Mutated plant lectin library useful to identify different cells

Mijung Yim, Takashi Ono, and Tatsuro Irimura*

Laboratory of Cancer Biology and Molecular Immunology, Graduate School of Pharmaceutical Sciences, University of Tokyo, Tokyo 113-0033, Japan

The 24 nucleotides comprising the carbohydrate-recognition domain of *Maackia amurensis* hemagglutinin (MAH) cDNA were randomly mutated. The mutant lectins were expressed as glutathione-*S*-transferase fusion proteins in *Escherichia coli* and 16 clones were randomly chosen. Although all of 16 recombinant lectins reacted strongly with anti-MAH polyclonal antibody, the carbohydrate-recognition domain of each was unique. As shown by agglutination studies, each mutant MAH lectin was able to bind to erythrocytes from one or more of five animal species in very distinct patterns. Thus, novel plant lectin libraries can be used to discriminate in a highly specific manner among a variety of cell types. This technology may prove to be very useful in a number of different applications requiring a high level of specificity in cell identification.

Plant lectins have great potential as tools in the identification, purification, and stimulation of specific glycoconjugates (1, 2). They also have been widely used to distinguish between cell types. A great deal of work has been conducted over the last half century to screen plant resources for unknown lectins and to characterize their carbohydrate specificities. Furthermore, as a result of recent developments in structural biology, much has been learned about the molecular basis of carbohydrate recognition by lectins (3). The general principles derived from studying the early structures have now been confirmed many times. It appears that legume lectins consist of a conserved scaffold composed of a basic set of essential and conserved residues, among which occur a limited number of variable residues that direct the specificity of the lectin. These observations suggest that it may be possible to engineer novel specificities into a legume lectin by incorporating random mutations in the sequence. This technology has the potential to vastly expand the limited range of natural lectin specificities and to permit the generation of novel carbohydrate-binding reagents that can identify a wide variety of cells.

Mammalian cells are known to bear a variety of sialylated glycoconjugates with subtle structural differences. The production of lectins that could distinguish these diverse structures would be highly useful. We speculated that such lectins could be obtained by engineering *Maackia amurensis* hemagglutinin (MAH). This lectin is unique among legume and other lectins because it recognizes carbohydrate chains containing sialic acid residues, namely sequences of Neu5Acα2–3Galβ1–3(Neu5Acα2–6)GalNAc (ref. 4; T.O., M.Y., and T.I., unpublished work) together with another isolectin engineering *Maackia amurensis* leukoagglutinin (MAL) specific for Neu5Acα2–3Galβ1–4GlcNAc (5). MAH consists of dimer of subunits with a relative molecular mass of 29,000. The nucleotide sequence of the cDNA that encodes MAH and its deduced amino acid sequence (6) indicate MAH is composed of 287 aa, including a 30-aa signal peptide. The putative carbohydrate-recognition domain of MAH was identified by comparing its amino acid sequence with those of other legume lectins (7), and early work with genetic mutations of this domain defined those amino acids that endow MAH with its binding specificity (5). Computer modeling of the three-dimensional structure of MAH with Neu5Acα2–3Galβ1–3(Neu5Acα2–6)GalNAc (8) con-firmed these observations. Together, these studies suggest that Asp135 plays an important role in determining the specificity of MAH. Asp127 and His132 also appear to be involved in the interaction with the carbohydrate substrate. The latter amino acids also participate in binding the metal ion, which presumably explains the metal-ion requirement for carbohydrate binding.

In the present study, we describe our genetic engineering of MAH and the unique specificities of the resulting lectins. The utility of these lectins in distinguishing various types of erythrocytes is shown. Given that only a limited number of natural lectins exist, this strategy represents an unique method by which to prepare an unlimited variety of plant lectins that can be used to discriminate between cells, even in those cases where genomic approaches are unable to detect any differences.

