Journal of Chromatography, 597 (1992) 213–219
Elsevier Science Publishers B.V., Amsterdam

CHROM. 23 895

Purification and characterization of carbohydrate-binding peptides from *Lotus tetragonolobus* and *Ulex europeus* seed lectins using affinity chromatography

Yukiko Konami*, Kazuo Yamamoto and Toshiaki Osawa☆

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

ABSTRACT

Carbohydrate-binding peptides of several anti-H(O) leguminous lectins were obtained from endoproteinase Asp-N or Lys-C digests of L-fucose-binding *Lotus tetragonolobus* lectin (LTA) and *Ulex europeus* lectin I (UEA-I) and from that of a di-N-acetylchitobiose-binding *Ulex europeus* lectin II (UEA-II) by affinity chromatography on columns of Fuc-Gel (for LTA and UEA-I) and on a column of a mixture of several oligomers of N-acetyl-D-glucosamine (GlcNAc) coupled to Sepharose 4B (GlcNAc oligomer–Sepharose 4B) (for UEA-II). These peptides were retained on the Fuc-Gel or GlcNAc oligomer-Sepharose 4B column and were presumed to have an affinity for the columns. The amino acid sequences of the retarded peptides were determined using a protein sequencer.

INTRODUCTION

Lectins are widely used to study the carbohydrate constituents of cell surfaces and glycoproteins. They are also widely used for the isolation of glycoproteins by affinity chromatography. Leguminous lectins resemble each other in their physicochemical properties, although they differ in their carbohydrate specificities. They usually consist of two or four subunits having a relative molecular mass of 25 000–30 000, and each subunit has one carbohydrate-binding site. Their interaction with carbohydrates requires tightly bound Ca^{2+} and Mn^{2+} ions.

The primary structures of the L-focuse-binding *Lotus tetragonolobus* anti-H(O) lectin (LTA) [1] and *Ulex europeus* anti-H(O) lectin I (UEA-I) [2], and the di-N-acetylchitobiose-binding *Ulex europeus* anti-H(O) lectin II (UEA-II) [2] and *Laburnum alpinum* anti-H(O) lectin I (LAA-I) [3] have already been determined using a protein sequencer and the galactose-binding *Bauhinia purpurea* lectin (BPA) [4] by a nucleotide sequence analysis of the cloned cDNA.

A carbohydrate-binding peptide has also been isolated from BPA by affinity chromatography on a column of lactose–Sepharose 4B and the amino acid sequence of this peptide has been determined [5].

This paper reports the isolation of the carbohydrate recognition peptides of LTA, UEA-I and UEA-II with immobilized L-fucose (for LTA and UEA-I) and N-acetyl-D-glucosamine (GlcNAc) oligomer immobilized on Sepharose 4B (for UEA-II) after digestion of these lectins with an endoproteinase (Asp-N or Lys-C). The amino acid sequences of these peptide fragments were subsequently determined. The peptide sequences implicated in the carbohydrate-binding site were compared with the primary sequence of concanavalin A (Con A), the three-dimensional structure of which has already been elucidated, and the sequences were found to

☆ Present address: Yakult Central Institute for Microbiological Research, Yaho, Kunitachi-shi, Tokyo 186, Japan.

0021-9673/92/$05.00

correspond to a part of the metal binding region which is relatively conserved in several leguminous lectins.

EXPERIMENTAL

Material

Lotus tetragonolobus and *Ulex europeus* seeds were purchased from F. W. Schumacher Co. (Sandwich, MA, USA). A column of Fuc-Gel for affinity chromatography was obtained from E. Y. Lab. (San Mateo, CA, USA) and Sepharose 4B was obtained from Pharmacia (Uppsala, Sweden). A C_{18} μBondasphere (100 Å) column for reversed-phase chromatography was obtained from Waters (Burlington, MA, USA). Endoproteinases Asp-N (*Pseudomonas fragi*) and Lys-C (*Lysobacter enzymogenes*) were purchased from Boehringer (Mannheim, Germany).

Purification of LTA, UEA-I and UEA-II

LTA, UEA-I and UEA-II were isolated and purified by affinity chromatography according to the methods reported previously [1,6,7].

