

The primary structure of the *Laburnum alpinum* seed lectin

Yukiko Konami, Kazuo Yamamoto, Satoshi Toyoshima and Toshiaki Osawa

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

Received 16 April 1991; revised version received 14 May 1991

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

Laburnum alpinum anti-H(O) lectin; Amino acid sequence

1. INTRODUCTION

The anti-H(O) hemagglutinating activity in extracts from the seeds of *Laburnum alpinum* was first discovered by Renkonen [1] and further confirmed by Morgan and Watkins [2,3]. The hemagglutination-inhibition studies on crude extracts from the seeds revealed a specificity towards di-*N*-acetylchitobiosyl residues [3–5], and human A, H and neuraminidase-treated human Le^a blood-group substances were also good inhibitors of the *Laburnum* lectin [5]. In a previous paper [6] we have shown the specific purification and characterization of two kinds of *Laburnum alpinum* lectin, a *Cytisus*-type di-*N*-acetylchitobiose-binding lectin (LAA-I) and a new type lectin which is inhibited by lactose or galactose (LAA-II). We have also determined the primary structures of the *Lotus tetragonolobus* anti-H(O) lectin (LTA) [7] and two types of the *Ulex europeus* anti-H(O) lectins (UEA-I and II) [8], and compared them with those of several lectins. Extensive homologies were found among them. In this study, we determined the complete amino-acid sequence of the anti-H(O) *Laburnum alpinum* lectin I (LAA-I) by using a protein sequencer. After digestion of the lectin with two kinds of endoproteinases, Lys-C and Asp-N, the resulting peptides were purified by reversed-phase high-performance liquid chroma-

tography (HPLC) and subjected to the sequence analysis. The complete primary structure of this lectin was compared with those of 14 lectins already determined, including those of LTA [7] and UEA-I and II [8]. Among these lectins, extensive homologies, especially between LAA-I and UEA-II, were found.

2. MATERIALS AND METHODS

The seeds of *Laburnum alpinum* were obtained from F.W. Schumacher Co., Sandwich, MA. C₄ and C₁₈ μ Bondaspheres (100 Å) for reversed-phase chromatography were obtained from Waters (Burlington, MA). Endoproteinases Lys-C (*Lysobacter enzymogenes*) and Asp-N (*Pseudomonas fragi*) were purchased from Boehringer (Mannheim, Germany).

The *Laburnum alpinum* lectin I was isolated and purified according to the methods reported previously [6]. This affinity-purified lectin was further purified by reversed-phase HPLC on a column of C₄ using a linear gradient (0–100%) of 2-propanol/acetonitrile (7:3) in distilled water containing 0.1% trifluoroacetic acid (TFA) in 60 min at a flow rate of 1 ml/min. Then the lectin thus purified (0.5 mg in 150 μ l of 50 mM phosphate buffer, pH 8.0) was digested with 5 μ g of Lys-C or with 2 μ g of Asp-N for 18 h at 37°C. The peptide fragments obtained were separated by reversed-phase HPLC on a column of C₁₈ 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. Elution profiles were monitored at 220 nm. The peptide fragments were collected manually.

Sequence analyses of the intact proteins and of the purified peptides were performed on a 6600 ProSequencer solid-phase protein sequencer (MilliGen/Biosearch, Burlington, MA, USA) and a PSQ-I gas-phase protein sequencer (Shimadzu, Kyoto, Japan).

Hydropathy plots were generated by the method of Kyte and Doolittle [9] and secondary structure predictions were carried out by using the DNA Strider (Centre d'Etudes Nucleaires de Saclay, Gif-sur-Yvette Cedex, France).

3. RESULTS

3.1. Determination of the primary structure

Purification and sequencing of the peptides, obtained after digestion of LAA-I with endoproteinases Lys-C, Fig. 1 and ASP-N, Fig. 2 provided enough overlapping

Correspondence address: T. Osawa, Division of Chemical Toxicology and Immunochemistry, Faculty of Pharmaceutical Sciences, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan. Fax: (81) (3) 8159344.

