The High Specificities of *Phaseolus vulgaris* Erythro- and Leukoagglutinating Lectins for Bisecting GlcNAc or β1–6-Linked Branch Structures, Respectively, Are Attributable to Loop B*

Received for publication, December 26, 2001, and in revised form, February 25, 2002
Published, JBC Papers in Press, February 25, 2002, DOI 10.1074/jbc.M112382200

Yuko Kaneda, Robert F. Whittier, Hidenori Yamanaka, Enrique Carredano‡, Masanori Gotoh, Hiroyuki Sota, Yukio Hasegawa, and Yasuro Shinohara§

From Tokyo Research and Development, Amersham Biosciences, 3-25-1, Hyakunincho, Shinjuku-ku, Tokyo, 169-0073 Japan and ‡Uppsala Research and Development, Amersham Biosciences, Björgkatan 30, Uppsala, SE-751 25, Sweden

Despite very similar tertiary structures based upon a common framework, legume lectins exhibit an amazing variety of sugar binding specificities. While most of these lectins recognize rather discrete sugar linkages, *Phaseolus vulgaris* erythroagglutinating and leukoagglutinating lectins (E4- and L4-PHA) are unique in recognizing larger structures. E4- and L4-PHA are known to recognize complex type N-glycans containing bisecting GlcNAc or a β1,6-linked branch, respectively. However, the detailed mechanisms of molecular recognition are poorly understood. In order to dissect the contributions of different portions of each lectin, we carried out region-swapping mutagenesis between E4- and L4-PHA. We prepared six chimeric lectins by exchanging different combinations of loop B and the central portion of loop C, two of four loops thought to be important for the recognition of monosaccharides (Sharma, V., and Suroliya, A. (1997) *J. Mol. Biol.* 267, 433–445). The chimeric lectins' sugar binding activities were evaluated quantitatively by surface plasmon resonance. These comparisons indicate that the high specificities of E4- and L4-PHA toward bisecting GlcNAc and β1,6-linked branch structures are almost solely attributable to loop B. The contribution of the central portion of loop C to the recognition of those structural motifs was found to be negligible. Instead, it modulates affinity toward LacNAc residues present at the nonreducing terminus. Moreover, some of the chimeric lectins prepared in this study showed even higher specificities/affinities than native E4- and L4-PHA toward complex sugar chains containing either a bisecting GlcNAc residue or a β1,6-linked branch.

The legume lectins are a family of sugar-binding proteins found mainly in the seeds of plants belonging to the *Leguminosae* family (1–3). Lectins from leguminous plants constitute a large family of homologous proteins displaying remarkable divergence in their carbohydrate specificity. Elucidation of the mechanism by which these lectins can possess such a broad range of binding specificities while maintaining a strikingly similar three-dimensional monomer structure will be key to understanding the essence of carbohydrate-protein interactions. At present, the crystal structures of ~20 legume lectins have been solved. These various lectin monomers share a common so-called "jellyroll" structure composed primarily of a six- and a seven-stranded antiparallel β-sheet. Each monomer binds a manganese and a calcium ion that are both essential for carbohydrate binding. Amino acid sequence comparisons, x-ray crystallographic analysis, and mutagenesis studies revealed that the differences in carbohydrate specificity appear to be due primarily to differences in amino acid residues residing in loops adjacent to the carbohydrate binding site.

The primary carbohydrate-binding site of legume lectins is a shallow depression on loops associated with the concave face of the seven-stranded curved β-sheet (5). It is constructed mainly by residues from four sequentially separate regions, which are described by Sharma and Suroliya (6) as loops A, B, C, and D. Side chains of two highly conserved residues, Asp and Asn (contributed by loops A and C, respectively), together with the backbone chain NH of a Gly or Arg residue from loop B, play crucial roles in carbohydrate recognition, since they participate in four key hydrogen bonds with the monosaccharide. The carbohydrate is further stabilized by stacking interactions with hydrophobic residues present in loop C. Gross differences in the size of loop D were shown to be crucial for distinguishing between Man/Glc and Gal/GalNAc (6).

Although legume lectins have been subdivided into categories based on their monosaccharide specificities (1, 7), there are marked differences in the fine specificities of the lectins within a single category. Despite clear delineation of the primary binding site, the mechanism by which legume lectins expand fine specificities and improve binding affinity by subsite multivalence (8) is largely unknown.

Among legume lectins, E4- and L4-PHA2 are exceptional in

---

* This work was performed under the management of the Research Association for Biotechnology as part of the R&D Project on Basic Technology for Future Industries supported by the New Energy and Industrial Technology Development Organization in the Ministry of Economy, Trade, and Industry. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. E-mail: yasuro.shinohara@jp.amershambiosciences.com.

