

# Biochemical Properties and cDNA Cloning of Two New Lectins from the Plasma of *Tachypleus tridentatus*

TACHYPLEUS PLASMA LECTIN 1 AND 2+\*

Received for publication, September 13, 2000, and in revised form, December 7, 2000  
Published, JBC Papers in Press, December 22, 2000, DOI 10.1074/jbc.M008414200

Shang-Chiung Chen‡, Chon-Ho Yen, Maw-Sheng Yeh, Chang-Jen Huang, and Teh-Yung Liu§

From the Institute of Biological Chemistry, Academia Sinica, Nankang Borough, Taipei 115, Taiwan, Republic of China

A Sepharose CL-4B-binding protein, *Tachypleus* plasma lectin 1 (TPL-1), and a lipopolysaccharide (LPS)-binding protein, *Tachypleus* plasma lectin-2 (TPL-2), have been isolated from the plasma of *Tachypleus tridentatus* and biochemically characterized. Each protein is coded by a homologous family of multigenes. TPL-1 binds to Sepharose CL-4B and was eluted with buffer containing 0.4 M GlcNAc. The deduced amino acid sequence of TPL-1 consisted of 232 amino acids with an N-glycosylation site, Asn-Gly-Ser at residues 74–76. It shares a 65% sequence identity and similar internal repeats of about 20 amino acid motifs with tachylectin-1. Tachylectin-1 was identified as a lipopolysaccharide-agarose binding nonglycosylated protein from the ameobocytes of *T. tridentatus*. TPL-2 was eluted from the LPS-Sepharose CL-4B affinity column in buffer containing 0.4 M GlcNAc and 2 M KCl. The deduced amino acid sequence of TPL-2 consisted of 128 amino acids with an N-glycosylation site, Asn-Cys-Thr, at positions 3–5. It shares an 80% sequence identity with tachylectin-3, isolated from the ameobocytes of *T. tridentatus*. TPL-2 purified by LPS-affinity column from the plasma predominantly exists as a dimer of a glycoprotein with an apparent molecular mass of 36 kDa. Tachylectin-3 is an intracellular nonglycosylated protein that also exists as a dimer in solution with an apparent molecular mass of 29 kDa. It recognizes Gram-negative bacteria through the O-antigen of LPS. Western blot analyses showed that, in the plasma, TPL-1 and TPL-2 exist predominantly as oligomers with molecular masses above 60 kDa. They both bind to Gram-positive and Gram-negative bacteria, and this binding is inhibited by GlcNAc. Possible binding site of TPL-1 and TPL-2 to the bacteria could be at the NAc moiety of GlcNAc-MurNAc of the peptidoglycan. The physiological function of TPL-1 and TPL-2 is most likely related to their ability to form a cluster of interlocking molecules to immobilize and entrap invading organisms.

\* This work was supported in part by grants from Academia Sinica, and the Chinese Petroleum Corp. 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.

The amino acid sequences of these proteins can be accessed through NCBI Protein Database under NCBI accession numbers AF264067 (*Tachypleus* plasma lectin 1) and AF264068 (*Tachypleus* plasma lectin 2).

‡ Portions of this work were submitted in partial fulfillment for the degree of Master of Science, National Taiwan University, Taipei, Taiwan.

§ To whom correspondence should be addressed. Present address: OriGene Technologies, Inc., 6 Taft Ct., Suite 300, Rockville, MD 20850. Tel.: 301-365-3085; Fax: 301-365-7950; E-mail: dliu@origene.com and darrellliu@aol.com.

The innate and the adaptive immunities are the two general systems that mediate resistance to infectious agents. Although a certain form of adaptive immunity is present in all vertebrates, the invertebrates have developed only the innate immune system that has been thought of as an evolutionary rudiment, whose only function is to limit infection until adaptive immune response is induced. Recent studies have shown that the innate immune system has the capacity to induce costimulatory signals necessary for the activation and differentiation of lymphocytes (1–3). This finding has renewed interest on the studies of invertebrate and vertebrate innate immunology.

The innate immune system uses germline-encoded receptors for recognition of common antigens on the surface of microbial pathogens. This feature distinguishes the innate immune system found in invertebrates from the adaptive immune system of the vertebrates that possess a repertoire of specific antigen receptors and antibodies. The conserved constituents or patterns, displayed by microorganisms, are recognized by pattern recognition molecules or receptors (4). These patterns, called pathogen-associated molecular patterns, seem to be shared among groups of pathogens. The lipopolysaccharides (LPS)<sup>1</sup> of Gram-negative bacteria, lipoteichoic acid of Gram-positive bacteria, glycolipids of mycobacterium, and mannans of yeast are some examples. The innate defense system is designed to recognize those pathogen-associated molecular patterns.

The horseshoe crab, an arthropod, has evolved only a non-clonal, or innate, defense system. The hemolymph and the hemocytes carry this defense system. Whereas the hemocytes, also named ameobocytes, contain large and small granules that are filled with defense molecules, such as coagulation factors (5–7), protease inhibitors (8), and antimicrobial peptides (6), the hemolymph contains three major proteins: hemocyanin, C-reactive proteins (CRPs), and  $\alpha_2$ -macroglobulin. Hemocyanin functions as an oxygen-carrying protein. CRPs are lectins that bind to phosphocholine of the pneumococcus C-polysaccharide (9) and to the chromatin of damaged cells (10).  $\alpha_2$ -Macroglobulin exhibits protease inhibitory activity with a broad specificity that can block the activities of proteases secreted from invading microorganisms (11). The *Limulus* CRPs, along with the C3 homologue,  $\alpha_2$ -macroglobulin, participate in a complement-like hemolytic activity in horseshoe crab hemolymph.

<sup>1</sup> The abbreviations used are: LPS, lipopolysaccharide; CRP, C-reactive protein; HPLC, high performance liquid chromatography; PC, phosphocholine; PEA, phosphoethanolamine; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; PAGE, polyacrylamide gel electrophoresis; GSP, gene-specific primer; TL, tachylectin; TPL, *Tachypleus* plasma lectin; PTH, phenylthiohydantoin; BSA, bovine serum albumin; .

Several lectins with a broad range of specificity have been identified in the amoebocytes of horseshoe crab (12, 13). These lectins have been proposed to function in concert to defend horseshoe crabs from invading pathogens. However, since these lectins are present mostly in the granules of the amoebocytes, they are unlikely to be involved in the immediate-early response of host-pathogen interaction.

The plasma of horseshoe crab also contains lectin-like innate defense molecules (14, 15). In the previous study from this laboratory, we described the isolation and characterization of proteins that bind to Sepharose CL-4B, lipopolysaccharide of *Escherichia coli*, and protein A of *Staphylococcus aureus* from the plasma of *Tachypleus tridentatus* (15). In the present study, we report biochemical characterization and cDNA cloning of two of the proteins, the Sepharose CL-4B-binding protein (TPL-1) and the lipopolysaccharide-binding protein (TPL-2), which we believe are involved in the innate immunity of horseshoe crabs.

#### MATERIALS AND METHODS

**Reagents**—*E. coli* O55:B5 LPS was purchased from Sigma. Sepharose CL-4B, CNBr-activated Sepharose CL-4B, molecular weight standards, and staphylococcal protein A-Sepharose CL-4B were from Amersham Pharmacia Biotech (Uppsala, Sweden). Trypsin and complete protease inhibitor tablets were from Roche Molecular Biochemicals. Streptavidin-agarose and EZ-Link NHC-LC-Biotin were from Pierce. All other chemicals were of the highest quality commercially available.