Abbreviations: Con A, concanavalin A; HPLC, high-performance liquid chromatography; LCL, *Lens culinaris* (lentil) lectin; LTA, *Lotus tetragonolobus* lectin; PHA, *Phaseolus vulgaris* lectin; SBA, *Glycine max* lectin; SL, *Onobrychis vicifolia* (sainfoin) lectin; UEA, *Ulex europeus* lectin; DBA, *Dolichos biflorus* lectin; LOL, *Lathyrus ochrus* lectin; LBL, *Phaseolus limensis* lectin; ECorL, *Erythrina corallodendron* lectin; TFA, trifluoroacetic acid

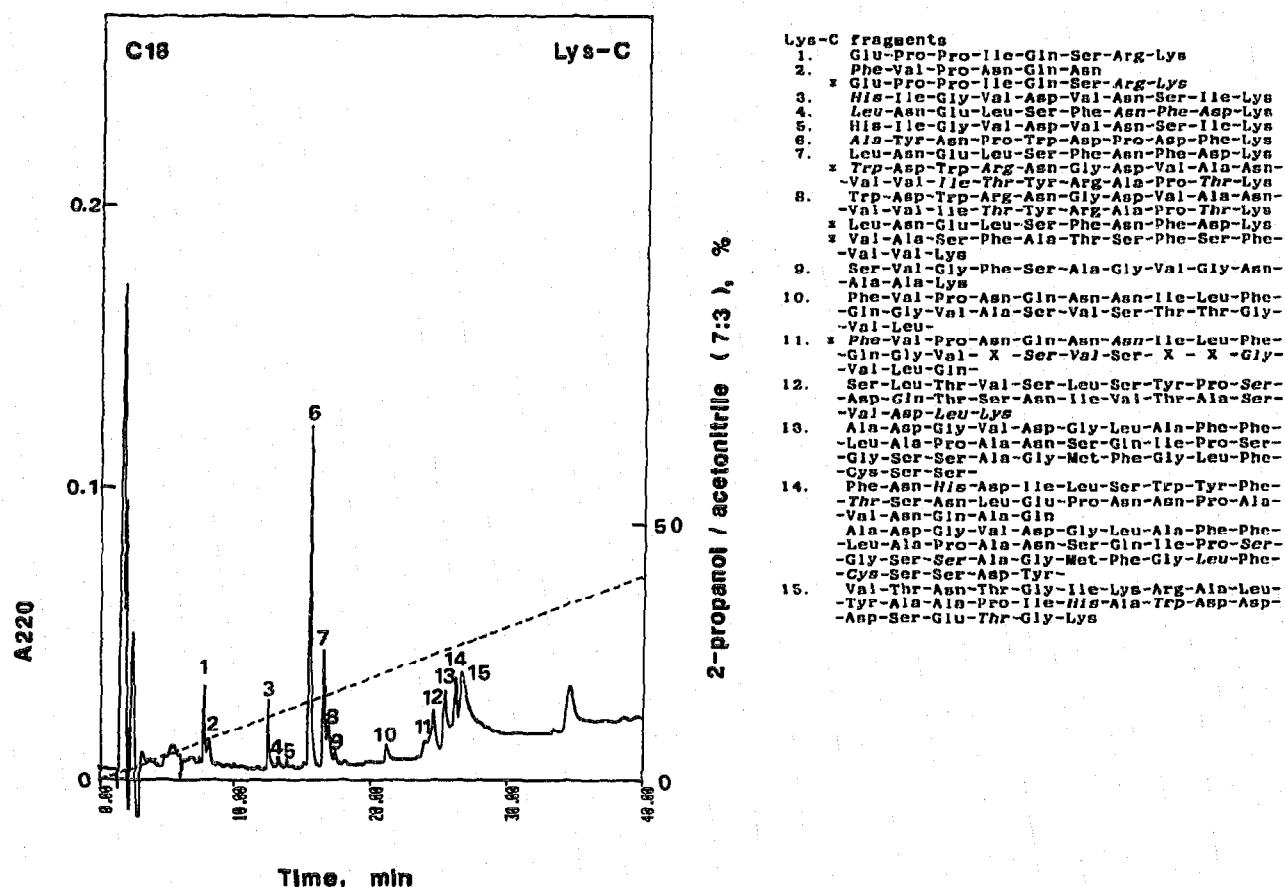


Fig. 1. Reversed-phase HPLC of Lys-C digest of purified LAA-I on a column of C₁₈. The residues which were difficult to identify are written in italics. A minor peptide in a fragment is marked with an asterisk.