Affinity chromatography of peptides of an Asp-N or Lys-C digest

A mixture of tetra-N-acetylchitotetraose, penta-N-acetylchitopentaose and hexa-N-acetylchitohexaose (1:1:1, w/w/w), prepared by the method of Rupley [8], was coupled to amino-Sepharose 4B according to the method of Baues and Gray [9] (GlcNAc oligomer–Sepharose 4B). Purified lectins (0.2–0.5 mg in 200 μl of 50 m*M* Tris–HCl, pH 8.5) were digested with 2 μg of Asp-N or with 5 μg of Lys-C for 18 h at 37°C. The peptide fragments obtained after digestion of LTA and UEA-I with endoproteinase Asp-N and those of UEA-I digested with endoproteinase Lys-C were separately applied to individual columns of Fuc-Gel (5 ml). An endoproteinase Asp-N digest of UEA-II was applied to a column of GlcNAc oligomer–Sepharose 4B (5 ml). The columns were washed with 10 m*M* Tris–HCl (pH 6.8) containing 0.15 *M* NaCl, 1 m*M* $CaCl_2$ and 1 m*M* $MnCl_2$ at a flow-rate of 1.5 ml/h. Each 0.5 ml aliquot was collected and analysed by reversed-phase high-performance liquid chromatography (HPLC) on a column of C_{18} with a linear gradient (0–60%) of propan-2-ol–acetonitrile (7:3) in distilled water containing 0.1% trifluoroacetic acid in 30 min at a flow-rate of 1 ml/min. The elution was monitored by the absorbance at 220 nm. Amino acid sequence analyses of the peptides fractionated by HPLC were performed on a PSQ-1 gas-phase protein sequencer (Shimadzu, Kyoto, Japan).

RESULTS AND DISCUSSION

Affinity chromatography of peptide fragments of an Asp-N or Lys-C digest

The Asp-N digests of anti-H(O) lectins were fractionated on a column of Fuc-Gel (for LTA and UEA-I) or GlcNAc oligomer–Sepharose 4B (for UEA-II) and each fraction was then separately subjected to reversed-phase HPLC on a column of C_{18}. Figs. 1–3 show elution profiles of the HPLC analysis of the affinity fractions obtained from LTA (Fig. 1), UEA-I (Fig. 2) and UEA-II (Fig. 3). In these figures, the elution profile of the original mixture of an Asp-N digest is shown in the left-hand panel. Comparison of the profiles from LTA with this original profile clearly shows that most of the Asp-N fragments of LTA were recovered in affinity fraction 9 (Fig. 1). Interestingly, a peptide fragment indicated by arrows in Fig. 1 appears to be retained on the column of Fuc-Gel, because this fragment is a major component of fraction 10 and also clearly observed in fraction 11, suggesting that this peptide specifically interacts with fucose. The presence of carbohydrate-binding peptides in the Asp-N digests of UEA-I (Fig. 2) and UEA-II (Fig. 3) was similarly shown. For the Asp-N digest of UEA-II, two peptides (shown by two arrows in fractions 11 and 12 of Fig. 3) were retained on the column of GlcNAc oligomer–Sepharose 4B.

These carbohydrate-binding peptides were purified by reversed-phase HPLC on a column of C_{18} and their amino acid sequences were determined with a gas-phase protein sequencer. Fig. 4 summarizes the amino acid sequences of these peptides. On the basis of the amino acid sequences on the respective lectins [1,2], the peptide DSYHNIW from LTA corresponds to the sequence from Asp-126 to Trp-132, the peptide DTIGSPVNFW from UEA-I to the sequence from Asp-128 to Trp-137 and the peptides DSYFGKTYNPW and DSYFGKTYNPWDP from UEA-II to the sequences from Asp-132 to Trp-142 and from Asp-132 to Pro-144.