§ This work was performed under the management of the Research Association for Biotechnology as part of the R&D Project on Basic Technology for Future Industries supported by the New Energy and Industrial Technology Development Organization in the Ministry of Economy, Trade, and Industry. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 The 3D Lectin Data Bank is available on the World Wide Web at webenligne.cernav.cnrs.fr/databank/lectine/

2 The abbreviations used are: E4-PHA, *P. vulgaris* erythroagglutinating lectin; L4-PHA, *P. vulgaris* leukoagglutinating lectin; AUC, area under the curve; AUC_, AUC of the association phase; AUC, AUC of dissociation phase extrapolated over an infinite time interval; E_b-w, chimera lectin of E4-PHA of which loop B is replaced with that of L4-PHA; E-c_, chimera lectin of E4-PHA of which the central portion of loop C is replaced with that of L4-PHA; E-b-c, chimera lectin of E4-PHA of which loop B and the central portion of loop C are replaced with those of L4-PHA; L-b-w, chimera lectin of L4-PHA of which loop B is replaced with that of E4-PHA; L-c_, chimera lectin of L4-PHA of which the central portion of loop C is replaced with that of E4-PHA; L-b-c, chimera lectin of L4-PHA of which loop B and the central portion of loop C are replaced with those of E4-PHA; LacNAc, N-acetylactosamine; BPH, 4-(biotinamido) phenylacetylatedrazide; SPR, surface plasmon resonance; NA2, NA2B, NA3, NA4, and NA2G, see Fig. 1; NA4-HD, NA4 hendraecasaccharide.
displaying considerably greater degrees of oligosaccharide specificity than had previously been appreciated. Both are isolectins isolated from *Phaseolus vulgaris* (red kidney bean) seeds (9). Despite differences in binding specificities and hemagglutinating properties, subunits from E 4- and L 4-PHA are similar in molecular weight as well as carbohydrate and amino acid compositions. E 4- and L 4-PHA exhibit greatest affinity toward complex type N-glycans containing either bisecting GlcNAc or $\beta$1,6-linked LacNAc, respectively (10–14). The rather narrow differences between these two lectins make them well suited for the elucidation of subsite multivalence. Mirkov and Chrispeels (15) reported that mutation of Asn128 to Asp in loop C of L 4-PHA eliminates carbohydrate binding and biological activity, suggesting that this residue plays a key role in primary recognition. However, the mechanism of L 4-PHA specificity toward $\beta$1,6-linked branch structures is not known.

Although the three-dimensional structure of L 4-PHA has been recently solved (16, 17), neither the crystal structure of the complex nor the three-dimensional structure of E4-PHA has yet been reported. In light of these outstanding questions, we chose to evaluate the effect of loop substitution on the sugar binding specificities for the identification of subsites. To be concrete, the primary sequences of E4- and L4-PHA were compared, and loop B and the central portion of loop C were chosen for swapping, since the biggest differences among the four loops (A–D) were concentrated in these two. Sugar binding specificities of six prepared chimeric lectins were extensively analyzed using a quantitative assay system based on immobilized oligosaccharides. This study clearly suggested that each loop functions to recognize different parts of an oligosaccharide, and the combination of these loops determines both the specificities and affinities of sugar binding.

**EXPERIMENTAL PROCEDURES**

**Materials and Instrumentation**—Fetuin (insolubilized on 4% beaded agarose) was purchased from Sigma. BIACORE® (BIACORE AB, Uppsala, Sweden), which are based on surface plasmon resonance (SPR), were used to measure the biomolecular interactions. E 4- and L 4-PHA were purchased from Honen Corp. (Tokyo, Japan). The abbreviations and structures of the oligosaccharides used in this study are shown in Fig. 1. All of the oligosaccharides, except for NA4-HD, shown in Fig. 1, were purchased from Oxford GlycoSystems (Abingdon, UK). The BIACORE SensorChip SA, Surfactant P20 was obtained from BIACORE Biosciences) using the HBS buffer comprised 10 mM HEPES (pH 7.4), 150 mM NaCl, 1m M CaCl2 and 0.05% Surfactant P20 in distilled water. Lectins were purified by Superdex® 200, PC3.2/30 (Amersham Biosciences) using the HBS buffer as a solvent. Quantitation of the lectin was carried out with ultraviolet absorbance at 280 nm monitored by a SMART® (Amersham Biosciences) UV monitor. 4-(Biotinamido) phenylacetylhydrazide (BPH) was synthesized as previously reported (18).

**Preparation of Total RNA from *P. vulgaris* and First-strand cDNA Synthesis**—Total RNA was isolated from midmaturation cotyledons (0.7 g total weight, 7–10 mm in length) of *P. vulgaris* cv. Tendergreen, according to the procedure reported by Hoffman et al. (19), using the QuickPrep® Total RNA Extraction Kit (Amersham Biosciences). RNA quantitation was done by extinctional measurement of 260 nm monitored by a SMART® (Amersham Biosciences) UV monitor. 4-(Biotinamido) phenylacetylhydrazide (BPH) was synthesized as previously reported (18).

**PCR Amplification of Products for Cloning**—All PCRs used for cloning purposes were carried out with the proofreading enzyme KOD polymerase (Toyobo) according to the manufacturer's instructions.

**Cloning of E 4-PHA and L 4-PHA cDNAs and Removal of Signal Peptide Coding Regions**—E 4- and L 4-PHA cDNAs were amplified from the first-strand cDNA via the PCR. For amplifying L 4-PHA, we used the oligonucleotide primers 5′-CATGAAATTCTACATTTGAGTTCAGTCACCCACCGTGTTCT-3′ and 5′-TGGAGGTTGAGTACCTAGTAGATTGTTGAGGA-3′, and for E 4-PHA, we used the primers 5′-CATGAAATTCTACATTTGAGTTCAGTCACCCACCGTGTTCT-3′ and 5′-TGGAGGTTGAGTACCTAGTAGATTGTTGAGGA-3′. After digestion with EcoRI, each PCR-generated fragment was inserted into the Smal and EcoRI site of pBluescript II SK(+) .
Loop-swapping Mutagenesis of $E_4$- and $L_4$-PHA

Hydrophobicity Comparison—$E_4$-PHA and $L_4$-PHA were compared along their length with respect to hydrophobicity using Grease software based on the algorithms of Kyte and Doolittle (21) and Pearson and Lipman (22).