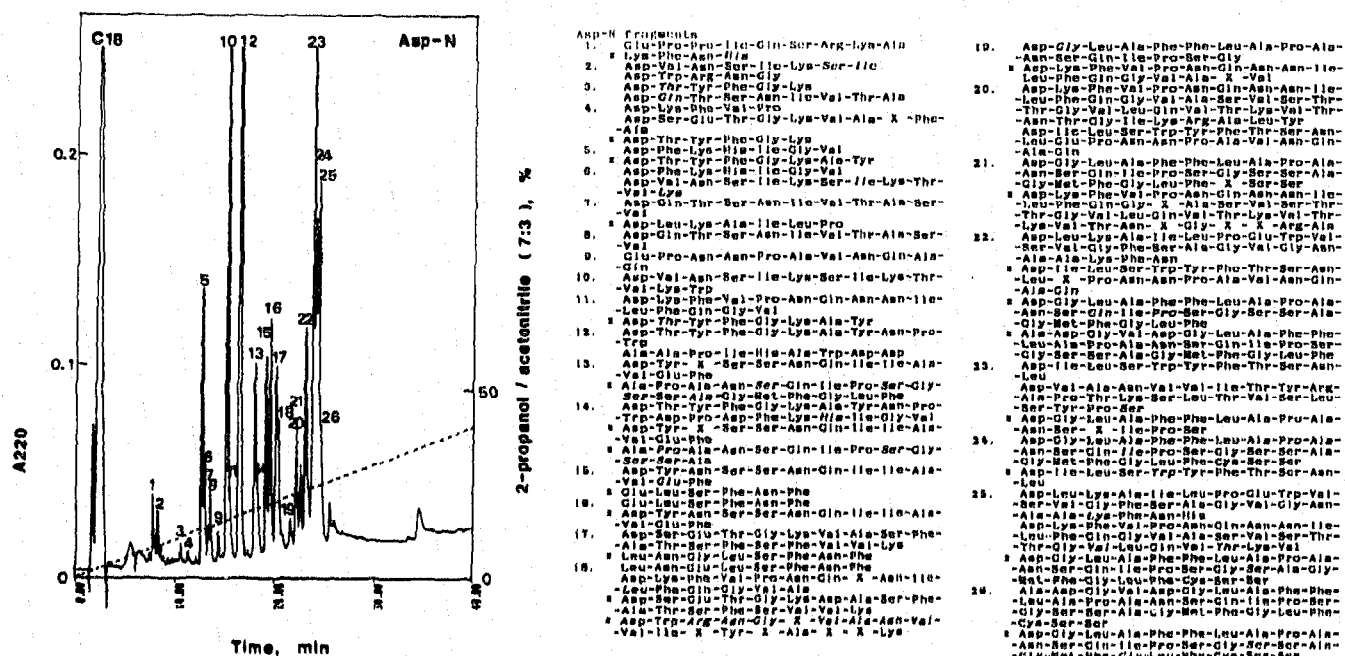


Fig. 2. Reversed-phase HPLC of Asp-N digest of purified LAA-I on a column of C₁₈. The residues which were difficult to identify are written in italics. A minor peptide in a fragment is marked with an asterisk.