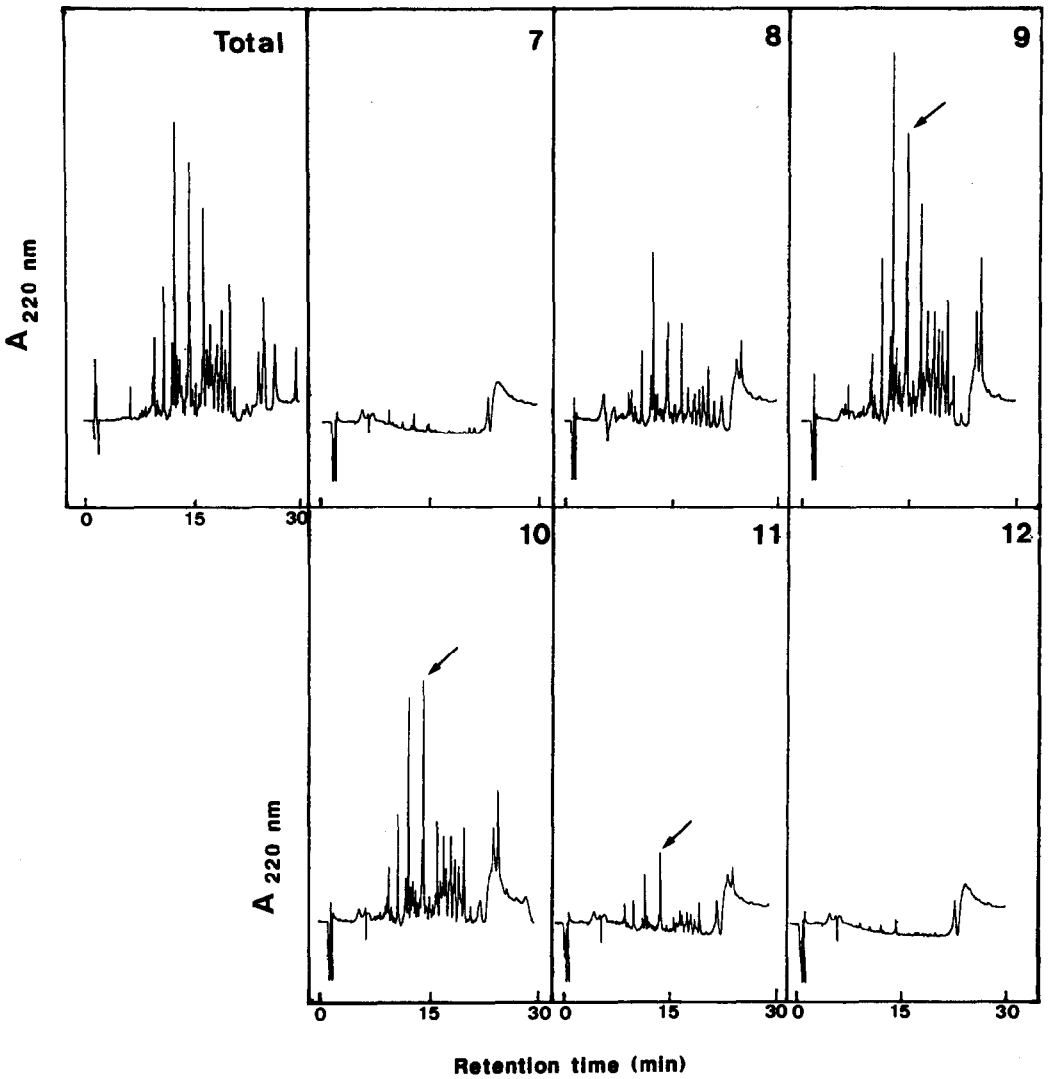

Fig. 1. Reversed-phase HPLC elution profiles of the fractions obtained by affinity chromatography (fraction number indicated in the upper right corner of each panel) of the Asp-N digest of LTA on a C_{18} column. Purified LTA was digested with an endoproteinase Asp-N at 35°C for 18 h. The reaction mixture was applied to a column of Fuc-Gel and the column washed with Tris-buffered saline containing 1 m*M* $CaCl_2$ and 1 m*M* $MnCl_2$. Each fraction was analysed by reversed-phase HPLC on a column of C_{18}. Arrows indicate the peptide of LTA retarded on the affinity column.

Lys-C fragments of UEA-I were also prepared. The peptide DNDSSYQTVAVEFDTIGSPVNFWDPGFPHIGIDVNRVK was retained on the affinity column of Fuc-Gel (Fig. 5). The peptide DTIGSPVNFW, which was obtained from UEA-I by treatment with Asp-N and retained on the affinity column of Fuc-Gel, was included in this sequence as shown by underlining in Fig. 4.