Experimental Procedures

Mutagenesis of the Carbohydrate-Recognition Domain of MAH. Site-directed mutagenesis by overlap extension (9, 10) was performed to randomize the carbohydrate-recognition domain of MAH. Four primers were used for the overlap extension:

Primer a: 5′-CCCCGGATCCACATCAGATGAGCTT-TCT-3′ (base pairs 89–105)

Primer b: 5′-ATGTCGATAATTTGGATCTNNMNNATC-MNNMNNATGMNNMNNMNNMNNGTCAAACTCTACA-GC-3′ (base pairs 457–519) (N, mixture of A/T/G/C; M, mixture of A/C)

Primer c: 5′-GATCCAAATTATCGACATATCGGAATT-GAT-3′ (base pairs 502–531)

Primer d: 5′-CCCGAATTCAGATCATGCAGTGTA-ACG-3′ (base pairs 847–864).

MAH cDNA fragments were separately amplified by PCR with primers a and b or c and d by using AmpliTaq Gold DNA polymerase (Perkin–Elmer). Standard PCR conditions were used, namely, an initial incubation for 9 min at 95°C and 1 min at 94°C was linked to 30 cycles of 1 min at 94°C, 1 min at 54°C, and 1 min at 72°C, followed by one 10-min cycle at 72°C to ensure that all amplified fragments were completely extended to the 3′ termini. The 3′ ends of the PCR products were polished by *pfu* DNA polymerase (Stratagene). The PCR products were purified by preparative agarose gel electrophoresis and quantitated. Equal molar amounts of the two PCR fragments were combined by 7 cycles of 1 min at 94°C and 4 min at 63°C. The products were used as the template DNA in a second reaction with primers a and d to generate the full-length mutant MAH cDNA.

Expression and Purification of Glutathione-*S*-Transferase (GST) Fusion Proteins. The PCR products containing the full-length mutant MAH cDNA were digested with *Bam*HI and *Eco*RI and ligated into *Bam*HI- and *Eco*RI-digested pGEX-2T. *Escherichia coli*
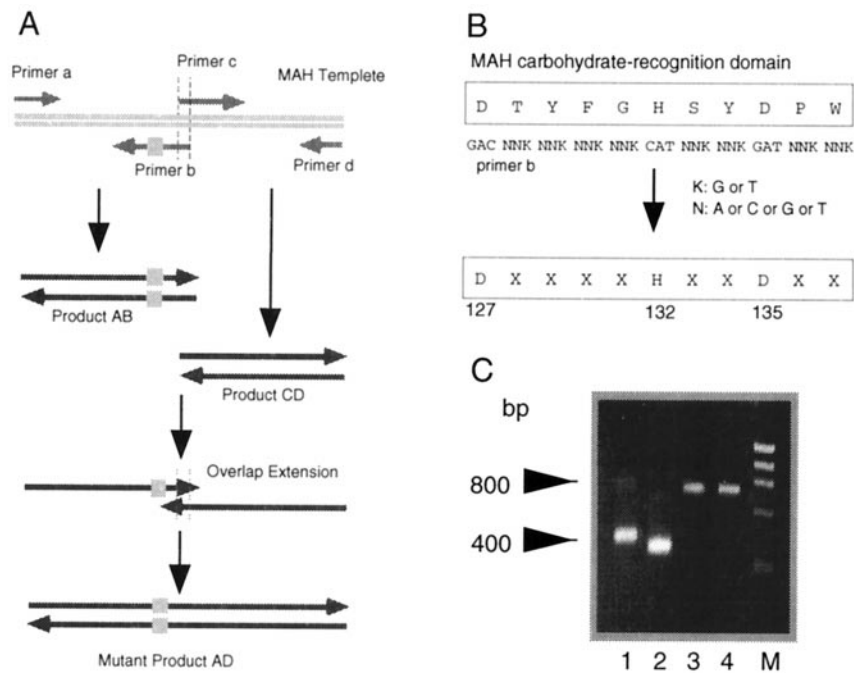

Fig. 1. Mutagenesis of MAH carbohydrate-recognition domain. (*A*) Schematic diagram describing the overlap-extension method of introducing mutations. (*B*) Mutagenesis of carbohydrate-recognition domain of MAH. (*C*) Identification of mutant MAH DNAs by agarose gel electrophoresis. Lane 1, product AB (≈400 bp); lane 2, product CD (≈400 bp); lane 3, product AD (≈800 bp); lane 4, wild-type MAH DNA (≈800 bp); and lane M, DNA marker (100-bp DNA Ladder; Life Technologies).