Lectin Domain Swapping and Construction of Expression Plasmids—The $E_4$- and $L_4$-PHA genes each contain two EcoI sites flanking the region encoding the divergent B loop along with most of the divergent C loop. To facilitate use of these sites for domain swapping, the genes lacking signal peptides were subcloned via gent C loop. To facilitate use of these sites for domain swapping, the region encoding the divergent B loop along with most of the divergent C loop. To facilitate use of these sites for domain swapping, the genes lacking signal peptides were subcloned via gent C loop. To facilitate use of these sites for domain swapping, the region encoding the divergent B loop along with most of the divergent C loop. To facilitate use of these sites for domain swapping, the genes lacking signal peptides were subcloned via gent C loop.

| Kinetic parameters obtained for the interaction analysis of $E_4$-PHA with immobilized NA2, NA2B, and NA3 |
|---------------------------------|
| **$k_c$** $\times 10^3$ | **$k_d$** $\times 10^{-4}$ |
| NA2 | 3.1 | 11.5 |
| NA2B | 1.3 | 1.2 |
| NA3 | 1.6 | 5.4 |

BPH Tagging of the Oligosaccharides—BPH tagging of oligosaccharides was performed under the previously reported conditions (18). Briefly, the oligosaccharide dissolved in water was incubated with a 4-fold molar excess of BPH in 30% acetonitrile at 90 °C for 1 h. After the reaction, 50 μM formate buffer (pH 3.5) was added and stored at 4 °C for 12 h to promote tautomerization from the acyclic Schiff base type hydrazone to stable β-glycosides. The reaction mixture was injected directly into reverse phase high pressure liquid chromatography to purify the BPH adducts.

Binding Activity Analysis by SPR Biosensor—Oligosaccharide binding activities of $E_4$-PHA, $L_4$-PHA, and PHA mutants were measured by SPR using a BIACORE instrument. In this system, a BPH-labeled oligosaccharide (10 pmol) was introduced onto a streptavidin sensor surface (SensorChip SA; BIACORE AB, Tokyo, Japan). Lectins (10 μg/ml in HBS) were introduced onto the surface at a flow rate of 20 μl/min. The interaction between lectin and oligosaccharide was monitored as the change in SPR response at 25 °C. After 3 min, flow was switched from sample to HBS buffer in order to initiate dissociation.

Sensor surfaces were regenerated with 50 mM H3PO4.

Medium with 0.1 mg/ml ampicillin and then induced by the addition of isopropyl-β-thiogalactoside to 1 mM. Cells were harvested by centrifugation at 6,000 × g for 10 min.

Protein Extraction and Purification—Extracts of expressed cells were obtained by sonicating cells in phosphate-buffered saline and collecting the supernatant after centrifugation at 20,000 × g for 10 min. The collected supernatant was applied to a fetuin-agarose (Sigma) column, which had been equilibrated with 7.5 volumes of HBS buffer at 4 °C. After washing the column with HBS buffer, protein was eluted with 50 mM H3PO4, followed by desalting using a PD-10 column (Amersham Biosciences). The protein was concentrated by lyophilization and purified by Superdex 200 PC3.2/30 (Amersham Biosciences) using HBS as the running buffer.

| Extracts of expressed cells |
|-----------------------------|
| NA2B | 1.3 |
| NA2B | 1.6 |

Expression of PHA and PHA Mutants in Escherichia coli—The constructed plasmids were introduced into E. coli BL21(DE3) grown on an LB plate containing 100 μg/ml ampicillin. The BL21(DE3) cells containing the plasmids were grown to the midlog phase at 37 °C in LB medium with 0.1 mg/ml ampicillin and then induced by the addition of isopropyl-β-thiogalactoside to 1 mM. Cells were harvested by centrifugation at 6,000 × g for 10 min.

Protein Extraction and Purification—Extracts of expressed cells were obtained by sonicating cells in phosphate-buffered saline and collecting the supernatant after centrifugation at 20,000 × g for 10 min. The collected supernatant was applied to a fetuin-agarose (Sigma) column, which had been equilibrated with 7.5 volumes of HBS buffer at 4 °C. After washing the column with HBS buffer, protein was eluted with 50 mM H3PO4, followed by desalting using a PD-10 column (Amersham Biosciences). The protein was concentrated by lyophilization and purified by Superdex 200 PC3.2/30 (Amersham Biosciences) using HBS as the running buffer.

BPH Tagging of the Oligosaccharides—BPH tagging of oligosaccharides was performed under the previously reported conditions (18). Briefly, the oligosaccharide dissolved in water was incubated with a 4-fold molar excess of BPH in 30% acetonitrile at 90 °C for 1 h. After the reaction, 50 μM formate buffer (pH 3.5) was added and stored at 4 °C for 12 h to promote tautomerization from the acyclic Schiff base type hydrazone to stable β-glycosides. The reaction mixture was injected directly into reverse phase high pressure liquid chromatography to purify the BPH adducts.