Comparison of carbohydrate-binding peptides

The amino acid sequences, which include the sequences of the putative carbohydrate-binding peptides of LTA, UEA-I and UEA-II retarded on each affinity column (underlined in Fig. 4), and their homologies with the relatively conserved regions of other leguminous lectins are shown in Fig. 6. This region has already been shown to contain a carbo-

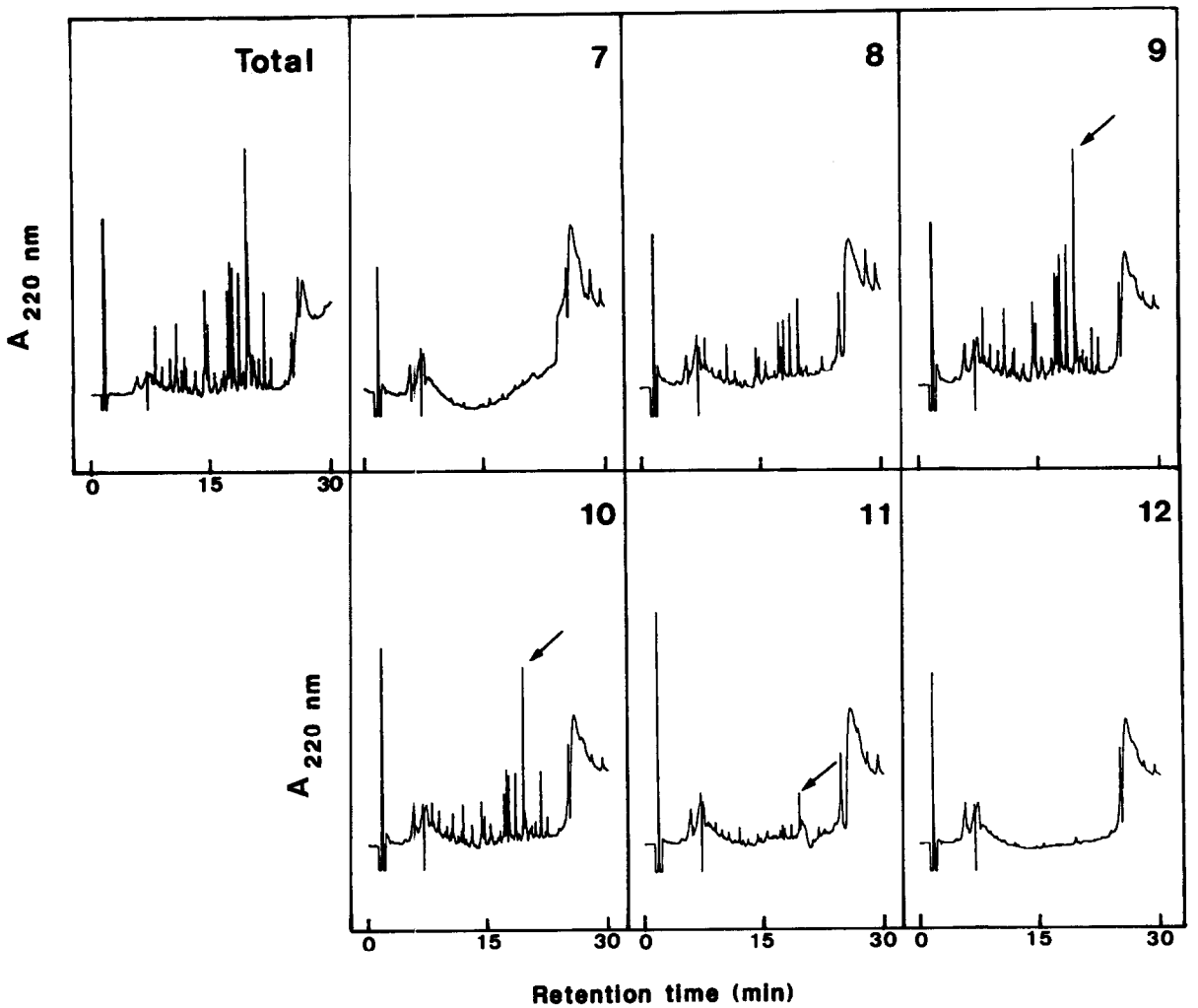

Fig. 2. Reversed-phase HPLC elution profiles on a column of C_{18} of the fractions obtained by affinity chromatography (fraction number indicated in the upper right corner of each panel) of the Asp-N digest of UEA-I on a column of Fuc-Gel. Arrows indicate the peptide of UEA-I retarded on the affinity column.