**Horseshoe Crab and Hemolymph**—*T. tridentatus* were captured on the beaches of Quimo Island, Taiwan. Horseshoe crabs were bled by cardiac puncture, and hemolymph was collected in a conical tube containing equal volume of chilled sterile 3% NaCl supplemented with 2 mM propranolol and protease inhibitor tablets (1 tablet/50 ml) to maintain the isotonic condition and to prevent the lysis of amoebocyte (16). The amoebocytes were separated from plasma by centrifugation at  $140 \times g$  for 15 min at 4 °C. The supernatant was transferred to a new conical tube under sterile condition, filtered through a 0.2- $\mu$ m pyrogen-free filter, and loaded immediately into column.

**Preparation of LPS-Sepharose CL-4B Affinity Resin**—LPS affinity resin was prepared by coupling LPS from *E. coli* O55:B5 with CNBr-activated Sepharose CL-4B according to the instruction of manufacturer with the ligand concentration of  $2 \times 10^{-6}$  mol/ml of drained gel by assuming the average molecular mass of LPS to be 5,000 daltons.

**Purification of TPL-1 and TPL-2 from Hemolymph**—Five hundred milliliters of filtered, protease inhibitor-supplemented hemolymph was passed sequentially through three 10 mm  $\times$  10-cm tandemly linked affinity columns, packed with Sepharose CL-4B, staphylococcal protein A-Sepharose CL-4B, and LPS-Sepharose CL-4B, respectively. The columns were pre-equilibrated with initial buffer (10 mM Tris-Cl, pH 7.4, 150 mM NaCl, 10 mM CaCl<sub>2</sub>) and at the end of sample loading, washed with at least 10 column volumes of the initial buffer containing 1 M KCl until a steady base line was obtained. The columns were detached from each other. To recover TPL-1 from the Sepharose CL-4B, which served as the affinity matrix, the column was eluted with the initial buffer containing 0.4 M GlcNAc. To recover the lipopolysaccharide-binding protein (TPL-2), the LPS column was eluted with the initial buffer containing 0.4 M GlcNAc and 2 M KCl. Solid ammonium sulfate was added to the effluent fractions containing the adsorbed proteins to 50% saturation. The precipitate was collected by centrifugation at  $10,000 \times g$  for 10 min and dissolved in initial buffer. The entire purification procedure was performed at 4 °C.

**Reverse Phase HPLC Analysis**—High performance liquid chromatography was performed on an HP1100 (Hewlett-Packard) HPLC system with a C<sub>4</sub> column (214TP54, Vydac) using a flow rate of 0.25 ml/min for protein and with a C<sub>18</sub> column (218TP52, Vydac) using a flow rate of 0.15 ml/min for protease-digested peptides. The compositions of Buffer A and Buffer B were acetonitrile:water:trifluoroacetic acid at 10:90:0.1 and at 90:10:0.1, respectively. Proteins and peptides were eluted from columns with linear gradient of 0–100% Buffer B. Absorbency for proteins and peptides were monitored at 280 and 214 nm, respectively.

**Proteolytic Digestion**—Protein purified by HPLC was lyophilized and dissolved in 0.4 M NH<sub>4</sub>HCO<sub>3</sub> containing 8 M urea. After reduction with dithiothreitol and S-alkylation with iodoacetamide, three volumes of distilled H<sub>2</sub>O were added. The protein was then digested with trypsin

(E/S = 1/25, w/w) at 37 °C for 24 h. The peptides generated were separated by reversed-phase HPLC as described above using C<sub>18</sub> column (218TP52, Vydac).

**Sugar Analysis**—Periodic acid-Schiff stain was performed to assay for glycoprotein. At the end of SDS-PAGE, gel was fixed with trichloroacetic acid, oxidized with periodic acid, followed by staining with Schiff's reagent and destaining with acetic acid as described (17). Monosaccharide contents were analyzed by gas chromatograph-mass spectroscopy using the Hewlett-Packard model 6890 gas chromatograph, connected to a Hewlett-Packard 5973 mass selective detector. Samples for analysis were subjected to methanolysis, re-N-acetylation, and trimethylsilylation and dissolved in hexane prior to splitless injection into a HP-5MS fused silica capillary column (30 m  $\times$  0.32 mm, inner diameter, Hewlett-Packard). The column head pressure was maintained at around 8.2 p.s.i. to give a constant flow rate of 1 ml/min using helium as carrier gas. Oven temperature was held at 60 °C for 1 min, increased to 90 °C in 1 min, and then to 290 °C in 25 min. The trimethylsilyl derivatives were analyzed by gas chromatograph-mass spectroscopy on the Hewlett-Packard system using a temperature gradient of 60–140 °C at 25 °C/min, and then increased to 300 °C at 10 °C/min.

**Protein Sequencing and Sequence Analysis**—Sequencing of samples recovered from the reverse-phase HPLC and from SDS-PAGE/electroblottings were performed on an ABI 492 Procise automatic protein sequencer (PerkinElmer Life Sciences). The initial yield ranged from 10 to 20 pmol. The sequences were then analyzed by the GCG package (Genetics Computer Group Inc.).

**Preparation of Anti-TPL-1, Anti-TPL-2 Polyclonal Antibodies**—To raise antiserum against TPL-1 and TPL-2, proteins recovered from the eluate of Sepharose CL-4B and LPS-affinity column, respectively, were further purified by HPLC to obtain a 30-kDa TPL-1-species and a 36-kDa TPL-2-species, as judged by SDS-PAGE (15). Each of the purified protein was mixed with Freund's complete adjuvant and injected into a female New Zealand White rabbit by the intrasplenic route (18). Blood samples were collected after 4 weeks and subsequently every 7 days for the following 6–8 weeks (50 ml each time). Sera obtained were stored at –20 °C. TPL-1/TPL-2-specific IgGs were affinity-purified from immunized rabbit plasma by staphylococcal protein A column chromatography. Antibodies recovered were concentrated by ammonium sulfate precipitation (50% saturation) and used in the Western blot and immunoassay. The titer of the specific antibody was assayed by either immunoblotting or immunodiffusion.

**Western Blot Analysis**—Proteins were electrophoresed on 12% SDS-PAGE and transferred electrophoretically to nitrocellulose using an electroblot apparatus (Hoefer TE70 semidry transfer unit, Amersham Pharmacia Biotech) with constant current of 0.8 mA/cm<sup>2</sup>. The membranes were blocked with 5% (w/v) skim milk in phosphate-buffered saline supplemented with 0.1% Tween 20 and probed with specific antibody. Blots were incubated with horseradish peroxidase-conjugated anti-rabbit immunoglobulin (IgG) in the second step and developed by the enhance chemiluminescence method (ECL system, Amersham Pharmacia Biotech).