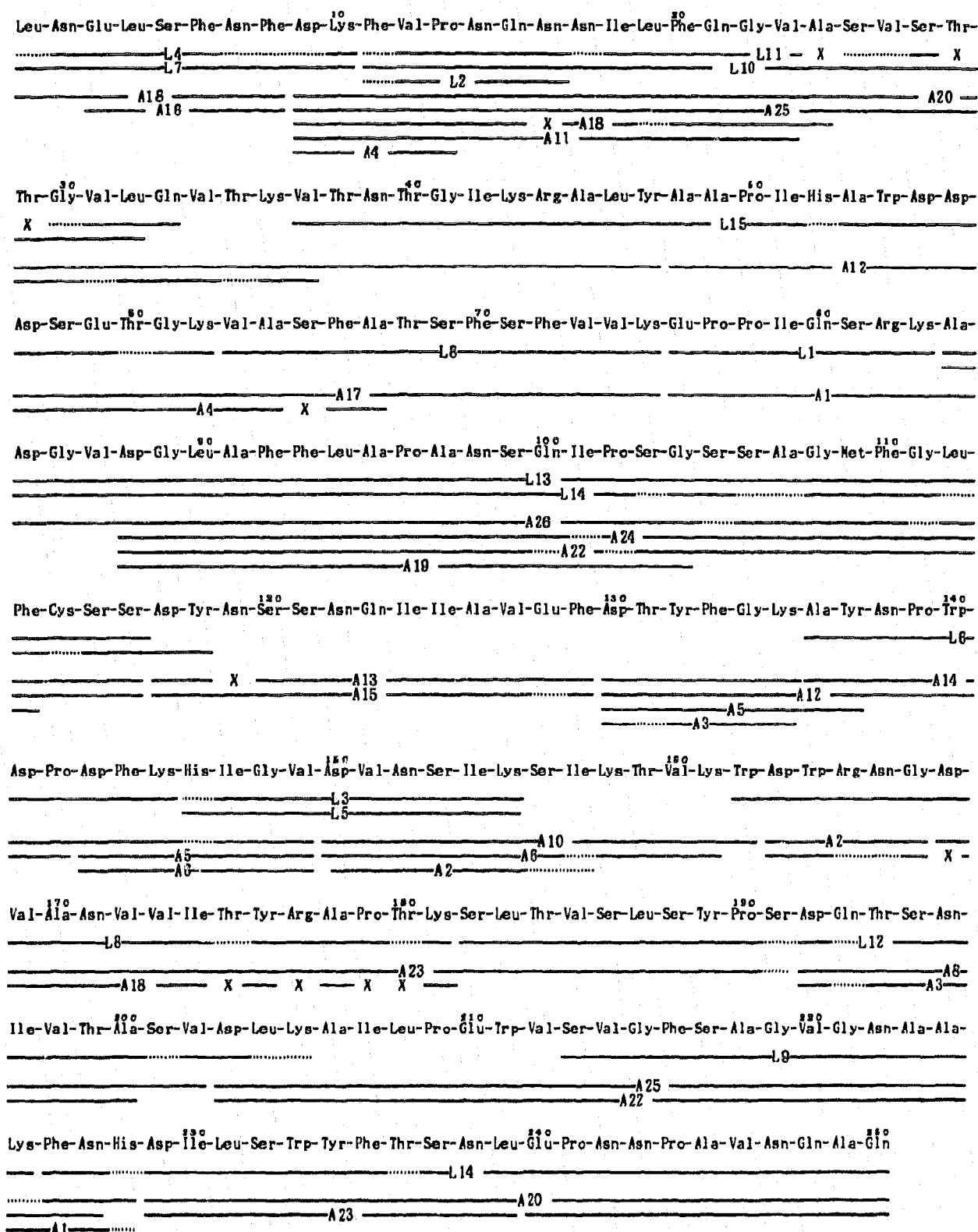


Fig. 3. Complete amino-acid sequence of LAA-I. Primary structure in terms of the peptides obtained after digestion with Lys-C (L) and Asp-N (A). Residues were identified by automated sequence analysis of the Lys-C fragments (---L---) or of the Asp-N fragments (---A---). Dashes show residues not determined clearly at this position.