strain BL 21 containing the appropriate expression plasmid was grown in 2× yeast–tryptone (YT) medium supplemented with 50 μg/ml of ampicillin at 37°C until an *A* at 660 nm of 0.8 was reached. GST fusion protein expression was induced by the addition of isopropyl β-D-thiogalactoside (final concentration 1 mM), and incubation at 37°C was continued for 3–4 h. Bacterial cell pellets were resuspended in Tris-buffered saline (TBS), then the suspensions were ultrasonicated. The insoluble fraction was washed twice with TBS and solubilized in 8 M urea in TBS. The solubilized inclusion bodies were renatured by 50-fold dilution with TBS/1 mM $MnCl_2$/1 mM $CaCl_2$. The supernatant was loaded on a Glutathione-Sepharose 4B column (Amersham Pharmacia), and GST fusion proteins were purified according to the manufacturer's instructions.

Detection of GST Fusion Proteins by Western Blot Analysis. Bacterial lysates (0.5 μg of protein per ml) were separated by SDS/PAGE, transferred to nitrocellulose membranes, and then probed with polyclonal anti-MAH antibodies, followed by incubation with alkaline-phosphatase-linked anti-rabbit IgG antibodies (Zymed). Bound antibodies were visualized by the alkaline phosphatase substrate kit SK-5200 (Vector Laboratories) according to the manufacturer's instructions.

DNA Sequencing of Recombinant MAH Mutants. DNA sequences of 16 randomly chosen clones were determined by using an ABI Prism Dye-Terminator Cycle-Sequencing Ready-Reaction Kit (Applied Biosystems) with the sense primer 5′-TTCTTGCAC-CACCTGATTCTC-3′ and the antisense primer 5′-CCGCCGT-TAATCCAATCCCAT-3′.

Preparation of Erythrocytes. Human group O venous blood was drawn from a donor into syringes previously treated with heparin. The heparinized blood was transferred to glass cylinders, and the erythrocytes were allowed to sediment by gravity. After removal of leukocyte-rich plasma and buffy coat, the erythrocyte layer was washed three times by centrifugation with PBS [140 mM NaCl/2.7 mM KCl/10 mM Na_2HPO_4/1.8 mM KH_2PO_4 (pH 7.3)], each time carefully removing the top layer of cells. The erythrocytes thus obtained were found to be free from leukocytes or cell debris. Bovine, equine, porcine, and chicken erythrocytes were prepared in the same way.

Hemagglutination Assay. The purified recombinant proteins from the 16 mutant lectins were serially diluted 2-fold in a U-bottom plate (Nunc), and to each dilution (0.05 ml) an equal volume of 2% erythrocyte suspension was added. The mixture was kept for 1 h at room temperature and then examined for agglutination.

Results

Mutagenesis of the Carbohydrate-Recognition Domain of MAH. The overlap extension method (9, 10) was used to introduce mutations in the carbohydrate-recognition domain of MAH. This strategy is depicted in Fig. 1*A*. MAH cDNA cloned into pGEX-2T was used as a template for PCR amplification to generate two separate DNA fragments (AB and CD). Fragment AB was amplified with primers a and b. As Asp135, Asp127, and His132 all seem to be involved in the binding of MAH to sialic acid, primer b was designed to introduce sequence changes at the carbohydrate-recognition domain of MAH while fixing the Asp127, His132, and Asp135 residues (Fig. 1*B*). Fragment CD was amplified with primers c and d. The 5′ end of primer c bears the same 12-bp sequence found at the 3′ end of fragment AB. The common 12-bp sequences made it possible for fragments AB and CD to be joined by the overlap-extension method into the AD product containing the mutant MAH cDNA. The length of the mutant MAH cDNA (≈800 bp) was confirmed by agarose gel electrophoresis (Fig. 1*C*).