**Immunoprecipitation of TPL-1-TPL-2 Heteromer from the Plasma with Biotinylated Anti-TPL-1 Antibodies Coupled to Streptavidin-agarose**—Anti-TPL-1 antibodies were biotinylated with EZ-Link NHS-Biotin (Pierce) as described by the manufacturer. The biotinylated antibodies were incubated with streptavidin-agarose (Pierce), pretreated with 1% BSA for 1 h at 4 °C to block nonspecific binding. The gel (25  $\mu$ l) was washed with the initial buffer, and incubated with horseshoe crab plasma (2  $\mu$ g/50  $\mu$ l) overnight at 4 °C. After washing with the initial buffer, the pellet was re-suspended in the nonreducing Laemmli SDS-PAGE buffer, boiled for 5 min and the supernatant (5  $\mu$ l) was subjected to Western blot analysis using anti-TPL-2 antiserum.

**Analysis of TPL-1 and TPL-2 Binding to Bacterial Cells by Enzyme-linked Immunosorbent Assay**—To the enzyme-linked immunosorbent assay plates (Greiner F-form), suspension of bacteria in a mixture of chloroform and ethanol 1:9 (v/v), were added ( $5 \times 10^7$  cells/well), and the solvent was evaporated under a stream of warm air (19). The concentration of bacteria in the culture was determined by measuring the scattered light of the culture at optical density of 600 nm with a spectrophotometer. The number of cells/ml was estimated assuming 0.1 optical density unit is roughly equivalent to  $10^8$  cells/ml. The microplates with the adsorbed bacteria were washed with wash buffer (0.05% Tween 20 in phosphate-buffered saline), and the unbound sites were blocked with 1% BSA dissolved in the wash buffer. A serially diluted TPL-1 or TPL-2 in diluent buffer (1% BSA in wash buffer) was added to each well and incubated for 2 h at room temperature. After washing, rabbit anti-TPL-1 or anti-TPL-2 antiserum were added to each well and incubated for 2 h at room temperature.

The antiserum against TPL-1 and TPL-2 used were preadsorbed with immobilized bacteria (*Streptococcus pneumoniae* R36A, *E. coli* Bos-12, or *Vibrio parahaemolyticus*) to minimize cross-reactivity with these bacteria. After washing, horseradish peroxidase-linked anti-rabbit immunoglobulin antibody was added to each well and the plates were incubated for 2 h at room temperature. After washing with the wash buffer, 0.1 ml of 0.1 mg/ml 3,3',5,5'-tetramethylbenzidine (Sigma) in substrate buffer was added to each well and incubated at room temperature for exactly 10 min. The reaction was terminated by the addition of 0.1 ml of 2 M H<sub>2</sub>SO<sub>4</sub> and the absorbency at 450 nm was read. Since 0.4 M GlcNAc and 0.4 M GlcNAc plus 2 M KCl inhibit the binding of TPL-1 and TPL-2, respectively, to bacteria, these samples served as controls for the binding assay.

**Mass Spectrometry**—Mass spectrometric analysis of HPLC-purified TPL-1 and TPL-2 was performed on a model DE-RP MALDI-TOF (PE Biosystems, Framingham, MA). All samples were dissolved in 3,5-dimethoxy-4-hydroxycinnamic acid (sinipinic acid) at 10 mg/ml and analyzed in the positive ion mode.

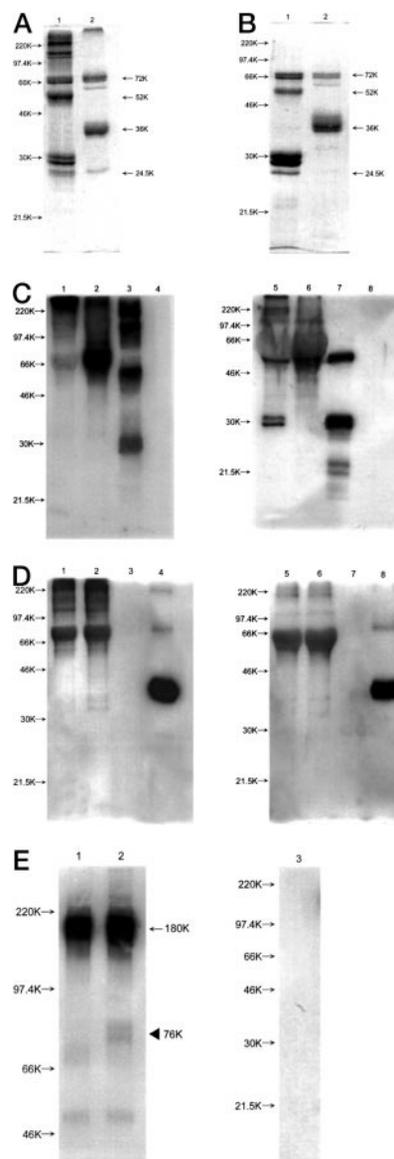
**cDNA Synthesis**—Tissues were obtained from an adult male of *T. tridentatus*. Immediately after dissection, the hepatopancreas, muscle, and hemocytes were excised and placed in liquid nitrogen. Total RNAs were prepared from hepatopancreas, using the RNazol B kit (Biotex), and poly(A)<sup>+</sup> RNAs were purified using QuickPrepR Micro mRNA purification kit with oligo(dT)-cellulose chromatography (Amersham Pharmacia Biotech). The first strand cDNA synthesis was primed with a hybrid oligo(dT) linker-primer and random primers and was transcribed using moloney murine leukemia virus reverse transcriptase (Life Technologies, Inc.). The synthesized cDNA was used as a template in subsequent PCR.

**PCR**—The primers were synthesized as follows: TPL-1, sense primer 5'-GA(A/G)TGGAC(A/C/G/T)CA(C/T)AT(A/C/T)AA(C/T)AA-3' from EWTHING (residues 1–7) of N-terminal sequence and antisense primer 5'-TT(A/G)TC(A/C/G/T)GA(C/T)TG(C/T)TT(A/C/G/T)AG(A/G)TA(A/C/G/T)CC-3' from GYKQXDN (residues 195–201) of peptide try-2; TPL-2, sense primer 5'-GA(A/G)GG(A/C/G/T)AA(A/G)(C/T)T(A/C/G/T)ATGAA(A/G)CA(C/T)CC-3' from EGKLMKHP (residues 13–20) of N-terminal sequence and oligo(dT) as antisense primer. The PCR of cDNA template was performed in a Biomtra personal cyclotherm with the following program: cycle 1, 96 °C for 2 min; cycles 2–41, 96 °C for 30 s, 50 °C for 30 s, 72 °C for 1 min; cycle 42, 72 °C for 10 min. Amplified DNA fragments were analyzed on a 1% agarose gel, ligated into pGEM-T easy vector (Promega), and transformed into *E. coli* strain JM109 by the calcium chloride method.

**DNA Sequence Analysis**—DNA sequence reaction was performed using the PRISM Ready Reaction Dye/Deoxy Terminator sequencing kit (PE Applied Biosystems). Samples were subjected to electrophoresis on an ABI 310 DNA sequencer, read automatically, and recorded using ABI Prism model version 2.1.1 software (PE Applied Biosystems).

