primary structure of the *Lotus tetragonolobus* seed lectin

Yukiko Konami, Kazuo Yamamoto and Toshiaki Osawa

Division of Chemical Toxicology and Immunochemistry, Faculty of Pharmaceutical Sciences, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Received 14 May 1990

The complete amino acid sequence of the *Lotus tetragonolobus* lectin was determined by a protein sequencer after digestion with endoproteinases of Lys-C and Asp-N, and compared with those of other leguminous plant lectins.

Lotus tetragonolobus anti-H(O) lectin; Amino acid sequence

INTRODUCTION

Extracts of many seeds, mainly of leguminous plants, have been found to contain soluble lectins, which have widely been used as powerful tools in the research on cell-surface complex-carbohydrates owing to their unique ability to bind specific sugars or sugar-containing macromolecules. Among them several anti-H(O) lectins have already been isolated and purified. In this study, we determined the complete amino acid sequence of the anti-H(O) *tetragonolobus* lectin (LTA) by a protein sequencer. After digestion with two kinds of endoproteinases, Lys-C and Asp-N, the resulting peptides were purified by reversed-phase high-performance liquid chromatography (HPLC). These segments of the continuous sequence accounting for the entire protein, were aligned through the comparison of sequences with those of several homologous leguminous lectins to give a final structure. Extensive homology was found among the primary structures of this *Lotus* lectin and the other seven lectins previously reported [1–7].

MATERIALS AND METHODS

The seeds of *Lotus tetragonolobus* were obtained from F.W. Schumacher Co., Sandwich, MA, USA. TSK-Gel G3000 SW for gel filtration was obtained from TOSOH Co., C_4 and C_{18} μBondaspheres (100 A) for reversed-phase chromatographies were obtained from Waters. Endoproteinases Lys-C (*Lysobacter enzymogenes*) and Asp-N (*Pseudomonas fragi*) were purchased from Boehringer Mannheim. Finely powdered *Lotus tetragonolobus* seeds was extracted with 0.9% NaCl. To the yellow clear supernatant obtained by centrifugation, solid $(NH_4)_2SO_4$ was added to give 0.7 saturation. The precipitate formed (crude lectin) was dialyzed against PBS (10 mM sodium phosphate buffer (pH 7.3) containing 0.15 M NaCl) and applied to a fucose-Sepharose 4B affinity column equilibrated with PBS. The column was washed with PBS until the absorbance at 280 nm of the eluate was approximately below 0.01 and then eluted with 0.1 M fucose in PBS. Fractions of 2 ml were collected at a flow rate of 0.4 ml/min. The SDS-polyacrylamide gel electrophoresis was carried out according to the method of Laemmli [8] with 10% acrylamide gel.

The purified lectin thus obtained was further applied to a reversed-phase HPLC on a column of C_4 by using a linear gradient (0 – 100%) of 2-propanol/acetonitrile (7:3) in distilled water containing 0.1% trifluoracetic acid (TFA) in 60 min at a flow rate of 1 ml/min. The peptide fragments obtained after digestion of the lectin with endoproteinases Lys-C and Asp-N were separated with a reversed-phase HPLC on columns of C_{18} and C_4, respectively by using a linear gradient (0 – 60%) of 2-propanol/acetonitrile (7:3) in distilled water containing 0.1% TFA in 60 min at a flow rate of 1 ml/min. Sequence analyses of the purified peptides were performed on a PSQ-1 gas-phase protein sequencer (Shimadzu, Kyoto, Japan), and a 6600 Pro-Sequencer solid-phase protein sequencer (Milligen/Biosearch, Burlington, MA, USA).