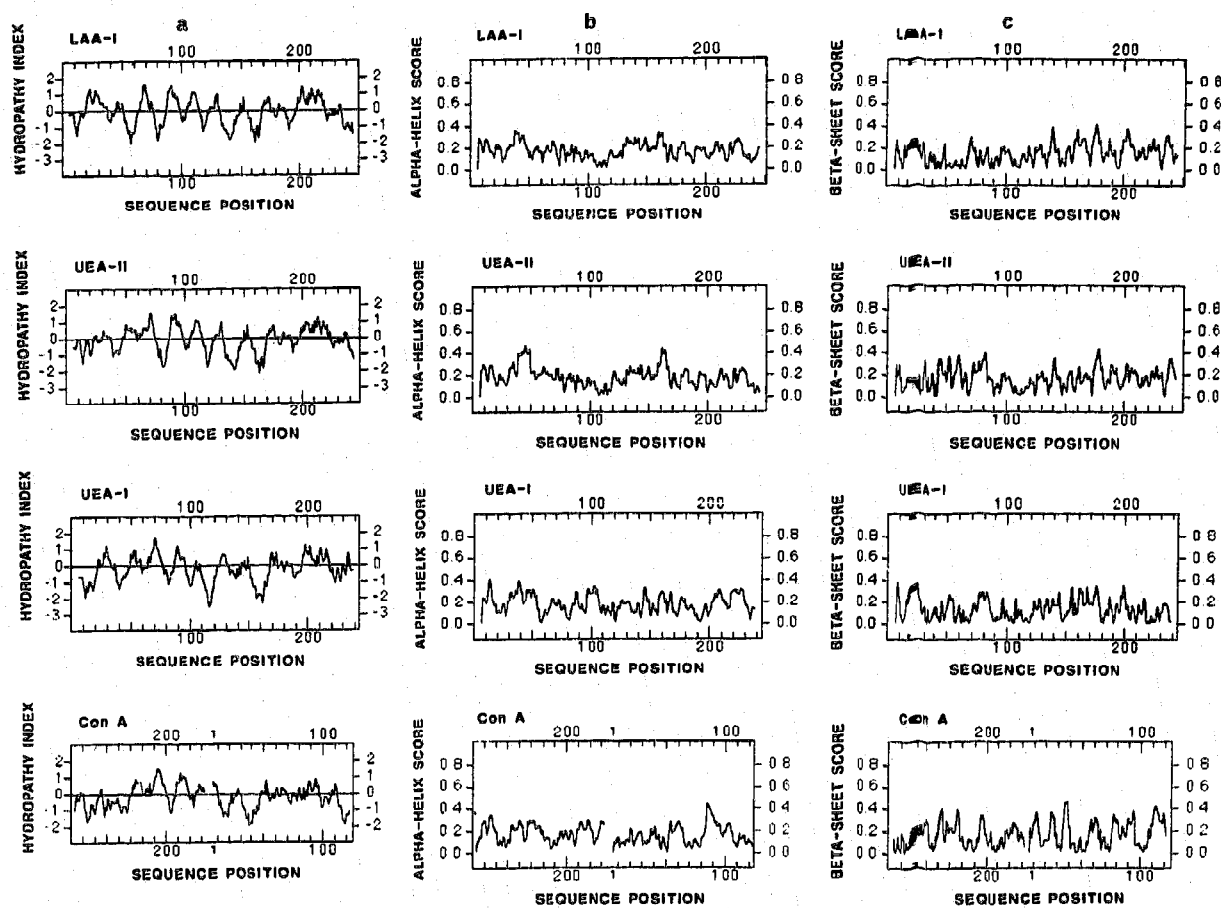


Fig. 5. Comparison of the hydropathy plots and predicted secondary structure analyses of LAA-I, UEA-I, UEA-II and Con A. The hydropathic scale is given on the left according to Kyte and Doolittle [9] with (–) indicating relative hydrophilicity and (+) indicating relative hydrophobicity. The β -sheet and α -helix scores are given on the left. The sequence of Con A is presented in a permuted arrangement to maximize homology. a, hydropathy profiles; b, α -helix; c, β -sheet.

sequences to obtain the complete amino-acid sequences of LAA-I, Fig. 3. The structure of this lectin contains 250 amino-acid residues. The molecular weight of this lectin calculated on the basis of the sequence is 27 150.83. The apparent molecular weight determined by SDS polyacrylamide gel electrophoresis described previously (32 000) [6] is in good agreement with this finding. As shown in Fig. 4, complete sequence of this *Laburnum alpinum* lectin I was compared with those of *Ulex europaeus* lectins I and II (UEA-I and II) [8], *Lotus tetragonolobus* lectin (LTA) [7], Con A [10], *Glycine max* (soy bean) lectin (SBA) [11], *Dolichos biflorus* anti-A lectin (DBA) [12], *Vicia faba* lectin (Favin) [13], *Lens culinaris* (lentil) lectin (LCL) [14], *Pisum sativum* lectin (pea) [15], *Onobrychis vicifolia* (sainfoin) lectin (SL) [16], *Phaseolus vulgaris* lectin (PHA) [17], *Lathyrus ochrus* lectin (LOL) [18], *Erythrina corallodendron* lectin (ECoRL) [19] and *Phaseolus limensis* lectin (LBL) [20]. Overall identity values between LAA-I and the other sequenced proteins, UEA-I and II, LTA, Con A, SBA, DBA, Favin, LCL, pea, SL, PHA,