**5'-Rapid Amplification of cDNA Ends (RACE) and 3'-RACE**—The sequences of 5' end cDNA of TPL-1 and TPL-2 were determined by using the MARATHON cDNA cloning system (CLONTECH) with gene-specific primers (GSPs). For TPL-1, GSP1 was 5'-GTATCC-AACTCCC-ATTACCATCCACTGG-3' and GSP2 was 5'-GGTACACGGTTCGCGTG-CACCTGAAGATG-3', corresponding to nucleotides 1063–1090 and 935–962, respectively; for TPL-2, GSP1 was 5'-AAATCACATTTACAC-AAGAGTTTCTCC-3' and GSP2 was 5'-AAATATTTGATTGAATACAT-TCCCAGT-3', corresponding to nucleotides 446–474 and 418–444, respectively. The primers for 3' RACE of TPL-1 were synthesized as follows: GSP1, 5'-ATATTTTCGTCGACCCGTTGACGG-3'; GSP2, 5'-AT-CTTTTCGCTGCAAGAAACCTTGC-3', corresponding to nucleotides 1276–1298 and 1390–1413, respectively. All the procedures were performed according to the manufacturer's recommendations. The amplified fragments representing overlapping 5' and 3' cDNA for TPL-1 and 5' cDNA for TPL-2 were gel-purified and cloned as described previously (20). These cDNAs were subsequently ligated together to create full-length cDNAs.

**Multiple Gene Analysis by PCR**—Genomic DNA was purified from hepatopancreas as described (20). The primers used for PCR were based on consensus amino acid sequences (residues 13–20 and 101–108) between TPL-2 and tachylepsin-3 (21): sense primer, 5'-GA(A/G)GG(A/C/G/T)AA(A/G)(C/T)T(A/C/G/T)ATGAA(A/G)CA(C/T)CC-3'; and antisense primer, 5'-TTGATTTTGTTCAGTC(A/C/G/T)AAACA-3'. The PCR reaction was carried out under the following conditions: cycle 1, 96 °C for 2 min; cycles 2–4, 96 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min; cycles 5–7, 96 °C for 30 s, 50 °C for 30 s, 72 °C for 1 min; cycles 8–41, 96 °C for 30 s, 45 °C for 30 s, 72 °C for 1 min; cycle 42, 72 °C for 15 min. The PCR product was gel-purified, cloned, and sequenced as described previously (20).



**FIG. 1. SDS-PAGE and Western blot analyses of TPL-1 and TPL-2.** \*, purified TPL-1 (5  $\mu$ g) and TPL-2 (8  $\mu$ g) were subjected to SDS-PAGE with (R) or without reduction (NR) by 2.5% 2-mercaptoethanol. The gel was stained with Coomassie Blue. The positions of the mass, in kDa, of standard proteins are indicated at the left. A, lane 1, TPL-1, no reduction; lane 2, TPL-2 no reduction. B, lane 1, TPL-1, with reduction; lane 2, TPL-2 with reduction. \*, Western blot analysis of pre- and post-column whole plasma (8  $\mu$ g/ $\mu$ l), purified TPL-1 (0.1  $\mu$ g/ $\mu$ l), and TPL-2 (0.16  $\mu$ g/ $\mu$ l) with and without reduction. C, using antiserum against TPL-1. Without reduction: lane 1, pre-column plasma; lane 2, post-column plasma; lane 3, TPL-1; lane 4, TPL-2. With reduction: lane 5, pre-column plasma; lane 6, post-column plasma; lane 7, TPL-1; lane 8, TPL-2. D, using antiserum against TPL-2. Without reduction: lane 1, pre-column plasma; lane 2, post-column plasma; lane 3, TPL-1; lane 4, TPL-2. With reduction: lane 5, pre-column plasma; lane 6, post-column plasma; lane 7, TPL-1; lane 8, TPL-2. E, detection of TPL-1-TPL-2 heteromer in the plasma. Condition of immunoprecipitation is described under "Materials and Methods." Supernatant (5  $\mu$ l) recovered from the boiling of the gel suspension with the nonreducing SDS buffer were subjected to Western blot using antiserum against TPL-2. Lane 1, buffer alone incubated with biotinylated anti-TPL-1 antibodies coupled to streptavidin-agarose; lane 2, plasma incubated with biotinylated anti-TPL-1 antibodies coupled to streptavidin-agarose; lane 3, plasma incubated with streptavidin-agarose without biotinylated anti-TPL-1 antibodies coupled to it.

## RESULTS AND DISCUSSION

**Purification of TPL-1 and TPL-2**—By passing plasma of horseshoe crabs through tandemly linked affinity columns, two new lectins were purified and characterized: TPL-1, which

1	GGGTTAAATCACTGCAAAATATGACGGCATTTTGACATGGTAATTAATCATGTATGC
64	CATAAGTGCCTCACTGATAAATCTGAGGACTATAATGCTAAAACGGGGCTTACATACCTTACAGGACACAGCAGATAGCTACTGCGAGCTTGGCCCTAACACAGAGATAA
184	ATAAACTTACTCGTGTGAAGGTTCTTAACTAACTTATTTATGTATATGATAATAGTTAAATGATGTTAATATATTAATTAATGTAACAGAGAAATAGATGGATAGGCC
304	AAAACTACTATTAAACGATTAACCTGCTATGCAAGAAGCTACATATATCTAAATGCTGATCCACTAATAGCTGAGATATATAGATATACCGATTAACAGTGGT
424	AATATCTGTAATTTAGTCTGGATGCGTCAAAATTTGACGACTTAATCAAGTTATGCTCAATCAATCAATTAACGCAACAGCAAGATGTTGACATATAAAG
544	CACGGGAGAAAATAGCTAGGGAAGCAAGTACGAATCTGCTCTAGTACTGCTTAAACAGAAATGATTTTATAGATCTTTGGGAGTATGGCTT
664	CCTGTTATTTGTTTATTTACTTATCTGCTCATGCGATGAGATGGGAAAGTGGTCTGCAATATAAGAAAGCTGATCACTGCAACAGCAATTAAGATTCACAGAGCCAGC
784	ATGAAAAATATAATGTATTTCAGTTTACTGTTAGTTACGTTGTTGCTGACCTTTCTTGTGCTGTCTCCAACATTTG
-24	<i>M K N I M Y F S L V T L L L T F L V V S P T L</i>
853	GCAGAATGGACCCACATTAATGAAAGCTTTCACACCTTACAGTTACACCAAGGTTTGTATGGGGCGTT
-1	<i>A E W T H I N G K L S H L T V T P R F V W G V</i>
922	AATAATGTTTCATGACATCTTCAGGTGCACGCGACCGTGTACCGGTAGCAACTGGATAAAAAGTGAAGGA
23	<u>N N V H D I F R C T R P C T G S N W I K V E G</u>
991	AGCTTGAAACAAATGACGCCGATGACCACGAGGTTTGGGGAGTAAACTCGAATGACAACATATAAAG
46	<i>S L K Q I D A D D H E V W G V N S N D N I Y K</i>
1060	CGCCACTGGATGGTAATGGGAGTTGGATACAAAATCAAAGGTGGACTCAAAACAGCTCAGTGCTTGGC
69	<i>R P V D G N G S W I Q I K G G L K H V S A S G</i>
1129	TACGGCTACATCTGGGGTGTAACTCGAAGACCAAATCTTCAAATGCTCAAGCCTTGTAAATGGAGAA
92	<i>Y G Y I W G V N S K D Q I F K C P K P C N G E</i>
1198	TGGGAGTGGTAGATGCTCTAAAACAAGTGGACGGTGGACGAGATCGGTATATGGAGTTACTCAG
115	<i>W E L V D G S L K Q V D G G R D L V Y G V T Q</i>
1267	AATGACGAAATATTTCTGTCGACCCGTTGACGGCAGTGGTGTCTGGGTGAATATCCCTGGAAAACTAAAA
138	<i>N D E I F R R P V D G S G V W V N I P G K L K</i>
1336	CACATCAGTGGCTCTGGCAGCTGGGAAGTGTGGGGTAAACTGTAATGATCAGATCTTTCGCTGCAAG
161	<i>H I S G S G S W E V F G V N C N D Q I F R C K</i>
1405	AAACCTGCTCAGGACAGTGGTAAGCTGTCAGGATACCTGAAGCAGTGTGATGCCAGTGGAGACTCT
184	<i>K P C S G Q W V R L S G Y L K Q C D A S G D S</i>
1474	CTTCTGGGTGCAACTCTAACGATGACATTTTGTAGTCAGTCCCTGCTTCCAAGTCTGCTGGATGAAT
207	<i>L L G V N S N D D I F E S V P A S K S C W M N</i>
1543	CCTTCTATAGATCACTTCCAACCTCGAGAATATTGTAGATCTATTATGAAAAATGGTTTAAATGAA
230	<i>P F L *</i>
1612	ACTTATAGTATTTTGTATAGTAACTCAAGTATGTTACTAAAGCAATAAATACCACATTAACAGCTCTGTTGTTGTTTGTATTTTATGAGGAGATAATTT
1732	TGCTAAGTACACTATTACTCTTAAAGTTTTTTGTTGCCATGGTACCTAATGTTCACTTCTAGAGCTTCGGGTGATCAATTAAGTTTTTAAATTAATGCTCAGAGTAAGT
1852	TTTCACTGTTTTTTTTCTTCTGCTCAATATGTCGACGGCCGTTTTGGGTTTTGCAAGAATGCTGCTGATTTCAATATCTTCCACGCTGGATTTGAATGAACAAATAATG