RESULTS

3.1. *Purification of lectin*

The affinity purified lectin was further purified by gel filtration by use of HPLC on a column of TSK-Gel G3000 SW (0.75 × 60 cm) to obtain a single A_{215} peak corresponding to a molecular mass of 51000. The minimum hemagglutinating doses of the purified lectin was 3.3 μg/ml against neuraminidase-treated human O erythrocytes and 26.4 μg/ml and 105.6 μg/ml against

Correspondence address: T. Osawa, Division of Chemical Toxicology and Immunochemistry, Faculty of Pharmaceutical Sciences, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan

Abbreviations: Con A, concanavalin A; HPLC, high-performance liquid chromatography; LL, *Lens culinaris* (lentil) lectin; LTA, *Lotus tetragonolobus* lectin; PBS, 10 mM sodium phosphate buffer (pH 7.3) containing 0.15 M NaCl; PHA, *Phaseolus vulgaris* lectin; SBA, *Glycine max* lectin; SL, *Onobrychis vicifolia* (sainfoin) lectin

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50 © 1990 Federation of European Biochemical Societies

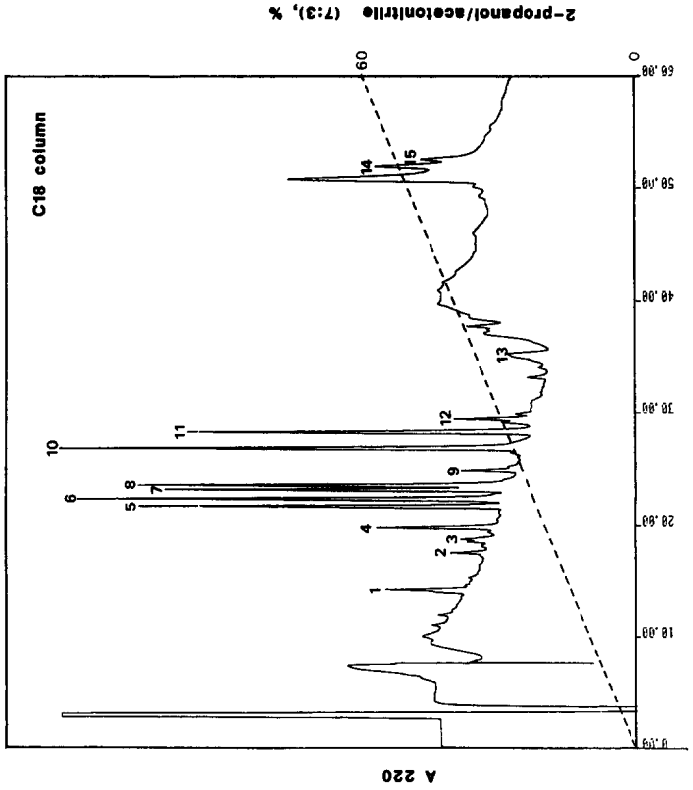

Fig. 1. Reversed-phase HPLC of Lys-C digest of purified *L. tetragonolobus* lectin on a column of C_{18}.

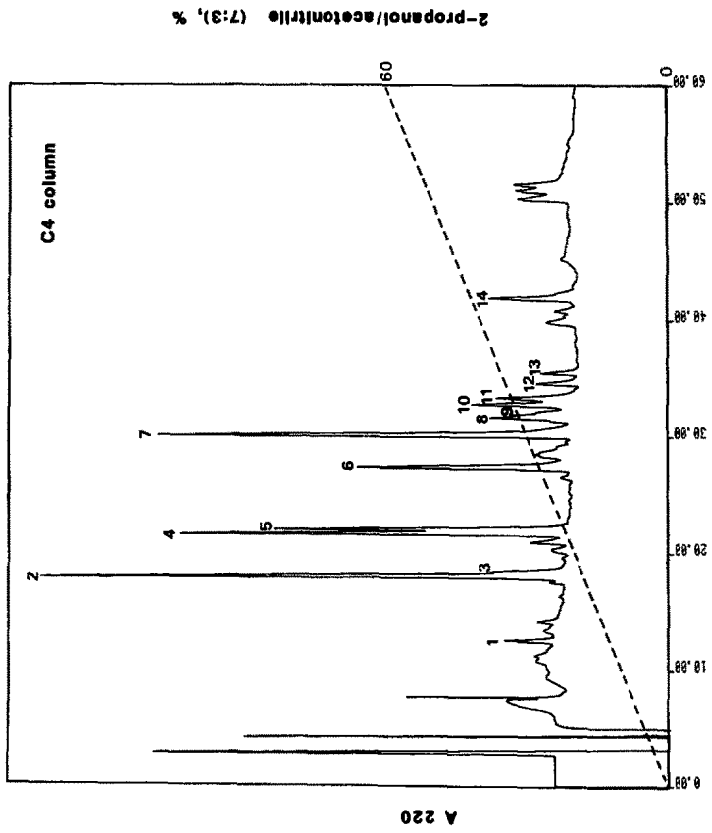

Fig. 2. Reversed-phase HPLC of Asp-N digest of purified *L. tetragonolobus* lectin on a colum of C_4.