Binding Activity Analysis by SPR Biosensor—Oligosaccharide binding activities of $E_4$-PHA, $L_4$-PHA, and PHA mutants were measured by SPR using a BIACORE instrument. In this system, a BPH-labeled oligosaccharide (10 pmol) was introduced onto a streptavidin sensor surface (SensorChip SA; BIACORE AB, Tokyo, Japan). Lectins (10 μg/ml in HBS) were introduced onto the surface at a flow rate of 20 μl/min. The interaction between lectin and oligosaccharide was monitored as the change in SPR response at 25 °C. After 3 min, flow was switched from sample to HBS buffer in order to initiate dissociation. Sensor surfaces were regenerated with 50 mM H3PO4.

The apparent $k_d$ was analyzed by fitting the dissociation phase directly to the following equation,

\[ R_t = R_s \exp(-k_d t) \]  

(Eq. 1)
Loop-swapping Mutagenesis of E₄⁻ and L₄-PHA

FIG. 3. Calculated AUC₉ (black bar) and AUC₉₋₆₋ (white bar) for the interactions of E₄-PHA (a) and L₄-PHA (b) with immobilized NGA2, NA2, NA2B, NA3, and NA4. After the injection of each lectin at a concentration of 10 μg/ml, association and dissociation were monitored for 3 min, respectively. The AUCs for both phases was calculated using the trapezoidal rule. AUC₉ was extrapolated over an infinite time interval using nonlinear regression. RU, response units.

Loop A
E₄-PHA VPNN S GPAD
L₄-PHA VPNN A GPAD

Loop B
E₄-PHA PVGSQPDKDGG LLGLP NNYKYD SN
L₄-PHA PVGSQPDKDGG FLLGF D_ _ GSN SN

Loop C
E₄-PHA DTLXYNHWDPPKR
L₄-PHA DTLXYNKWDPTER

Loop D
E₄-PHA GITKGNVETND L
L₄-PHA GINKNVETNDVL

FIG. 4. Comparison of primary sequences of loops A−D between E₄⁻ and L₄-PHA.

using nonlinear least squares analysis (23, 24). R₉ represents the amplitude of the dissociation process. The k₉ was then analyzed by fitting the association phase directly to the following equation,

\[ R₉ = \frac{C k₉ R_{max}}{(C k₉ + k₉)[1 - \exp(-(C k₉ + k₉)t)]} \]  

(Eq. 2)

where R_{max} represents the maximum lectin-oligosaccharide complex concentration (in resonance units), and C is the constant concentration of injected lectin.

The area under the curve (AUC) for both association and dissociation phases was calculated using the trapezoidal rule. AUC₉ was extrapolated over an infinite time interval using nonlinear regression.

\[ \text{AUC} = \int R₉ dt \]  

(Eq. 3)

Substituting Equation 1 into Equation 3, and integration from t₀, when dissociation was initiated, to ∞ gives Equation 4,

\[ \text{AUC}_{R₉₋₆₋} = R₉/k₉ \]  

(Eq. 4)

Substituting Equation 2 into Equation 3 and integration from 0 to t₀ gives the following:

\[ \text{AUC}_{₀₋₆₋} = \frac{(C k₉ R_{max} - R₉)(C k₉ + k₉)}{(C k₉ + k₉)} \]  

(Eq. 5)

**RESULTS**

**Quantitative Assay to Establish Sugar Binding Specificity Profiles for Native E₄⁻ and L₄-PHA**—Five oligosaccharides (NGA2, NA2, NA2B, NA3, and NA4) were chosen as model oligosaccharides based on the known sugar binding specificities of E₄⁻ and L₄-PHA. These were biotinylated with BPH and then immobilized onto streptavidin-coated sensor surfaces. The molecular amount of each immobilized oligosaccharide is presumed to be nearly constant, since the amount of streptavidin was constant, and an excess of purified BPH-oligosaccharide was introduced (18). A solution containing the lectin was passed over the sensor surface, and binding was monitored as changes in the SPR signal.

The sensorgrams obtained for native E₄⁻ and L₄-PHA with these model oligosaccharides are shown in Fig. 2. E₄-PHA exhibited the greatest increases in resonance signal with NA2, NA2B, and NA3, and the maximum response with each was nearly equal although the patterns for the association and dissociation phases differed quite markedly. E₄-PHA also interacted with NGA2 and NA4, displaying smaller increases in resonance signals. The apparent rate constant values were calculated for NA2, NA2B, and NA3 by fitting their sensorgrams using a simple one-to-one interaction model, and these are summarized in Table I. The differences in apparent k₉ were rather small, whereas the apparent k₉ differed by up to an order of magnitude. The other sensorgrams were difficult to fit to this model. In order to obtain quantitative values for comparison among all the sensorgrams, the AUCs for both the association and dissociation phases were calculated (see details under “Experimental Procedures”). As shown in Fig. 3a, the calculated AUC₀₋₆₋ was extremely high for NA2B, and the other values in descending order were NA3 > NA2 > NGA2 ≈...
NA4. This is a reasonable order in view of the previously reported sugar binding specificity of immobilized E4-PHA during affinity chromatography (12, 14, 27).