Expression and Characterization of MAH Mutant Lectins. The mutant MAH cDNAs were inserted into the *Bam*HI and *Eco*RI sites of an *E. coli* expression vector pGEX-2T that contains a GST gene
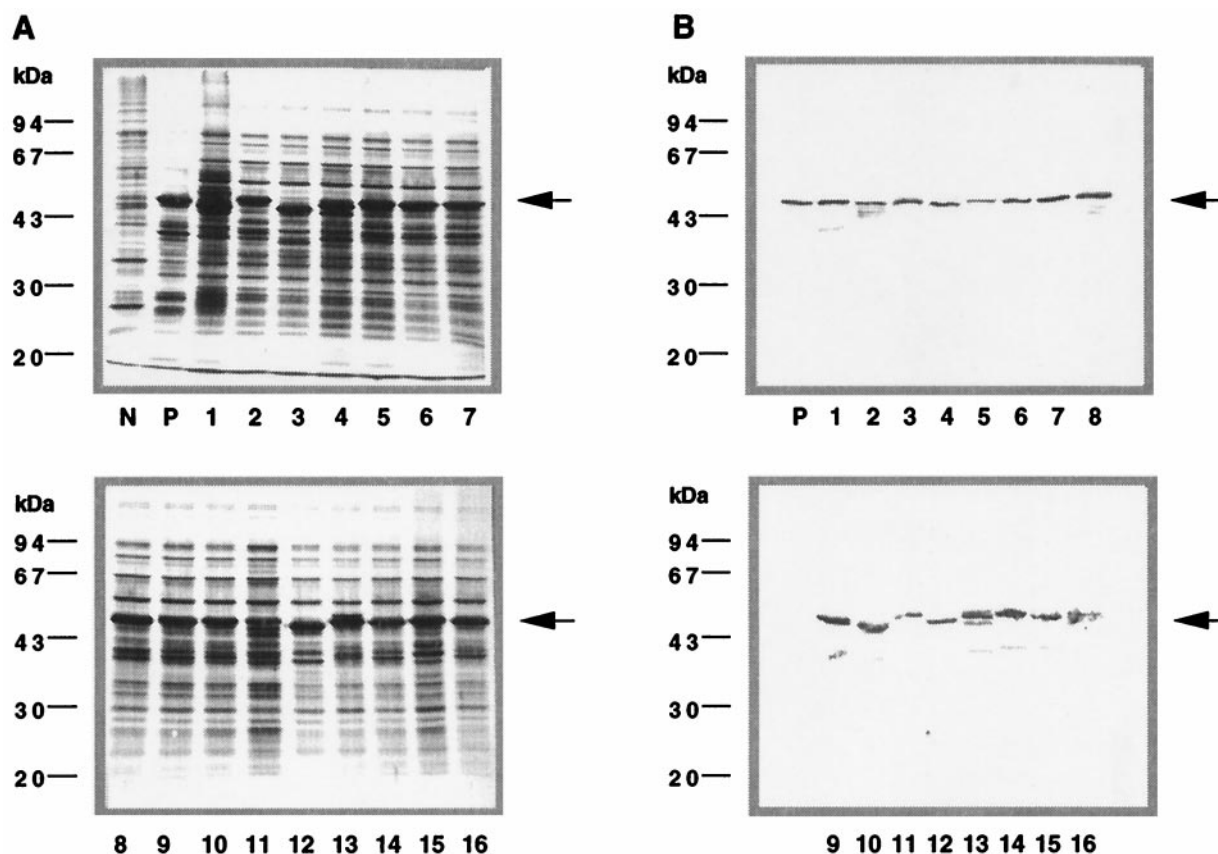
Fig. 2. Expression of the recombinant MAH mutant lectins in transformed *E. coli* induced by isopropyl thiogalactoside. (*A*) PAGE analysis of recombinant mutant MAH lectins. (*B*) Western-blotting analysis of recombinant mutant MAH lectins with anti-MAH polyclonal antibody. Lane N, *E. coli* containing the pGEX-2T plasmid; lane P, *E. coli* containing the pGEX-2T-wild type MAH plasmid; lanes 1–16, *E. coli* containing pGEX-2T-mutant MAH plasmids (clones 1–16, respectively). Relative molecular masses are indicated on the left. Arrowheads indicate recombinant MAH mutant lectins.

upstream of the insertion site. The GST-mutant lectin fusion protein was produced by using *E. coli* BL 21 as a host strain. Fifty randomly chosen clones were analyzed for their ability to synthesize recombinant mutant lectins by PAGE. The GST-mutant lectin fusion proteins of 16 clones migrated to the position corresponding to an apparent M_r of 55,600, which agrees with what was expected for a GST-mutant lectin fusion product (Fig. 2*A*).