neuraminidase-treated human B and A erythrocytes, respectively. This *Lotus* lectin was most inhibited by L-fucose (minimum amounts (mM) completely inhibiting 4 hemagglutinating doses: 0.001 mM). The SDS polyacrylamide gel electrophoresis of this purified lectin showed a single protein band corresponding to a subunit molecular mass of 26000.

3.2. *Determination of the primary structure*

Purification and sequencing of the peptides which were obtained by digestion with endoproteinases, Lys-C (Fig. 1) and Asp-N (Fig. 2), provided enough overlapping sequences to obtain the complete amino acid sequence of LTA (Fig. 3). The structure of this lectin contains 240 amino acid residues. The molecular mass of this lectin calculated on the basis of the sequence is 26273.17. The molecular masses approximated by gel filtration and SDS-polyacrylamide gel electrophoresis are in good agreement with this finding. Complete sequence of LTA was compared to those of concanavalin A (Con A) [1,2], *Glycine max* lectin (SBA) [2], *Vicia faba* lectin (favin) [2,3], *Lens culinaris* lectin (lentil lectin; LL) [4], *Pisum sativum* lectin (pea) [5], *Onobrychis vicifolia* lectin (sainfoin lectin; SL) [6] and *Phaseolus vulgaris* lectin (PHA) [7] (Fig. 4).

Overall homology between LTA and the other sequenced proteins, expressed as percentage identities/positions compared is 41.7%, 42.5%, 39.6%, 36.3%, 40.8%, 37.9% and 30.8% for Con A, SBA, favin, lentil lectin, pea, sainfoin lectin and PHA, respectively.

4. DISCUSSION

In the previous papers [9,10], it was reported that there exist three fucose-binding proteins (A, B, C) in the seeds of *L. tetragonolobus*. Since the lectin which we purified has a molecular mass of 56000, it might corres-