FIG. 2. Nucleotide and deduced amino acid sequences of TPL-1. Nucleotide and amino acid residues are numbered on the left. The underlines represent sequences determined by amino acid sequence analysis of the N terminus of the intact protein and a tryptic peptide. The residues in the gray box indicate corresponding nucleotide sequences used as primers for PCR. The broken underlines correspond to nucleotide sequences used in 5'-RACE and 3'-RACE. The putative signal sequence and glycosylation site are printed in *italics* and *bold*, respectively. An asterisk marks the stop codon.

binds to Sepharose CL-4B; and TPL-2, which binds to the LPS of *E. coli*. The use of Sepharose CL-4B as a "pre-column," prior to the passage of plasma through the LPS-Sepharose CL-4B, allowed the separation of TPL-1 from TPL-2. TPL-1 and TPL-2 could not be eluted from their respective affinity column with 2 M KCl, EDTA, galactose, or lactose. In the previous study, both proteins were eluted from the respective affinity column with buffer (10 mM Tris-Cl, pH 7.4, 150 mM NaCl) containing 4 M urea or 2 M guanidium chloride (15). Proteins eluted with 4 M urea or 2 M guanidium chloride gradually formed irreversible precipitate upon removal of the chaotropic agents. In the present study, 0.4 M GlcNAc was found effective in eluting TPL-1 from the Sepharose CL-4B column, while elution of TPL-2 from the LPS-affinity column required 2 M KCl in addition to 0.4 M GlcNAc. Proteins eluted with GlcNAc remain in solution after removal of GlcNAc and can be reabsorbed to the affinity column. One passage of the plasma through the affinity columns depleted all of the proteins that bind to Sepharose CL-4B and LPS-columns. Subsequent to the elution of the columns with GlcNAc, insignificant amount of other proteins were eluted with 4 M urea or 2 M guanidium chloride.

The mechanism of the binding of TPL-1 to the Sepharose CL-4B and its elution by GlcNAc is not known. Sepharose is a polymerized form of agarose consisting of repeating unit of  $\alpha$ -1,6-linked D-galactose and an unusual 3,6-anhydro-L-galactose. Sepharose CL is prepared from Sepharose by reacting

with 2,3-dibromopropanol under alkaline condition, resulting in cross-linkages between the 6-OH of D-galactose of one chain and 2-OH of the 3,6-anhydro-L-galactose of the other chain, via 2-hydroxypropyl bridges (22). Evidently, TPL-1 binds to this structure in a GlcNAc-dissociable manner.

Tachylectin-2 was isolated from the amebocyte of horseshoe crabs by dextran sulfate chromatography and shown to exhibit high affinity for both GlcNAc and GalNAc (23). Tachylectin-P was purified from the perivitelline fluid of horseshoe crab by using an affinity column consisting of bovine submaxillary gland mucin attached to Sepharose 4B, and eluted from the column with GlcNAc (24). From the plasma of horseshoe crabs, tachylectin 5A and 5B were purified, using an *N*-acetylated resin and elution of the proteins by GlcNAc (14). While exhibiting a common specificity of binding to the *N*-acetoamido moiety of hexoses, tachylectin-2, tachylectin-P, and tachylectin-5A/B do not share any sequence homology with each other. Whether these proteins would bind to unmodified Sepharose CL-4B like TPL-1, and be eluted from it by GlcNAc, is not known.

TPL-2 can be eluted from the LPS-affinity column with buffer containing 0.1% LPS and 2 M KCl. However, the LPS-eluted TPL-2 could not be completely separated from LPS. GlcNAc is a component of the LPS used in the preparation of LPS-affinity column. Thus, GlcNAc was used to elute LPS-binding protein.

```

TPL-1 : ----- : -
TL1 : ----- : -
TL-P : ----- : -
tectonin1 : ----- : -
tectonin2 : MATPVYIVLPSRPDHVSTAKDHAVQVTEHLHPVDSQTEGRLLWIFNDQYIRLAHHNLLVDVNGGA : 68

TPL-1 : ----- : -
TL1 : ----- : -
TL-P : ----- : -
tectonin1 : ----- : -
tectonin2 : AREGNTVLSYPPDKDHAKNQLWVKNKDCILHKLDESFLGVNDRGQVITQKKEQVILRAPASLEKR : 136

TPL-1 : --EETHINGKYSLEVTF--REVNGVNVVHDFRTRTRDTSNNWIKVSGSLKQIDA-DEHEVGVNRS : 63
TL1 : -VCHGHPGKEMHIFATP--RELVGVNSWQCYLGRQCTVDGCTQTSGLSKQVDA-DEHEVGVNRR : 64
TL-P : --CHGHPGKEMHIFATP--RELVGVNSWQCYLGRQCTVDGCTQTSGLSKQVDA-DEHEVGVNRR : 63
tectonin1 : MVNREKHEEISVYVCGAGSDHGVNHLGHWYD---SHKREKVDSEITVGVNHEEYVGVNRR : 64
tectonin2 : FSAERHEEGLNVVAVGAGNIVGVNHLGHWYD---SKRHQELSAATNLSVGLDGTGVNKA : 200

TPL-1 : DNIYKRPVDSGSSNIOIKSGIKRHSASGYGYIWGVS*DOTFCPCPCNG*HELVDGSLKQVGGRA- : 130
TL1 : DNIYKRPVDSGSSNIVRS*SLRHSASGYGYIWGVS*DOTFCPCPCNG*AWCNRR*HQIDGGCS- : 131
TL-P : DNIYKRPVDSGSSNIVRS*SLRHSASGYGYIWGVS*DOTFCPCPCNG*AWCNRR*HQIDGGCS- : 130
tectonin1 : HNIYRLDR-SNNK*QTPSEIVGVSGSHHGVNHLDEYK*DHHRN-K*OKIDGATN*VSGK*G : 130
tectonin2 : HEIYRLDR-ETNK*SIWPEIVGVSGSHHGVNHLDEYK*W*ADSR-S*WTFVGD*TN*VSGK*G : 266

TPL-1 : ILYGVTONDEIFRRPVDGSGVWVNFPAKHHISGSSSHEVFGVNC*DOE*RCN*PC*SG*V*E*E*Y*E : 198
TL1 : ILYGVTONDEIFRRPVDGSGVWVNFPAKHHISGSSSHEVFGVNC*DOE*RCN*PC*SG*V*E*E*Y*E : 198
TL-P : MYGVNS*ANLYRRPVDGSGS*CS*SS*SKH*TS*SS*IE*EV*GV*ND*Q*LY*RG*FC*CS*G*NS*LD*G*E* : 198
tectonin1 : TVYGVNRGHQIYWDGS---KVDLVLSGVVCHVSDAKRIVGVNHLDEYK*LEH--GKDKKEDG*E* : 193
tectonin2 : TVYGVNRGHQIYHYNGN---S*DAV*SG*E*V*CH*V*H*V*AN*K*LV*GV*N*K*AG*H*V*E*L*E*H--GKDKKEDG*E* : 329

TPL-1 : QCDRSGD-SLLVHNSDDFTEVFPASKSCWMPFL : 232
TL1 : QCDATGN-TIVVHNSDDVHNSG : 221
TL-P : QCDAT----- : 203
tectonin1 : MVSVGHGGEVGVNHLHHEKATL----- : 217
tectonin2 : WVAVGHGGEVGVNHLHHEKALL----- : 353