LOL, ECoRL and LBL are 51.4%, 82.7%, 42.2%, 39.9%, 48.2%, 45.4%, 40.0%, 34.8%, 41.0%, 37.6%, 45.2%, 42.7% and 37.7%, respectively. Fig. 5 shows the hydropathy plots and predicted secondary structures for LAA-I, UEA-I, UEA-II and Con A. The sequences are presented as linear hydropathy plots to illustrate the marked similarities in the distributions of hydrophilic and hydrophobic regions not evident from a comparison of their amino acid sequences.

4. DISCUSSION

The amino-acid sequence of the *Laburnum alpinum* lectin I (LAA-I) was compared with those of several lectins which have been determined so far, Fig. 4. This *Laburnum* lectin has a more striking homology with the *Ulex europaeus* lectin II (UEA-II) than with any other lectins including the *Ulex* isolectin, UEA-I. According to our previous paper [8] overall positional identity between UEA-I and II was 52.0%. On the other hand, overall identity between LAA-I and UEA-II is 82.7%.

Since these two lectins are di-*N*-acetylchitobiose-binding lectins, this extreme homology is presumed to be related to their sugar binding specificity.

Despite the differences in the properties among these lectins, the comparison of their primary sequences indicates specific residues and structural domains that are highly conserved. Most noteworthy residues are those implicated in the metal binding activities of Con A. Residues Glu⁻⁸, Asp⁻¹⁰, Asp⁻¹⁹ and His⁻²⁴ of Con A involved in manganese binding and Asp⁻¹⁰, Asn⁻¹⁴ and Asp⁻¹⁹ involved in calcium binding are all conserved in this *Laburnum* lectin [21,22].

Residues Pro⁻⁶⁸, Ser⁻⁷⁶, Phe⁻¹³⁰, Phe⁻¹⁷⁵, Pro⁻¹⁷⁸, Val⁻¹⁸⁸ and Ser⁻¹⁸⁹ which represent one-half of the residues of Con A involved in subunit-subunit interactions [23] are all but Phe⁻¹⁷⁵ conserved. This amino acid, Phe⁻¹⁷⁵, is substituted by a homologous residue, Tyr. As for the residues related to the hydrophobic cavity (Tyr⁻⁵⁴, Leu⁻⁸¹, Leu⁻⁸⁵, Val⁻⁸⁹, Phe⁻¹¹¹, Val⁻¹⁷⁹, Phe⁻¹⁹¹ and Phe⁻²¹²) all of the residues except Val⁻¹⁷⁹ are conserved in this lectin. Val⁻¹⁷⁹ is exchanged by a homologous residue, Ile. These extensive homologies demonstrate an evolutionary conservation. As for the residues of Con A involved in carbohydrate binding, only 4 residues (Tyr⁻¹², Asn⁻¹⁴, Gly⁻⁹⁸, Asp⁻²⁰⁸), out of the 6 residues implicated in the carbohydrate binding of Con A [22], are conserved in LAA-I.

The glycosylation sites were assumed to be at Asn⁻¹¹⁹. This is located in the unique sequence of -Tyr-Asn-Ser-Ser-. In the peptide fragments obtained after digestion of LAA-I with Asp-N, A15, and A13 (as shown in Fig. 2) were found to have sequences Asp-Tyr-Asn-Ser-Ser-Asn-Gln-Ile-Ile-Ala-Val-Glu-Phe- and Asp-Tyr-X-Ser-Ser-Asn-Gln-Ile-Ile-Ala-Val-Glu-Phe-, respectively. This fact suggests that LAA-I is partly glycosylated at position 119. Among the homologous lectins which contain carbohydrate, the glycosylation position is not always conserved. UEA-I is glycosylated at Asn⁻¹⁰ and Asn⁻¹¹⁶, UEA-II at Asn⁻¹¹⁸ and Asn⁻²⁴⁵, Favin at Asn⁻¹⁶⁹, SBA at Asn⁻⁷⁵, SL at Asn⁻¹¹⁸, DBA at Asn⁻¹¹⁴, and LTA at Asn⁻⁴.