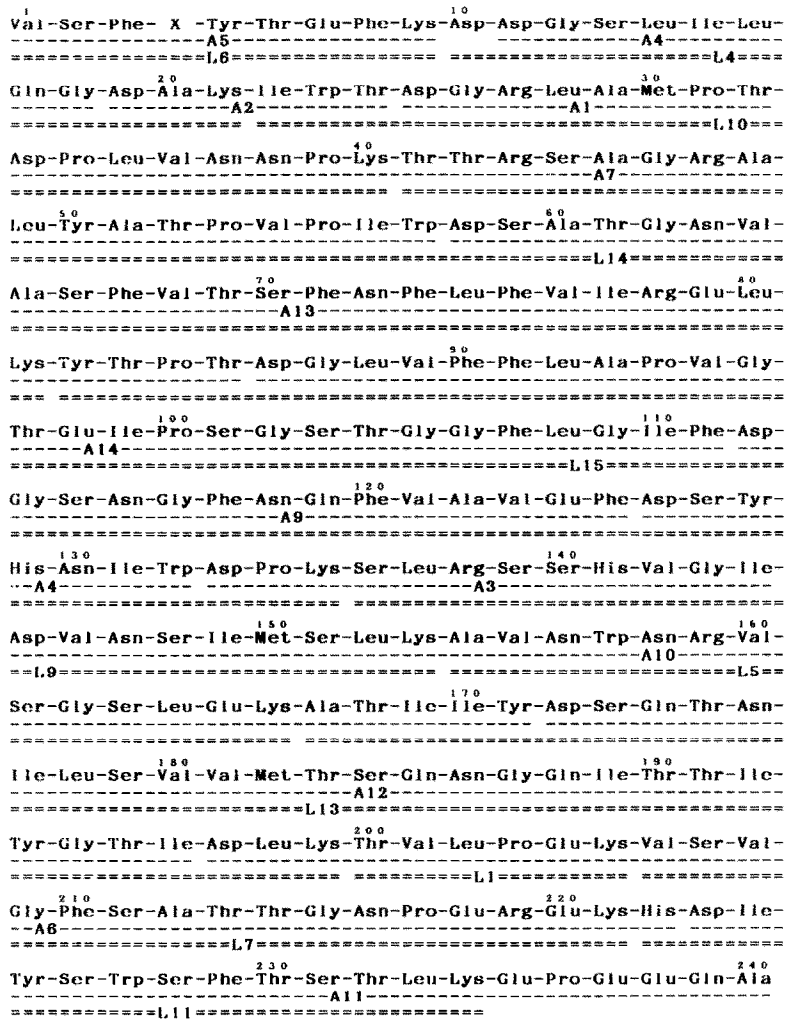

Fig. 3. Complete amino acid sequence of *L. tetragonolobus* lectin. Primary structure in terms of the peptides obtained after digestion with Lys-C (L) and Asp-N (A). Residues were identified by automated sequence analysis of the Asp-N fragments (------) or that of the Lys-C fragments (= = = = = =).

pond to the compound B with a molecular mass of 58000. However, in our case no other isolectins could be isolated even if we tried to use the same types of ion-exchange chromatography and isoelectric focusing. We assume that this may be due to the difference in *Lotus* seeds.0The comparison of the amino acid sequence of LTA to those of several lectins which have been determined so far was shown in Fig. 4. The homology among these lectins is so close that it is likely that the three-dimensional structures of these lectins will be found to be similar.0When comparing the sequence of LTA to that of Con A, the only one leguminous plant lectin whose three-dimensional structure has been completely elucidated, a striking homology exists in residues 1–240 of LTA with the sequences 123–237 and 1–122 of Con A. For example, of 51 residues in the central β-pleated sheet of Con A, 31 residues are identical in LTA, and 25 of 56 residues in the back β-sheet of Con A are conserved in LTA. Moreover, all of the direct metal ligands in Con A (Glu^{-8}, Asp^{-10}, Tyr^{-12}, Asn^{-14}, Asp^{-19}, His^{-24}, His^{-24} and Ser^{-34}) are conserved in LTA sequence. As for the residues related to the hydrophobic cavity, all of the residues in Con A (Tyr^{-54}, Leu^{-81}, Leu^{-85}, Val^{-89}, Phe^{-111}, Val^{-179}, Phe^{-191} and Phe^{-212}) are conserved in LTA. The conservation of the hydrophobic cavity supports its essential role in the function of plant lectins through evolution [11].0The glycosylation site was assumed to be at position 4 of LTA and it was located in the unique sequence -Phe-X-Tyr-Thr-. The carbohydrate attachment position, among the homologous lectins which contain carbohydrate, is not conserved. Favin is glycosylated at Asn-169 [54], the soybean agglutinin at Asn-75 [12] and sainfoin lectin at Asn-118 [6].0The sequence of *Lotus* lectin reported here, along with the other homologous structures, supports the conclusion that the leguminous lectins are a family of related proteins whose structures have been evolutionarily conserved.

REFERENCES

[1] Cunningham, B.A., Wang, J.L., Waxdal, M.J. and Edelman, G.H. (1975) J. Biol. Chem., 250, 1502–1512.

[2] Hemperly, J.J. and Cunningham, B.A. (1983) Trends Biochem, Sci. 8, 100–102.

[3] Hopp. T.P., Hemperly, J.J. and Cunningham, B.A. (1982) J. Biol. Chem. 257, 4473–4483.

[4] Foriers, A., Lebrun, E., Van Rappenbusch, R., De Neve, R. and Strosberg, A.D. (1981) J. Biol. Chem. 256, 5550–5560.

[5] Higgins, T.J.V., Chandler, P.M., Zurawski, G., Button, S.C. and Spencer, D. (1983) J. Biol. Chem. 258, 9544–9549.

[6] Kouchalakos, R.N., Bates, O.J., Bradshaw, R.A. and Hapner, K.D. (1984) Biochemistry 23, 1824–1830.