hydrate-binding peptide of BPA [10] and seems to correspond also to a metal-binding domain. The crystal X-ray analysis of con A [11] suggests that Mn^{2+} and Ca^{2+} ions are each bound to the protein by four bonds to the amino acid side-chain; the Ca^{2+} ligands are Asp-10 and the backbone carbonyl of Tyr-12, Asn-14 and Asp-19, and the Mn^{2+} ligands are Glu-8, Asp-10, Asp-19 and His-24. Val-32 and Ser-34 are suggested to be involved in the binding of Mn^{2+} to the protein by water molecules. The amino acid residues which are required for the carbohydrate binding of the anti-H(O) lectins, and discussed in this study, seem to be located in a part of the Ca^{2+}-binding region. This assumption agrees with the observation [12] that the peptide, which has been isolated from endoproteinase digests of BPA as a carbohydrate-binding peptide (underlined in Fig. 6) and also synthesized chemically, showed lactose-binding activity in the presence of Ca^{2+} and Mn^{2+}. In the absence of Ca^{2+},

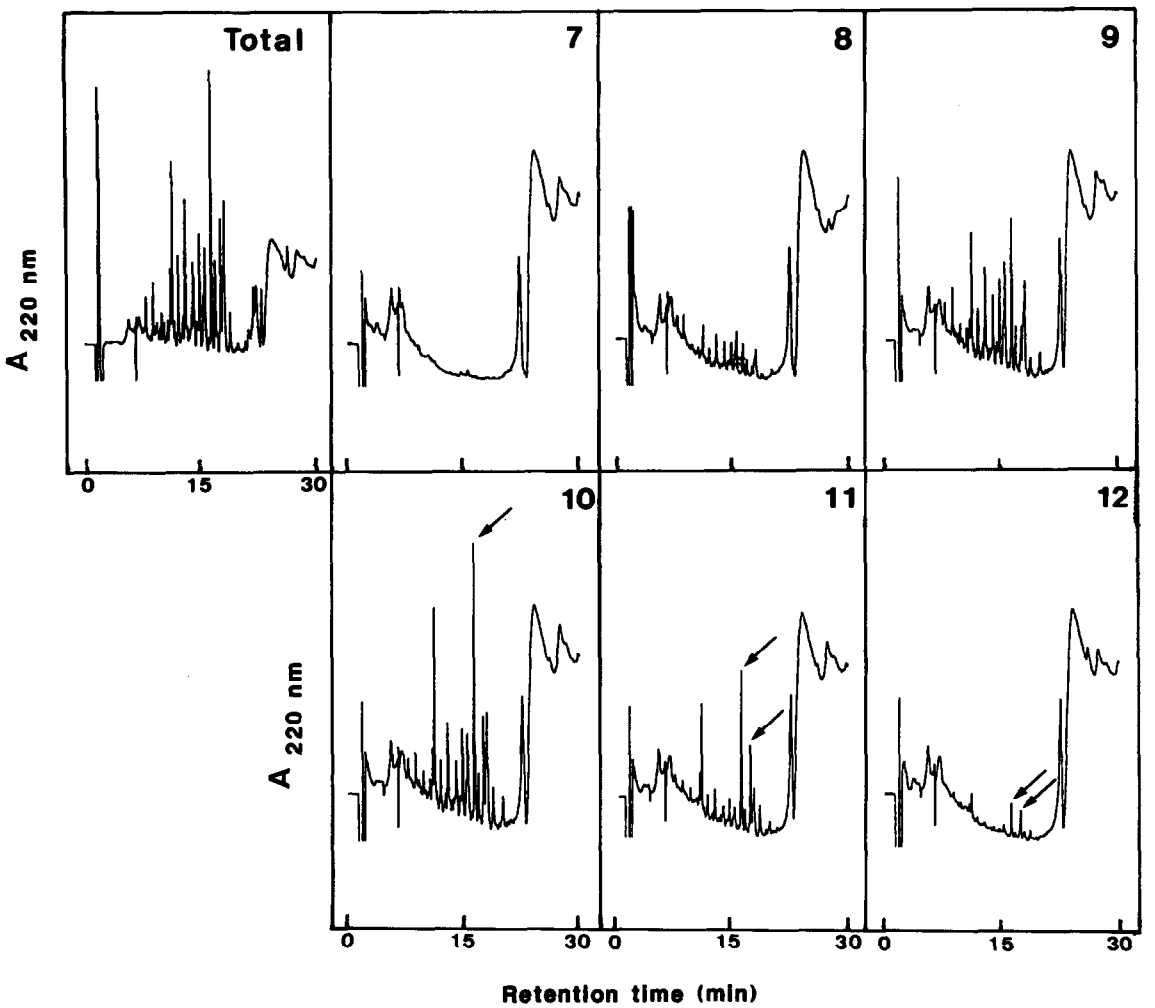

Fig. 3. Reversed-phase HPLC elution profiles on a column of C_{18} of the fractions obtained by affinity chromatography (fraction number indicated in the upper right corner of each panel) of the Asp-N digest of UEA-II on a column of GlcNAc oligomer–Sepharose 4B. Arrows indicate the two peptides of UEA-II retarded on the affinity column.