As for L4-PHA, it also bound to all of the oligosaccharides. The peak resonance signals obtained in descending order were NA2 > NA4 > NGA2 > NA2B ~ NA3. Patterns for association and dissociation phases again varied dramatically. As shown in Fig. 3b, the calculated AUC_0° values were extremely high for NA4 and decreased for the remaining four oligosaccharides in the order NA2 > NA3 > NGA2 > NA2B. This also agreed well with the order previously reported for elution from immobilized L4-PHA during analytical affinity chromatography (10, 12, 28). These results demonstrated the validity of this method as a quantitative procedure for comparing sugar binding specificities of chimeric lectins. The calculated AUC_0° showed no correlation with the results from affinity chromatography.

The other main observation from this study is that the AUC_0° values calculated for E4-PHA with the various oligosaccharides were generally higher than those for L4-PHA. The calculated AUC_0° values of NA2B was 1,400-fold higher for E4-PHA than for L4-PHA. The only exception, NA4, exhibited a value 1.7-fold higher for L4-PHA than for E4-PHA. The calculated AUC_0° values for the remaining oligosaccharides were uniformly higher with E4-PHA than L4-PHA (by 70–270-fold).

Selection of Regions to Be Swapped among Two PHA Lec-tins—The oligosaccharide binding profiles confirmed that both native lectins can recognize fairly long sequences and that the lectins’ binding specificities can differ considerably from each other. Whereas these differences must reflect differences in their structures, the three-dimensional structure of E4-PHA has yet to be solved, so no direct three-dimensional comparison can be made. Therefore, primary structures of four loops thought to play important roles in sugar recognition were compared (Fig. 4). The greatest differences resided in loop B; of 26 amino acid residues composing this loop in E4-PHA, five differed in identity, and two were absent altogether from L4-PHA. To date, the extent of knowledge about the role of loop B in legume lectins is that a Gly or Arg residue therein hydrogen-bonds with monosaccharides (6). How this longest loop contributes to fine sugar recognition is not known and therefore merits study. Moreover, two of four variant residues in loop C were also chosen. This loop contains calcium-binding sites and a strikingly important Asn residue, whose substitution to Asp eliminates sugar binding activity (15, 29). These amino acid differences between E4- and L4-PHA were clearly reflected in a hydrophobicity scale comparison (Fig. 5). Averaged over an 11-residue window, E4-PHA is markedly less hydrophobic than L4-PHA in the C-terminal portion of loop B, whereas the opposite pattern exists for loop C. Due to the reasons described, we focused on those two loops, and six chimeric lectins were designed for the study of their sugar binding specificities.

Preparation of Chimeric Lectins—Plasmid pET-11d, containing either L4-, E4-PHA, E-bL, E-cL, E-bLcL, L-bE, L-cE, or L-bEcE cDNA, was expressed in E. coli. Each PHA was separated and purified by fetuin-agarose affinity chromatography followed by gel filtration chromatography. Purified proteins were analyzed by SDS-PAGE. Silver staining detected only a single band for each purified lectin (data not shown). Although their molecular weights were slightly reduced compared with native PHA, this reduction was most likely attributable to the absence of glycosylation. The result of gel filtration chromatography indicated that all recombinant PHAs were tetrameric, just as the native PHAs. Therefore, the valence of each lectin should be unchanged.

Quantitative Sugar Binding Specificity Analysis of Each Chimeric Lectin—The results of the sugar binding specificities of all six chimeric lectins were analyzed quantitatively by the method described above. The calculated AUC_0° values are summarized in Fig. 6.

Comparing E-bL with native E4-PHA, loop B substitution decreased the calculated AUC_0° values for NA2B by 20-fold, while increasing that for NA4 by 12.8-fold. The effects of this loop substitution on NGA2, NA2, and NA3 were rather small.

In contrast, loop C substitution (E-cL) reduced the calculated AUC_0° values by a fairly uniform order of magnitude without altering the relative specificities.

Substitution of both loops yielded a pattern of sugar binding specificities quite similar to that of native L4-PHA, as reflected by the calculated AUC_0° values for E-bLcL. The AUC_0° values obtained for NA4 increased by more than 50-fold compared with native E4-PHA and even exceeded that of native L4-PHA.

**Fig. 5. Comparison of hydrophobicity profiles between E4- and L4-PHA.** Hydrophobicity values were computed according to Kyte and Doolittle (21) using a window size of 11 residues. The scoring matrix for individual amino acids varies from −4.5 for hydrophilic Arg to +4.5 for hydrophobic Ile. The break in the hydrophobicity curve of L4-PHA corresponds to an alignment gap at residues 112 and 113. The positions of loops A, B, C, and D are indicated above the plots.
by 28-fold. Thus, this chimeric lectin acquired even higher specificity and affinity toward NA4 than native L4-PHA.

The specificity pattern obtained for L-bL was similar to that of E-bL. Compared with native L4-PHA, the AUC_{d 0/3}/H11009 obtained for NA2B increased by 15-fold, while decreasing by a drastic 120-fold for NA4. Although the AUC_{d 0/3}/H11009 for NA2B was still 2 orders of magnitude lower than that of native E4-PHA, the specificity was improved, since cross-reactivity with NA2 and NA3 was relatively reduced.