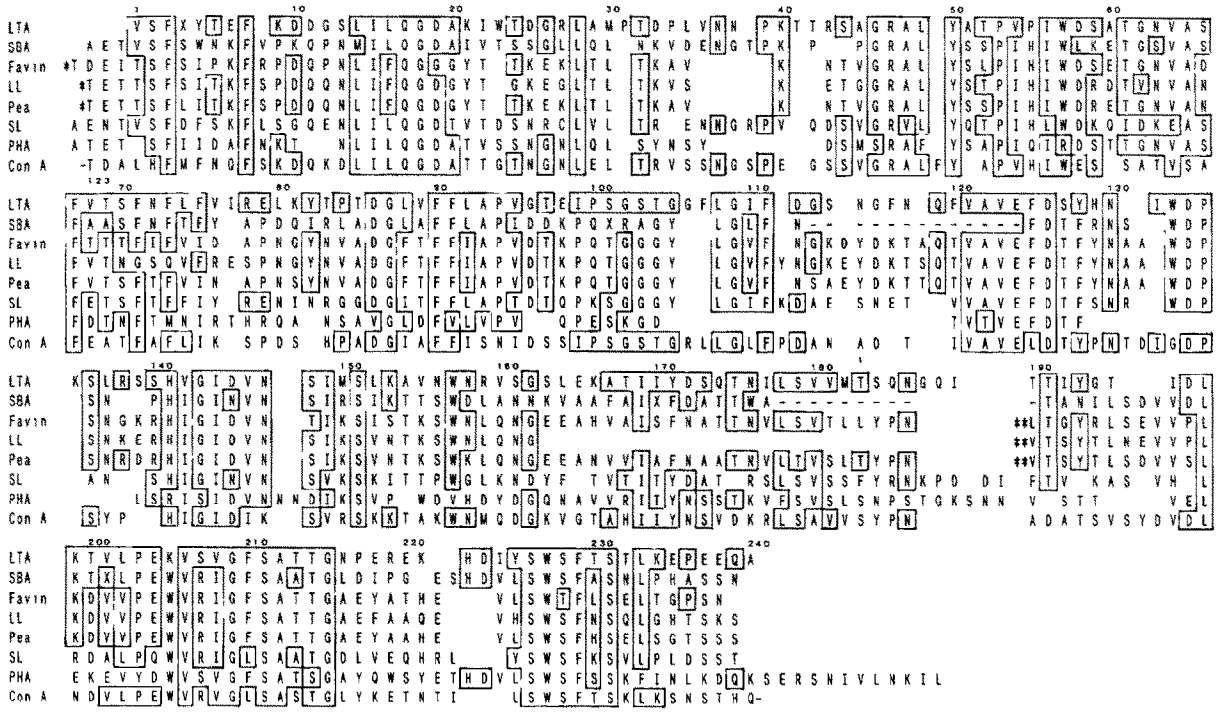

Fig. 4. Comparison of complete amino acid sequence of *L. tetragonolobus* lectin with those of Con A, *Glycine max* Lectin (SBA), *Vicia faba* Lectin (favin), *Lens culinaris* Lectin (LL), *Pisum sativum* Lectin (pea), *Onobrychis vicifolia* Lectin (SL) and *Phaseolus vulgaris* Lectin (PHA). Gaps are inserted in all sequences to maximize homology. Identical residues are enclosed in boxes. * and ** indicate the amino termini of the β and α chains of favin, LL and pea, respectively. (LL and pea α subunit at position 189, LL and pea β subunit at -3, favin α at 189, favin β at -4. The amino of Con A, SBA, SL and PHA (a single chain) occur at 115, 1, -3, -4 and -4, respectively. Dashes indicate regions of SBA that are as yet undetermined.

[7] Hoffman, L.M., Ma, Y. and Barker, R.F. (1982) Nucleic Acids Res. 10, 7819–7828.
[8] Laemmli, U.K. (1970) Nature 227, 680–685.
[9] Kalb, A.J. (1968) Biochim. Biophys. Acta 168, 532–536.
[10] Blumberg, S., Hildesherim,. J., Yariv, J. and Wilson, K.J. (1972) Biochim. Biophys. Acta 264, 171–176.
[11] Edelman, G.H. and Wang, J.L. (1978) J. Biol. Chem. 253, 3016–3022.
[12] Becker, J.W., Hemperly, J.J. and Cunningham, B.A. (1983) in: Chemical Taxonomy, Molecular Biology and Function of Plant Lectins (Goldstein, J.J. and Etzlar, M.E., Eds.) pp. 31–45, Alan R. Liss, New York.