```

FIG. 3. Sequence alignment of TPL-1 with tachylectin-1 (TL1) and tachylectin-P (TL-P) of *T. tridentatus* and tectonin-1 and tectonin-2 of *Myxomyces physarum polycephalum*. The alignment was performed using CLUSTALW and PILEUP program of GCG package. Identical residues are shown in gray boxes.

TPL-1 and TPL-2 isolated accounted for about 0.1% and 0.02–0.04%, respectively, of the total hemolymph proteins. Although the amount of TPL-1 remained fairly constant, the amount of TPL-2 decreased rapidly after the animals have been kept in captivity.

**Biochemical Properties of TPL-1**—On SDS-PAGE under non-reducing condition (Fig. 1A, lane 1), TPL-1 showed two major bands at around 30 kDa, a band at around 52 kDa, two bands at around 66 kDa, and additional bands at over 100 kDa. Under reducing condition (Fig. 1B, lane 1), protein bands at 30, 52, and 66 kDa were detected. In Western blot analysis, antibodies raised against the 30-kDa TPL-1, reacted with these protein bands in nonreducing (Fig. 1C, lane 3) and reducing (Fig. 1C, lane 7) SDS-PAGE. Protein bands with even higher molecular masses (52 kDa and above) were detected in both the pre- and post-column plasma samples, under nonreducing SDS-PAGE (Fig. 1C, lane 1, pre-column; lane 2, post-column) and reducing (Fig. 1C, lane 5, pre-column; lane 6, post-column). TPL-2 did not react with anti-TPL-1 serum (Fig. 1C, lanes 4 and 8), validating the specificity of the antibodies.

Previously (15), we described a protein band with a molecular mass of 40 kDa as the major protein eluted by 4 M urea from Sepharose CL-4B. With 0.4 M GlcNAc solution as eluent, the 40-kDa protein was not detected (Fig. 1, A (lane 1) and B (lane 1)). The 30-kDa protein (TPL-1) was shown to have identical amino-terminal sequence with the previously published sequence of GBP (15).

The gene sequence data predicted 232 amino acid residues for TPL-1 (Fig. 2). The deduced amino acid sequence of TPL-1 shares a 65% identity to tachylectin-1 (TL-1) (25), identified in the large granules of amoebocytes of the horseshoe crab, and 66% identity to tachylectin-P (TL-P) (24), an embryonic lectin in perivitelline fluid of the horseshoe crab (Fig. 3). A notable difference is the presence of a potential *N*-glycosylation site Asn<sup>74</sup>-Gly<sup>75</sup>-Ser<sup>76</sup> in TPL-1 and its absence in the other two intracellular proteins, TL-1 and TL-P. Although TL-1 and TL-P share 98% sequence homologies with each other, they manifest different biological and biochemical characteristics, in hemagglutinating activity, antibacterial activity, and affinity to other

endogenous proteins (24). TPL-1 also shows a 30–65% identity to tectonin I and tectonin II of myxomycete (Fig. 3) whose function is as yet not known (26). A sequence homology search showed no significant similarity between TPL-1 and any other proteins besides TL-1, TL-P, tectonin I, and tectonin II, including the galectins.

Assuming 8 of the 9 Cys in TPL-1 are involved in disulfide bond formation (based on sequence homologies and conservation of Cys positions with tachylectin-1 (25)), the calculated molecular mass of TPL-1 is 25,801.9 Da. This value agrees well with the 25,857.5-Da TPL-1 found by mass spectrometry (Fig. 4A). The calculated pI of TPL-1 is 8.04, making it a slightly basic protein.

Mass spectrometry analysis showed one other major TPL-1-species with a molecular mass of 26,699.7 Da, and minor species of 16,595.7, 17,578.6, 49,208.8, 50,45.7, 51,889.4, 52,604.2, and 77,903.2 Da. The mass difference of 842.2 between the 25,857.5-Da TPL-1 and the 26,699.7-Da TPL-1 can be attributed to the presence of two HexNAc and three hexoses on the 26,699.7-Da TPL-1. The minor 16,595.7-Da and the 17,578.6-Da TPL-1 represent proteolytic cleavage products of TPL-1 as reported previously (15).

The combined SDS-PAGE, Western blot analysis, and the mass spectrometry data suggest monomer/dimer relationship between the following species of TPL-1: 24,548.0 Da/49,208.8 Da; 25,857.5 Da/51,889.4 Da; and 26,699.7 Da/52,604.2 Da, respectively. The 77,903.2-Da species corresponds to a trimer of the 25,857.5-Da TPL-1.

**Biochemical Properties of TPL-2**—Upon SDS-PAGE, the purified TPL-2 showed major protein bands with a mass of about 36 kDa and a minor band of about 72 kDa, under both non-reducing (Fig. 1A, lane 2) and reducing conditions (Fig. 1B, lane 2). In Western blot analysis, these protein bands reacted with antiserum raised against the HPLC-purified 36-kDa TPL-2 (Fig. 1D, lane 4, NR; lane 8, R). The plasma samples, before and after passage through the affinity columns, showed protein bands of 72 kDa and higher molecular masses, reacting with anti-TPL-2 serum in the nonreducing SDS-PAGE (Fig. 1D, lane 1, pre-column; lane 2, post-column), and mainly of a 66-kDa protein band in the reducing SDS-PAGE (Fig. 1D, lane 5, pre-column; lane 6, post-column). TPL-1 did not react with anti-TPL-2 serum (Fig. 1D, lanes 3 and 7), affirming the specificity of anti-TPL-2 antibodies.