Plasmid DNAs of the 16 clones were sequenced by the dideoxy chain termination method (11). Fig. 3 shows the deduced amino acid sequences of the carbohydrate-recognition domains of these clones. The domains of all 16 of the clones differed from one another except at the Asp127, His132, and Asp135 residues, as expected. Despite these differences, Western blotting analysis with polyclonal antibodies specific for MAH detected all 16 mutant lectins (Fig. 2*B*).

The 16 GST-mutant lectin fusion proteins were purified by affinity chromatography on Glutathione-Sepharose columns. The concentration was determined by bicinchoninic acid protein assay kit (Pierce), using BSA as a standard. Profiles of purified fusion proteins after PAGE in the presence of SDS indicated that all of 16 fusion proteins had M_r 55,600. These proteins were used in hemagglutination assays.

Hemagglutination Activity of Mutant Lectins. Substantial differences in carbohydrate specificities between the 16 mutant lectins and wild-type MAH were revealed by agglutination assays using human, bovine, equine, porcine, and chicken erythrocytes (Table 1). The carbohydrate chains on the surface of these erythrocytes differ in their structures, quantity, and patterns of

presentations, as is indeed indicated by the distinct patterns of recognition by the mutant lectins compared with each other (Fig. 4). The use of these lectins thus readily permitted the identification of each type of erythrocyte.

			130					135		
MAH	:	D	T	Y	F	G	H	S	Y	D	P	W
Clone1	:	D	N	L	I	Y	H	I	N	D	M	A
Clone2	:	D	R	P	L	T	H	S	D	D	P	V
Clone3	:	D	A	N	V	L	H	K	L	D	R	R
Clone4	:	D	V	P	N	V	H	K	I	D	Y	R
Clone5	:	D	E	S	H	C	H	V	E	D	V	E
Clone6	:	D	V	M	P	L	H	L	Q	D	T	T
Clone7	:	D	T	S	L	P	H	V	V	D	V	G
Clone8	:	D	N	L	T	G	H	L	P	D	S	G
Clone9	:	D	M	H	V	V	H	N	S	D	N	L
Clone10:		D	L	H	S	N	H	T	L	D	K	Q
Clone11:		D	P	K	P	L	H	A	I	D	S	S
Clone12:		D	V	S	H	W	H	T	Q	D	Y	R
Clone13:		D	T	E	L	L	H	T	N	D	H	K
Clone14:		D	S	A	S	N	H	V	A	D	E	R
Clone15:		D	S	S	G	I	H	D	L	D	G	R
Clone16:		D	P	P	M	A	H	Y	R	D	W	P

Fig. 3. Deduced amino sequences of the carbohydrate-recognition domains of 16 mutant MAH clones.

Yim *et al.*

Table 1. Details of hemagglutinating activity of 16 clones

Clone	Hemagglutinating activity (µg/50 µl)				
	Human	Bovine	Equine	Chicken	Porcine
MAH	<1.56	≫100	<1.56	<1.56	≫100
1	3.12	12.5	25	25	≫100
2	6.25	≫100	50	25	50
3	6.25	≫100	50	25	12.5
4	≫100	≫100	≫100	≫100	≫100
5	≫100	≫100	≫100	≫100	≫100
6	6.25	≫100	50	25	≫100
7	3.12	>50	12.5	25	25
8	6.25	≫100	50	12.5	50
9	6.25	>50	≫100	25	≫100
10	3.12	>50	12.5	50	≫100
11	6.25	25	≫100	25	≫100
12	3.12	≫100	≫100	25	≫100
13	6.25	>50	3.12	12.5	50
14	6.25	50	12.5	25	≫100
15	1.56	≫100	12.5	25	≫100
16	3.12	≫100	25	25	≫100
GST	≫100	≫100	≫100	≫100	≫100

Hemagglutinating activity was shown by minimum hemagglutinating dose against erythrocytes.

Fig. 4. The hemagglutinating activity of 16 mutant lectins toward erythrocytes of different animals. The activity is indicated by reciprocal values of minimum hemagglutinating concentration relative to one of these mutant lectins having the highest hemagglutinating activity being 1. Nos. 1–16, clones 1–16; no. 17, negative control (GST).