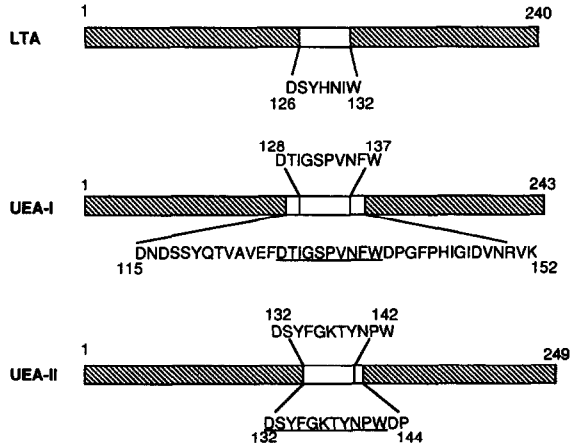

Fig. 4. Amino acid sequences of the peptides of LTA, UEA-I and UEA-II retarded on the affinity column of Fuc-Gel or GlcNAc oligomer–Sepharose 4B. Overlapped sequences are underlined.

the retardation of the peptide on the lactose–Sepharose column was not observed.

A chimeric lectin gene has been constructed using a cDNA clone coding BPA, the carbohydrate-binding peptide sequence WPNTEWS of which was replaced by the corresponding region of the mannose-binding *Lens culinaris* lectin [10]. The chimeric lectin expressed in *Escherichia coli* was found to bind α-mannosyl–bovine sereum albumin, and this binding was inhibited by mannose. This suggests that the carbohydrate-binding specificity of BPA was altered by substituting the seven amino acid residues.

Of particular interest is the amino acid sequence of arcelin, which is a major seed protein discovered in wild beans (*Phaseolus vulgaris*) and which has insecticidal activity but no carbohydrate-binding