The deduced amino acid sequence of TPL-2 (Fig. 5) showed a 68% identity with conservation of the 6 Cys positions to tachylectin-3 (Fig. 6). Although a potential *N*-glycosylation site, Asn<sup>3</sup>-Cys<sup>4</sup>-Thr<sup>5</sup> is present in TPL-2, this site is absent in tachylectin-3. Tachylectin-3 is a nonglycosylated intracellular protein isolated from the large granule of the amoebocyte.

In our previous report (15), the amino-terminal residues 1, 2, and 3 were left as blank, while residue 4 was shown as Tyr and residue 6 as Lys. Amino-terminal residue analysis of the GlcNAc eluted TPL-2 showed (residue number in superscript, recovery of PTH-amino acid (picomoles) in parentheses, and absence of PTH-amino acid denoted as X): E<sup>1</sup> (152)-D<sup>2</sup> (90)-X<sup>3</sup>-X<sup>4</sup>-T<sup>5</sup> (70)-X<sup>6</sup>-V<sup>7</sup> (75)-T<sup>8</sup> (61)-D<sup>9</sup> (82)-R<sup>10</sup> (61)-S<sup>11</sup> (36)-L<sup>12</sup> (66)-E<sup>13</sup> (43)-G<sup>14</sup> (56)-K<sup>15</sup> (75)-L<sup>16</sup> (46)-M<sup>17</sup> (40)-K<sup>18</sup> (78)-H<sup>19</sup> (25)-P<sup>20</sup> (40).

Gene sequence analysis (Fig. 5) predicts residue 3 as Asn and both residues 4 and 6 as Cys. The sequence analysis shown above corrects and confirms the earlier sequence analysis of TPL-2 (15). The absence of PTH-amino acid at Asn<sup>3</sup> supports the contention that the *N*-glycosylation site of TPL-2 at this position.

The gene sequence data predicted 128 amino acid residues for TPL-2 (Fig. 5). Assuming 6 of the 7 Cys in TPL-2 are engaged

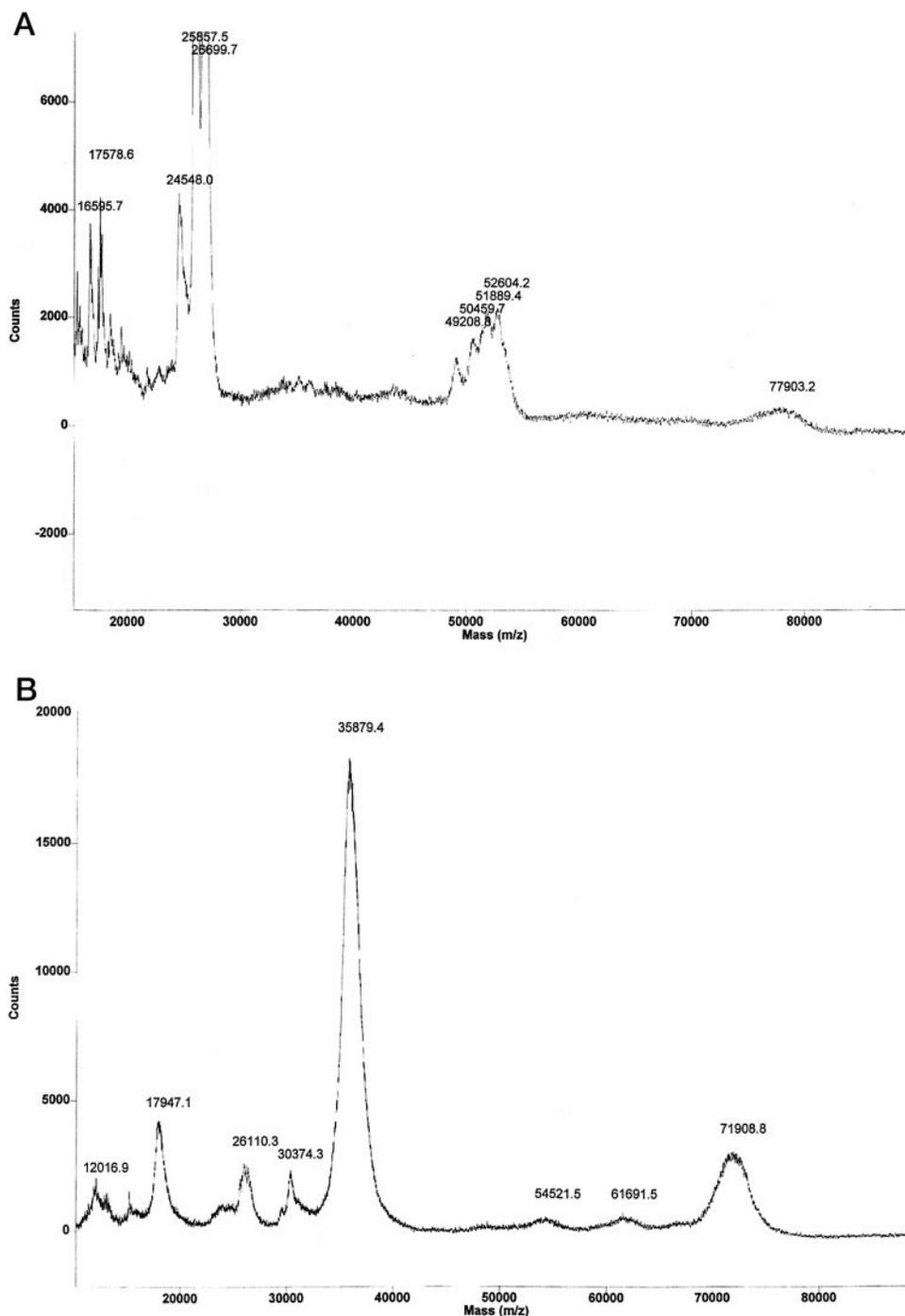


FIG. 4. Mass spectrometric analysis of TPL-1 and TPL-2. The HPLC-purified TPL-1 and TPL-2 eluted from the affinity-columns were subjected to mass spectrometric analysis as described under "Materials and Methods." A, mass spectrometric analysis of TPL-1. B, mass spectrometric analysis of TPL-2.

in disulfide bond formation based on sequence homologies and conservation of Cys positions with tachylectin-3 (21), the calculated molecular mass of TPL-2 will be 14,295.4 Da. Mass spectrometric analysis (Fig. 4B) showed a major TPL-2-species with a mass of 35,879.4 Da, and two minor ones with masses of 17,954.4 and 72,088.0 Da. If the difference of 3,659 Da between the 17,954.4-Da species determined by mass spectrometry and the calculated mass of 14,295.4 Da could be attributed to *N*-glycosylation, a number of possible complex type glycostructures, consisting of about 20 hexoses, can be accommodated. The presence of sugars in TPL-2 was confirmed by periodic acid-Schiff staining (data not shown) and by carbohydrate analysis of the HPLC-purified TPL-2 eluted by GlcNAc. The

molar ratio of hexoses determined in this study is very similar to the one reported earlier (15). Although, without knowing the structure of the carbohydrate, it will be difficult to calculate the exact contribution of carbohydrate to the molecular mass of a glycoprotein, the molar ratio of hexoses shows: Man, 3.1; Gal, 1.9; GlcNAc, 2.4; and GalNAc, 0.4. Setting GalNAc as 1.0, it then follows: Man, 7.8; Gal, 4.8; and GlcNAc, 6.0. This gives rise to a total of 20 hexoses with a calculated molecular mass of 3534, which is close to the difference of 3659 between the 17,954.4-Da species determined by mass spectrometry and the calculated mass of 14,295.4 Da calculated from the deduced amino acid sequence of the TPL-2.