Discussion

The life cycles of cells in multicellular organisms are regulated by stimuli from the extracellular milieu, and such stimuli are mediated at least in part by carbohydrates and their recognition molecules (12–14). The structural variety of cell-surface carbohydrates is believed to be as diverse as that of the antigen-specific receptors in the adaptive immune system, yet diversity of the molecules designed to recognize these carbohydrates, the lectins, are limited.

In the present study, we have described a method that allows the preparation of an unlimited number of novel lectins (novel lectin libraries) with diverse specificities. We produced 16 mutant MAH lectins, and although we did not fully characterize their carbohydrate specificities, it is notable that 14 of the mutants were able to agglutinate human and domestic animal erythrocytes in highly specific patterns. Such specificity in distinguishing the erythrocytes from different species has not been achieved previously with an exception of the use of specific antibodi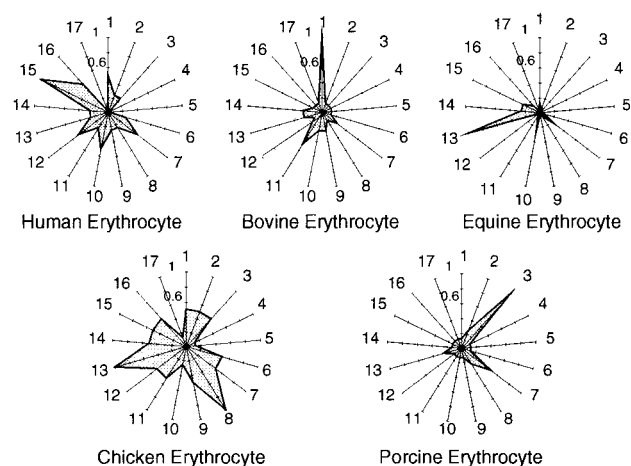es. This work proves the principle of this paper, namely that panels of genetically engineered lectins with a variety of specificities will be useful in distinguishing between different cell types.

In the past few years, new technologies have established DNA arrays to distinguish different cell types, through the profiling of the gene expression (15). However, cellular phenotypes are not solely determined by the gene expression. Posttranslational modification by glycosylation and regulation of protein localization are believed to be essential in determining the identity of a particular cell type at a given stage of differentiation. Therefore, different cells should be distinguished through profiling of cell-surface glycans. The use of a panel of lectins with diverse specificities is the most convenient and effective mean to achieve such profiling. The development of novel lectin libraries will greatly expand the utility of lectin and ultimately allow cellular identification to become even more accurate and reliable.

1. Lis, H. & Sharon, N. (1986) *Annu. Rev. Biochem.* **55,** 35–67.
2. Sharon, N. & Lis, H. (1989) *Science* **246,** 227–234.
3. Sharma, V. & Surolia, A. (1997) *J. Mol. Biol.* **267,** 433–445.
4. Konami, Y., Yamamoto, K., Osawa, T. & Irimura, T. (1994) *FEBS Lett.* **342,** 334–338.
5. Yamamoto, K., Konami, Y. & Irimura, T. (1997) *J. Biochem.* **121,** 756–761.
6. Yamamoto, K., Ishida, C., Saito, M., Konami, Y., Osawa, T. & Irimura, T. (1994) *Glycoconj. J.* **11,** 572–575.
7. Konami, Y., Ishida, C., Yamamoto, K., Osawa, T. & Irimura, T. (1994) *J. Biochem.* **115,** 767–777.
8. Imberty, A., Gautier, C., Lescar, J., Perez, S. & Wyns, L. (2000) *J. Biol. Chem.* **275,** 17541–17548.
9. Higuchi, R., Krummel, B. & Saiki, R. K. (1988) *Nucleic Acids Res.* **16,** 7351–7367.
10. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. (1989) *Gene* **77,** 51–59.
11. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
12. Hart, G. W. (1992) *Curr. Opin. Cell Biol.* **4,** 1017–1023.
13. Varki, A. (1993) *Glycobiology* **3,** 97–130.
14. Warren, C. E. (1993) *Curr. Opin. Biotechnol.* **4,** 596–602.
15. Gerhold, D., Rushmore, T. & Caskey, C. T. (1999) *Trends Biochem. Sci.* **24,** 168–173.

APPLIED BIOLOGICAL SCIENCES