Swapping of only the C loop (L-cE) increased the calculated AUC_{d 0/3}/H11009 for all oligosaccharides. This effect was most pronounced for NA2, whose value increased by 41-fold.

When both B and C loops were introduced from E4 into L4-PHA (L-bE/cE), like L-bL, the specificity pattern was similar to that of E4-PHA. Moreover, the AUC_{d 0/3}/H11009 calculated for NA2B was comparable with that of native E4-PHA, and the specificities toward NA2 and NA3 were relatively decreased compared with E4-PHA, making this chimera lectin even more specific than the native E4-PHA for NA2B while maintaining substantial affinity.

**DISCUSSION**

Structure binding activity relationship analyses were performed for E4-PHA and L4-PHA using region-swapping mutagenesis.

To evaluate the effect of structural changes on sugar binding specificities and binding affinity, we first established an evaluation system, in which we could detect binding to immobilized oligosaccharides. The sugar binding specificities of lectins have been traditionally evaluated by retardation or binding of oligosaccharides and/or glycopeptides on immobilized lectin affinity chromatography. Clearly, such an approach has limited quantitative potential. Since lectins are often used to detect glycosylation pattern of cells or tissues, and many sugar-lectin interactions take place on cell walls or membranes, a solid phase surface with immobilized oligosaccharides has great advantages in mimicking the natural situation. In this study, we were able to improve quantitation of interactions by monitoring lectin binding to oligosaccharides immobilized onto a sensor surface.

We employed AUC as a quantitative index for evaluating sensorgrams obtained from interactions between lectins and immobilized oligosaccharides. Since this method does not require model fitting except for extrapolation using nonlinear regression, the obtained values are free from miscalculations due to inadequate models. We observed good correlation between AUC_{d 0/3} and elution order derived from immobilized lectin affinity chromatography. The AUC_{d 0/3} is a meaningful...
the exclusivity of their erythro- (E4) and leuko- (L4) agglutinates made the binding conditions nonhomogeneous (8, 30). The immobilization of oligosaccharides and free multivalent lectins explained by an enhanced subunit multivalence effect whereby ties. This magnification of relative binding strength may be to magnify the differences in relative sugar binding specificities. To the left, in loop B, changes D110N and G111K together with the additional two residues present in between these in E4-PHA would alter the topology of the binding site, most probably providing steric hindrance to the binding of NA4-HD.

As is clear from the quantitative comparison of binding specificities between E4- and L4-PHA, the relative AUC_d,0→∞ value obtained for each immobilized oligosaccharide was always much higher for E4-PHA than for L4-PHA except in the case of NA4. E-c, maintained the relative sugar binding preferences of native E4-PHA, whereas AUC_d,0→∞ values were markedly decreased. On the other hand, the opposite loop swap (L-cE) markedly increased AUC_d,0→∞ for all tested oligosaccharides. These results suggest that loop C is important for both lectins in the recognition of terminal LacNAc residue(s), with the loop from E4-PHA conferring a much higher relative affinity than that from L4-PHA.

From the crystal structure of L4-PHA (17), it can be seen that Gly-111 from loop B is located at one of the turns of the loop (see Fig. 7), and the insertion of two additional residues next to it would modify the topology considerably. In an attempt to explain how this modification would interfere with the binding of NA4, hexasaccharide, NA4-HD (Fig. 1), was docked to the binding site formed in between loops B and C. For comparison, NA2 and NA2B were also docked to this site. It was found that NA4-HD and NA2 dock favorably to the binding site formed in between and around loops B and C. For E4-PHA, whereas AUC_d,0→∞ -AI variant 1 have been solved (34, 35), and both lack loop C. Quantitative sugar binding activity analysis of each chimeric lectin revealed distinct functions of loop B and C in subsite multivalence. One of the clear conclusions from this study is that the high specificities of E4- and L4-PHA toward bisecting GlcNAc and a β1,6-linked branch, respectively, were attributable almost solely to loop B. This conclusion can be drawn from the observations that 1) those binding specificities were destroyed by removing this loop from native PHAs and 2) those binding specificities can be transferred by swapping this loop between the two lectins. To our knowledge, the only known function in sugar recognition previously attributed to this longest loop lay in hydrogen bonding through the backbone chain NH group of the Gly or Arg residue located at the end of this loop. The current study reveals its central contribution to subsite multivalence for the first time.

**Fig. 7. Ribbon diagram representing the trace of L4-PHA focused at the carbohydrate binding site formed around and in between loops B (yellow) and C (green).** Part of loop D (magenta) contributes slightly. A chicken net lipophilic color-coded Connolly surface is shown, where brown denotes lipophilic, green polar, and blue charged areas. Shown in sticks are NA4-HD docked to the structure of L4-PHA at the binding site and some of the residues that may be particularly important for the specificity. To the left and coming from loop C, mutation N128D would cause electrostatic repulsion from the neighboring Asp-129 side chain. The side chain of the residue at position 129 would then provide steric hindrance to carbohydrate binding. To the right, in loop B, changes D110N and G111K together with the additional two residues present in between these in E4-PHA would alter the topology of the binding site, most probably providing steric hindrance to the binding of NA4-HD.

parameter, which is described by R_d/R_3 if the interaction can be approximated by a one-to-one interaction model (see “Experimental Procedures”).