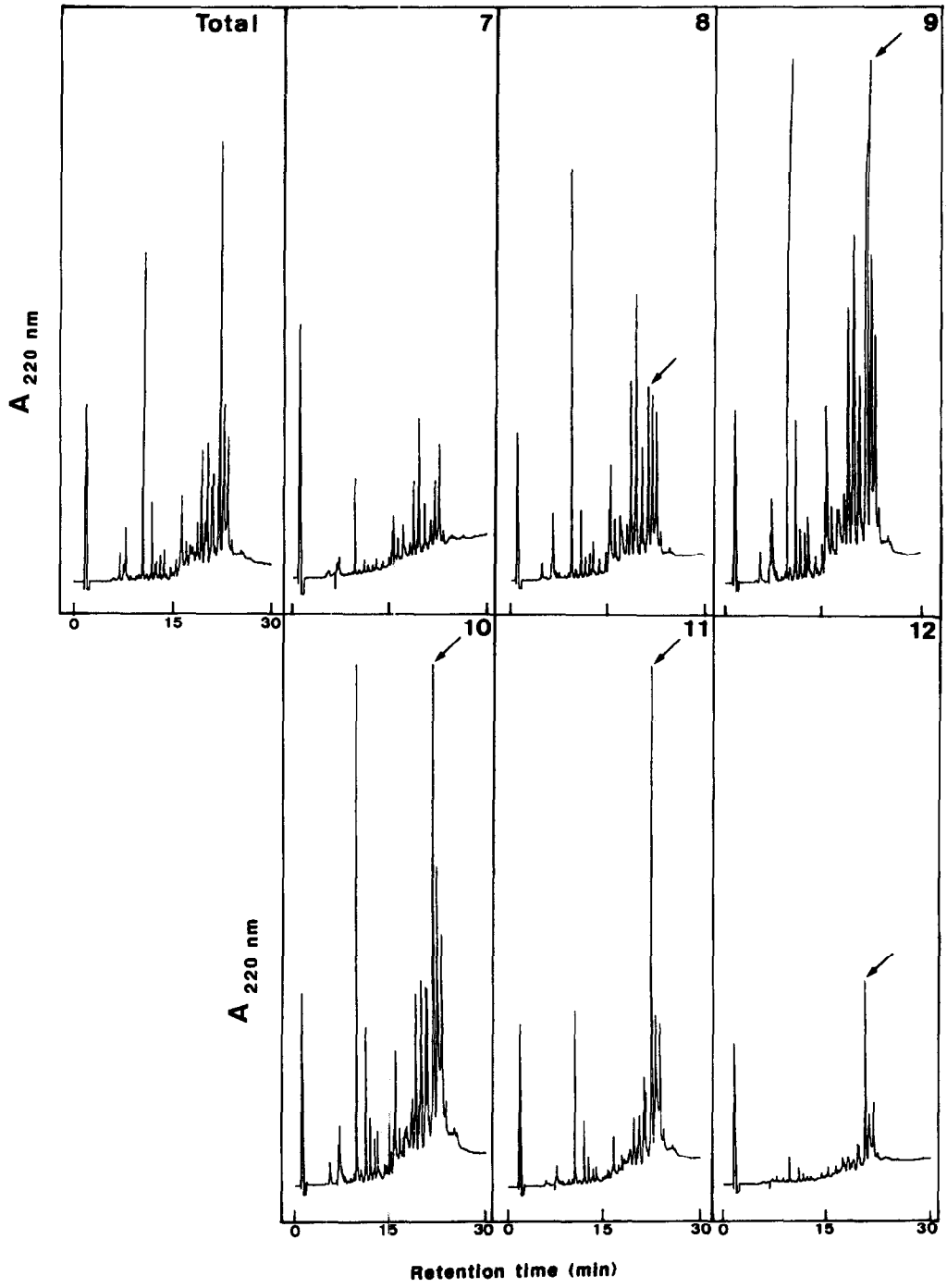

Fig. 5. Reversed-phase HPLC elution profiles on a column of C_{18} of the fractions obtained by affinity chromatography (fraction number indicated in the upper right corner of each panel) of the Lys-C digest of UEA-I on a column of Fuc-Gel. Arrows indicate the peptide of UEA-I retarded on the affinity column.

```
LTA     -VAVEFDSYH----NI-W-DPKSLRSSHVGIDVNSIMSLKA-
UEA-I   -VAVEFDTIGSP-VNF-W-DPG-FP--HIGIDVNRVKSINA-
UEA-II  -IAVEFDSYFGKTYNP-W-DPD-FK--HIGIDVNSIKSIKT-

BPA     -VAVEFDTWP----NTEWSDLRYP---HIGINVNSTVSVAT-
GS4     -VAVEFDTWI----NKDWNDPPYP---HIGIDVNSIVSVAT-
PHA-E   -VAVEFDTLY----NKDW-DPT--ER-HIGIDVNSIRSIKT-
PHA-L   -VAVEFDTLY----NVHW-DPK-P-R-HIGIDVNSIKSIKT-
ECorL   -LGVEFDTFS----NP-W-DPPQVP--HIGIDVNSIRSIKT-
DBA     -VAVEFDTLS----NSGW-DPSM-K--HIGIDVNSIKSIKT-
SBA     -VAVEFDTFR----NS-W-DPPNP---HIGINVNSIRSIKT-

PSA     -VAVEFDTFY----NAAW-DPSNRDR-HIGIDVNSIKSVNT-
LCA     -VAVEFDTFY----NAAW-DPSNKER-HIGIDVNSIKSVNT-
VFA     -VAVEFDTFY----NAAW-DPSNGKR-HIGIDVNTIKSIST-
LOL     -VAVEFDTFY----NTAW-DPSNGDR-HIGIDVNSIKSINT-
Con A   -VAVELDTYP----NTDIGDPS-YP--HIGIDIKSVRSKKT-
          8  10 12       14      19        24       32 34
```