As for the comparison of the hydropathy plots of LAA-I, UEA-II, UEA-I and Con A, Fig. 5, the most striking similarities occur between the region involving residues 100–230 of LAA-I and the region involving residues 220–110 of Con A, between residues 60–230 of LAA-I and residues 60–230 of UEA-II, and between

residues 140–220 of LAA-I and residues 140–220 of UEA-I. These notable similarities suggest that these lectins may have similar exposed and buried regions, suggesting similar polypeptide folding patterns. In addition to similarities in the hydropathicity profiles of these lectins, secondary structure predictions of this *Laburnum* lectin and two *Ulex* lectins show an extensive network of β -sheet structures as in Con A [21].

REFERENCES

- [1] Renkonen, K.O. (1948) *Ann. Med. Exp. Biol. Fenn.* 26, 66–72.
- [2] Morgan, W.T.J. and Watkins, W.M. (1953) *Br. J. Exp. Pathol.* 34, 94–103.
- [3] Watkins, W.M. and Morgan, W.T.J. (1962) *Vox Sang.* 7, 129–150.
- [4] Osawa, T. (1966) *Biochim. Biophys. Acta* 115, 507–510.
- [5] Matsumoto, I. and Osawa, T. (1971) *Vox Sang.* 21, 548–557.
- [6] Konami, Y., Yamamoto, K., Tsuji, T., Matsumoto, I. and Osawa, T. (1983) *Hoppe-Seyler's Z. Physiol. Chem.* 364, 397–405.
- [7] Konami, Y., Yamamoto, K. and Osawa, T. (1990) *FEBS Lett.* 268, 281–286.
- [8] Konami, Y., Yamamoto, K. and Osawa, T. (1991) *J. Biochem.* 109, 650–658.
- [9] Kyte, J. and Doolittle, R.F. (1982) *J. Mol. Biol.* 157, 105–132.
- [10] Cunningham, B.A., Wang, J.L., Waxdal, M.J. and Edelman, G.H. (1975) *J. Biol. Chem.* 250, 1502–1512.
- [11] Hemperly, J.J. and Cunningham, B.A. (1983) *Trends Biochem. Sci.* 8, 100–102.
- [12] Schnell, D.J. and Etzler, M.E. (1987) *J. Biol. Chem.* 262, 7220–7225.
- [13] Hopp, T.P., Hemperly, J.J. and Cunningham, B.A. (1982) *J. Biol. Chem.* 257, 4473–4483.
- [14] Foriers, A., Lebrun, E., Van Rappenbusch, R., De Neve, R. and Strosberg, A.D. (1981) *J. Biol. Chem.* 256, 5550–5560.
- [15] Higgins, T.J.V., Chandler, P.M., Zurawski, G., Button, S.C. and Spencer, D. (1983) *J. Biol. Chem.* 258, 9544–9549.
- [16] Kouchalakos, R.N., Bates, O.J., Bradshaw, R.A. and Hapner, K.D. (1984) *Biochemistry* 23, 1824–1830.
- [17] Hoffman, L.M., Ma, Y. and Barker, R.F. (1982) *Nucleic Acid Res.* 10, 7819–7828.
- [18] Yarwood, A., Richardson, M., Sousa-Cavada, B. and Rouge, P. (1985) *FEBS Lett.* 184, 104–109.
- [19] Adar, R., Richardson, M., Lis, H. and Sharon, N. (1989) *FEBS Lett.* 257, 81–85.
- [20] Imbrie-Milligan, C., Datta, P. and Goldstein, I.J. (1989) *J. Biol. Chem.* 264, 16793–16797.
- [21] Becker, J.W., Reeke, G.N., Jr., Wang, J.L., Cunningham, B.A. and Edelman, G.M. (1975) *J. Biol. Chem.* 250, 1513–1524.
- [22] Hardman, K.D., Agarwal, R.C. and Freiser, M.J. (1982) *J. Mol. Biol.* 157, 69–86.
- [23] Reeke Jr, G.N., Becker, J.W. and Edelman, G.M. (1975) *J. Biol. Chem.* 250, 1525–1547.