Although the binding specificities of E4- and L4-PHA are similar qualitatively, they differ quantitatively. In agreement with affinity chromatography studies using immobilized lectins, we found that E4- and L4-PHA most strongly recognized NA2B and NA4, respectively. However, since those oligosaccharides were only retarded by affinity chromatography and did not require hapten injection for elution, our results seemed to magnify the differences in relative sugar binding specificities. This magnification of relative binding strength may be explained by an enhanced subunit multivalence effect whereby immobilization of oligosaccharide and free multivalent lectin made the binding conditions nonhomogeneous (8, 30). The large differences in binding specificities we observed parallels the exclusivity of their erythro- (E4) and leuko- (L4) agglutinating activities (31). This analysis also showed that the high AUC_d,0→∞ values observed for NA2B and NA4 were due mainly to their extremely slow apparent dissociation rates.

Quantitative sugar binding activity analysis of each chimeric lectin revealed distinct functions of loop B and C in subsite multivalence. One of the clear conclusions from this study is that the high specificities of E4- and L4-PHA toward bisecting GlcNAc and a β1,6-linked branch, respectively, were attributable almost solely to loop B. This conclusion can be drawn from the observations that 1) those binding specificities were destroyed by removing this loop from native PHAs and 2) those binding specificities can be transferred by swapping this loop between the two lectins. To our knowledge, the only known function in sugar recognition previously attributed to this longest loop lay in hydrogen bonding through the backbone chain NH group of the Gly or Arg residue located at the end of this loop. The current study reveals its central contribution to subsite multivalence for the first time.

On the other hand, the replacements of loop C neither destroy nor transfer those sugar binding specificities. Instead, loop C was found to modulate affinity toward LacNAc residues present at a nonreducing terminus. This loop also contains the Asn residue whose substitution to Asp has been shown to eliminate sugar binding activity (15, 29). Furthermore, this loop is reported to be the primary binding site for mono- and disaccharides (32, 33). P. vulgaris contains two proteins with no known carbohydrate binding activity: the α-amylase inhibitor (α-AI) and arcelin. Both proteins share about 50–60% sequence identity with E4- and L4-PHA. The crystal structures of arcelin variant 5a and α-AI variant 1 have been solved (34, 35), and both lack loop C.
site with the ranking NA4-HD > NA2. Moreover, it was found that NA2B does not dock favorably to the binding site, all of this in agreement with the sensorgrams shown in Fig. 1b and the bars of Fig. 2b. Also, the docking conformation suggests a plausible explanation to why NA4 does not bind to E4-PHA, L-bE, and L-bE/cE, since all of these have two extra residues inserted in the B loop. From Fig. 7 it is quite clear that such an insertion would introduce prohibitive steric hindrance at the binding site.

Although both native E4- and L4-PHA are glycoproteins, none of the chimeric lectins prepared were glycosylated. At least with respect to L4-PHA, however, studies have shown that recombinant lectin produced in E. coli retains mitogenic and erythroagglutinating activities similar to those of native L4-PHA (36). To date, no similar studies have been reported for E4-PHA. To confirm this possibility, sugar binding specificities of recombinant E4-PHA produced in E. coli were compared with those of native E4-PHA. No difference was observed either qualitatively or quantitatively (data not shown). Therefore, it is quite unlikely that glycosylation differences significantly affect our measurements for the chimeric lectins.

All of the chimeric lectins prepared in this study showed qualitatively similar sugar binding specificities to E4- and L4-PHA. They are also quantitatively similar to either E4- or L4-PHA. However, some of them showed even higher specificities/affinities toward complex sugar chains containing either bisecting GlcNAc or a β1,6-linked branch. For instance, E-cL, L-bC, and L-bC/cA showed greater specificity toward NA2B than native E4-PHA, suggesting that they may be more useful for the specific detection of complex type oligosaccharides having bisecting GlcNAc. Similarly, E-bC/cA acquired greater specificity and binding affinity toward β1,6-linked branch structures.

These chimeric lectins may be useful for the histochemical study of N-glycans, since the carbohydrate moieties of cell surface glycoconjugates are known to play important roles in cell adhesion and metastasis (37). Increased expression of tri- or tetra-antennary β1,6-GlcNAc-bearing N-glycans has been correlated with metastatic potential in rodent tumor models (38) and also has been shown to be a marker of tumor progression in human breast and colon neoplasia (39). It has been recently reported that a shift in the expression of bisecting GlcNAc to highly branched β1,6-GlcNAc N-glycans plays an important role in modulating the function of cell surface glycoproteins involved in human glioma invasivity (40).

These findings may be further applicable to the rational design of de novo lectins whose sugar binding specificities could be highly tailored. Recently, Yamamoto et al. proposed the concept of artificial lectins with distinct and desired carbohydrate specificities. They introduced random mutations affecting loop C of Bauhinia purpurea lectin (4). An improved understanding of how lectin subsites contribute to the recognition of fairly long glycan sequences will improve the outlook for the fine engineering of sugar binding specificities in artificial lectins.