Fig. 6. Comparison of the amino acid sequences of the putative carbohydrate-binding domains of LTA [1], UEA-I and UEA-II [2] to those of other leguminous lectins, such as BPA [4], *Griffonia simplicifolia* lectin IV (GS4) [14], *Phaseorus vulgaris* erythroagglutinin (PHA-E) [15], *Phaseorus vulgaris* leukoagglutinin (PHA-L) [15], *Erythrina corallodendron* lectin (ECorL) [16], *Dolichos biflorus* lectin (DBA) [17], *Glycine max* (soybean) lectin (SBA) [18], *Pisum sativum* lectin (PSA) [19], *Lens culinaris* (lentil) lectin (LCL) [20], *Vicia faba* (favin) lectin (VFA) [21], *Lathyrus ochrus* lectin (LOL) [22] and Con A [23]. Amino acid residues involved in calcium-binding (bold letters) and manganase-binding (outlined letters) are indicated. Underlined residues represent carbohydrate-binding peptides studied in this work. The numbers along the bottom of the figure are for the Con A sequence.

activity. The primary structure of arcelon was deduced from the nucleotide sequence of arcelin-1 cDNA [13] and the derived amino acid sequence was 58–61% identical with the amino acid sequences of *Phaseorus vulgaris* erythroagglutinin and *Phaseorus vulgaris* leukoagglutinin. Arcelin lacks the eight amino acids corresponding to the putative carbohydrate-binding sequence of *Phaseolus vulgaris* lectins, indicating that this portion plays an important part in the carbohydrate-binding activity of leguminous lectins.

REFERENCES

1 Y. Konami, K. Yamamoto and T. Osawa, *FEBS Lett.*, 268 (1990) 281.
2 Y. Konami, K. Yamamoto and T. Osawa, *J. Biochem.*, 109 (1991) 650.
3 Y. Konami, K. Yamamoto, S. Toyoshima and T. Osawa, *FEBS Lett.*, 286 (1991) 33.
4 K. Kusui, K. Yamamoto, Y. Konami and T. Osawa, *J. Biochem.*, 109 (1991) 899.
5 K. Yamamoto, Y. Konami, K. Kusui and T. Osawa, *FEBS Lett.*, 281 (1991) 258.
6 L. G. Gürtler, *Biochim. Biophys. Acta*, 544 (1978) 593.
7 Y. Konami, T. Tsuji, I. Matsumoto and T. Osawa, *Hoppe-Seyler's Z. Physiol. Chem.*, 362 (1981) 983.
8 J. A. Rupley, *Biochim. Biophys. Acta*, 83 (1964) 245.
9 R. J. Baues and G. R. Gray, *J. Biol. Chem.*, 252 (1977) 57.
10 K. Yamamoto, Y. Konami, T. Osawa and T. Irimura, *J. Biochem.*, 111 (1992) 87.
11 J. W. Becker, G. N. Reele, J. L. Wang, B. A. Cunningham and G. M. Edelman, *J. Biol. Chem.*, 250 (1975) 1513.
12 K. Yamamoto, Y. Konami, K. Kusui and T. Osawa, *FEBS Lett.*, 281 (1991) 258.
13 T. C. Osborn, D. C. Alexander, S. S. M. Sun, C. Cardona and F. A. Bliss, *Science*, 240 (1988) 207.
14 L. T. J. Delbaere, M. Vandonselaar, L. Prasad, J. W. Quail, J. R. Pearlstone, M. R. Carpenter, L. B. Smillie, P. V. Nikrad, U. Spohr and R. U. Lemieux, *Can. J. Chem.*, 68 (1990) 1116.
15 L. M. Hoffman and D. D. Donaldson, *EMBO J.*, 4 (1985) 883.
16 R. Arango, S. Rozenblatt and N. Sharon, *FEBS Lett.*, 264 (1990) 109.
17 D. J. Schnell, D. C. Alexander, B. G. Williams and M. E. Etzler, *Eur. J. Biochem.*, 167 (1987) 227.
18 L. O. Vodkin, P. R. Rhodes and R. B. Goldberg, *Cell*, 34 (1983) 1023.
19 T. J. V. Higgins, P. M. Chandler, G. Zurawski, S. C. Button and D. Spencer, *J. Biol. Chem.*, 258 (1983) 9544.
20 A. Foriers, E. Lebrun, R. V. Rapenbusch, R. De Neve and A. D. Strosberg, *J. Biol. Chem.*, 256 (1981) 5550.
21 B. A. Cunningham, J. J. Hemperly, T. P. Hopp and G. M. Edelman, *Proc. Natl. Acad. Sci. U.S.A.*, 76 (1979) 3218.
22 A. Yarwood, M. Richardson, B. Soura-Cavada and P. Rouge, *FEBS Lett.*, 184 (1985) 104.
23 D. M. Carrington, A. Auffret and D. E. Hanke, *Nature (London)*, 313 (1985) 64.