REFERENCES

1. Sharon, N., and Lis, H. (1990) FASEB J. 4, 3198–3208
2. Loris, R., Hamelryck, T., Bockaert, J., and Wyns, L. (1998) Biochim. Biophys. Acta
3. Eizirik, M. R. (1998) Trends Glycosci. Glycotech. 10, 247–255
4. Yamamoto, K., Maruyama, I. N., and Osawa, T. (2000) J. Biochem. (Tokyo) 127, 137–142
5. Young, N. M., and Oomen, R. P. (1992) J. Mol. Biol. 228, 924–934
6. Sharma, V., and Surolia, A. (1992) J. Mol. Biol. 287, 433–445
7. Goldstein, I. J., and Puretz, R. D. (1986) In: The Lectins: Properties, Functions and Applications in Biology and Medicine (Lienier, I. E., Sharon, N., and Goldstein, I. J., eds) pp. 33–247, Academic Press, Inc., New York
8. Rini, J. M. (1995) Annu. Rev. Biophys. Biomol. Struct. 24, 551–577
9. Leavitt, R. V., Felsted, R. L., and Bachur, N. R. (1977) J. Biol. Chem. 252, 1393–1402
10. Cummings, R. D., and Kornfeld, S. (1982) J. Biol. Chem. 257, 11230–11234
11. Hammarstrom, S., Hammarstrom, M. L., Sundblad, G., Arnarp, J., and Persson, B. (1992) J. Mol. Biol. 228, 924–934
12. Green, D. E., and Baenziger, J. U. (1987) J. Biol. Chem. 262, 12018–12029
13. Bierhuizen, M. F., De Witt, M., Govers, C. A., Ferwerda, W., Koeleman, C., Pos, O., and Van Dijk, W. (1988) Eur. J. Biochem. 175, 397–394
14. Kobata, A., and Yamashita, K. (1989) Methods Enzymol. 179, 46–54
15. Mirkov, T. I., and Chrispeels, M. J. (1993) Glycobiology 3, 581–587
16. Dao-Thi, M. H., Hamelryck, T. W., Poortmans, F., Voelker, T. A., Chrispeels, M. J., and Wyns, L. (1988) Proteins 24, 134–137
17. Hamelryck, T. W., Dao-Thi, M. H., Poortmans, F., Chrispeels, M. J., Wyns, L., and Loris, R. (1996) J. Biol. Chem. 271, 20479–20485
18. Shinohara, Y., Sota, H., Gotoh, M., Hasebe, M., Hasegawa, M., and Tsuru, M., Nakao, J., Hasegawa, Y., and Shiga, M. (1996) Anal. Chem. 68, 2573–2579
19. Hoffman, L. M., Ma, Y., and Barker, R. F. (1982) Nucleic Acids Res. 10, 7819–7826
20. Hoffman, L. M., and Donaldson, D. D. (1985) EMBO J. 4, 883–889
21. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132
22. Pearson, W. R., and Lipman, D. J. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 2444–2448
23. O’Shanessy, D. J., Brigham-Burke, M., Soneson, K. K., Hensley, P., and Brooks, I. (1993) Anal. Biochem. 212, 457–468
24. Karlsson, R., Roos, H., Fagerstrom, L., and Persson, B. (1994) Methods 6, 99–119
25. Mastruykov, V. S., Chen, K., Yang, L. R., and Allinger, N. L. (1993) Theochem. 99, 193–204
26. Carey, M., Kramer, R., Lengauer, T., and Klebe, G. (1996) J. Mol. Biol. 261, 470–489
27. Kobata, A., and Endo, T. (1992) J. Chromatogr. 597, 111–122
28. Bierhuizen, M. F. A., Tedzes, H., Schiphorst, W. C. M., Van den Eijnden, D. H., and Van Dijk, W. (1988) Glycoconjug. J. 5, 85–97
29. van Eijsden, R. R., Hoedemaeker, F. J., Diaz, C. L., Lugtenberg, B. J., de Pater, B. S., and Kijne, J. W. (1992) Plant Mol. Biol. 20, 1049–1058
30. Shinohara, Y., Hasegawa, Y., Kaku, H., and Shibuya, N. (1997) Glycobiology 7, 1201–1208
31. Egorin, M. J., Bachur, S. Z., Felsted, R. L., Leavitt, R. D., and Bachur, N. R. (1979) J. Biol. Chem. 254, 894–898
32. Konami, Y., Yamamoto, K., and Osawa, T. (1992) J. Chromatogr. (Tokyo) 111, 436–439
33. Hamelryck, T. W., Poortmans, F., Goossens, A., Angenon, G., Van Montagu, M., Wyns, L., and Loris, R. (1996) J. Biol. Chem. 271, 32796–32802
34. Boppard-Gilles, C., Rousseau, P., Rouge, P., and Payan, F. (1996) Structure 4, 1441–1452
35. Hoffman, L. M., and Donaldson, D. D. (1987) Bio/Technology 5, 157–160
36. Hakomori, S. (1985) Cell 43, 339–345
37. Asada, M., Furukawa, K., Segawa, K., Endo, T., and Kobata, A. (1997) Cancer Res. 57, 1073–1080
38. Fernandez, B., Sagman, U., Auger, M., Demetrio, M., and Dennis, J. W. (1991) Cancer Res. 51, 718–723
39. Yamamoto, H., Sweger, J., Greene, S., Saito T., Hurh, J., Sweeney, C. L., Leestma, J., Mkrdichian, E., Cerullo, L., Nishikawa, A., Ibarra, Y., Taniguchi, N., and Moskal, J. R. (2000) Cancer Res. 60, 134–142