The results of SDS-PAGE, Western blot analysis, and mass

1 CCCCCGGCAGGTTGAAATATTAATAATTACCTAA  
 36 AATATCAGTGTACTAGTACAATTTGATATTTGAATAAAGAAGTAAACGAGTAAGAAG  
 96 AAGTCCCCTCGGTAAATACCTCCGAGGTTTACTTACGTTGTCGATTCACGTTGAGA  
 156 GCTGAGTTAACACAGAACATCTCGTTCAAATGTCGATAAAAAGAAACCCCTTTTGAG  
 216 ATGGGAATTTCAAAGTGGTGTACTATAATACTCGTCCACCTGGTACAGTAAACAGGA  
 -20 *M G I F K V W F T I I L V H L V H V T G*

276 GAAGATACTCCAGTGTGTGACAGACAGTCTCTCGAAGGAAAAGTATGAACACCCCT  
 1 E D N C T C V T D R S L E G K L M K H P

336 TCACCCGACAGTITACCAAAATTTGGATGGGTGCGCGGATGGTGGCCCAATCCCCC  
 21 S T P A V Y Q I L D G C R R L V P N P P

396 ACTTACAACAATCTACAAAACCTGGGAATGATTAATCAATCAATATTTGGAGAACTC  
 41 T Y N N I Y K N W E C I Q S N I L E K L

456 TTGTGTAATGTGATCTCTTTCCACCGTCCAGAACTTATCAAGGGAAGTGGAGATACT  
 61 L C K C D S L S N G A E L I K G S G D T

516 GTATATTTACTAAGTAATGGGTCAAAAGACCTATTCCTGACCCCTGAACITTTAATGGC  
 81 V Y L L S N G V K R P I A D P E T F N G

576 TTTTGTTCGACTGGAACAAATCAAGACTTATTCAGATATGTCATCAACAGCTTTCT  
 101 F C F D W N K I K T Y S D I V I N S L S

636 ACTGGCTATTATAATAATTAAGTAATATTTAAGITACATAAGITTTTATGAACITTT  
 121 T G P I I I I K \*

696 CATTTGAAACAGTCTCTTCGATTTGTTCTTTATAAATTAATAAACCAGITTTATTTTC  
 AAAAAAAAAAAAA

FIG. 5. Nucleotide and deduced amino acid sequences of TPL-2. Nucleotide and amino acid residues are numbered on the left. The underlines represent sequences determined by amino acid sequence analysis of the N terminus of the intact protein. The residues in the gray box indicate corresponding nucleotide sequences used as a primer for PCR. The broken underlines correspond to nucleotide sequences used for cDNA cloning and for multiple gene analysis. The putative signal sequence and glycosylation site are printed in italics and bold, respectively. An asterisk marks the stop codon.

Lib-14p :	-----*	20	-----*	40	-----*
Lib-2p :	-----		-----		-----
Lib-23p :	-----		-----		-----
Lib-15p :	-----		-----		-----
Lib-22p :	-----		-----		-----
TPL-2 :	<b>MGIFKVP</b> <b>IIIVH</b> <b>VEH</b> <b>CGEDN</b> <b>CTCV</b> <b>DRS</b> <b>LEGK</b> <b>LKH</b> <b>PST</b> <b>FAV</b> <b>Q</b> <b>L</b> <b>D</b>				
TL-3 :	<b>MTV</b> <b>N</b> <b>I</b> <b>N</b> <b>L</b> <b>T</b> <b>L</b> <b>S</b> <b>F</b> <b>Y</b> <b>H</b> <b>M</b> <b>H</b> <b>I</b> <b>R</b> <b>R</b> <b>G</b> <b>D</b> <b>C</b> <b>T</b> <b>C</b> <b>V</b> <b>T</b> <b>D</b> <b>N</b> <b>T</b> <b>D</b> <b>T</b> <b>L</b> <b>M</b> <b>K</b> <b>H</b> <b>P</b> <b>S</b> <b>T</b> <b>F</b> <b>A</b> <b>V</b> <b>Q</b> <b>L</b> <b>D</b>				

Lib-14p :	60	-----*	80	-----*	100
Lib-2p :	-----		-----		-----
Lib-23p :	-----		-----		-----
Lib-15p :	-----		-----		-----
Lib-22p :	-----		-----		-----
TPL-2 :	<b>GCR</b> <b>H</b> <b>V</b> <b>P</b> <b>N</b> <b>P</b> <b>P</b> <b>T</b> <b>N</b> <b>N</b> <b>L</b> <b>Y</b> <b>K</b> <b>N</b> <b>W</b> <b>E</b> <b>C</b> <b>I</b> <b>Q</b> <b>S</b> <b>N</b> <b>L</b> <b>E</b> <b>K</b> <b>L</b> <b>L</b> <b>C</b> <b>K</b> <b>C</b> <b>S</b> <b>L</b> <b>S</b> <b>N</b> <b>G</b> <b>A</b> <b>E</b> <b>L</b> <b>I</b> <b>K</b> <b>G</b> <b>S</b> <b>G</b> <b>D</b> <b>T</b>				
TL-3 :	<b>GCR</b> <b>H</b> <b>V</b> <b>P</b> <b>N</b> <b>P</b> <b>P</b> <b>T</b> <b>N</b> <b>N</b> <b>L</b> <b>Y</b> <b>K</b> <b>N</b> <b>W</b> <b>E</b> <b>C</b> <b>I</b> <b>Q</b> <b>S</b> <b>N</b> <b>L</b> <b>E</b> <b>K</b> <b>L</b> <b>L</b> <b>C</b> <b>K</b> <b>C</b> <b>S</b> <b>L</b> <b>S</b> <b>N</b> <b>G</b> <b>A</b> <b>E</b> <b>L</b> <b>I</b> <b>K</b> <b>G</b> <b>S</b> <b>G</b> <b>D</b> <b>T</b>				

Lib-14p :	-----*	120	-----*	140
Lib-2p :	-----		-----	
Lib-23p :	-----		-----	
Lib-15p :	-----		-----	
Lib-22p :	-----		-----	
TPL-2 :	<b>V</b> <b>Y</b> <b>L</b> <b>L</b> <b>S</b> <b>N</b> <b>G</b> <b>V</b> <b>K</b> <b>R</b> <b>P</b> <b>I</b> <b>A</b> <b>D</b> <b>P</b> <b>E</b> <b>T</b> <b>F</b> <b>N</b> <b>G</b> <b>F</b> <b>C</b> <b>D</b> <b>W</b> <b>N</b> <b>K</b> <b>I</b> <b>K</b> <b>T</b> <b>Y</b> <b>S</b> <b>D</b> <b>I</b> <b>V</b> <b>I</b> <b>N</b> <b>S</b> <b>L</b> <b>T</b> <b>G</b> <b>P</b> <b>I</b> <b>I</b> <b>I</b> <b>K</b>			
TL-3 :	<b>V</b> <b>Y</b> <b>L</b> <b>L</b> <b>S</b> <b>N</b> <b>G</b> <b>V</b> <b>K</b> <b>R</b> <b>P</b> <b>I</b> <b>A</b> <b>D</b> <b>P</b> <b>E</b> <b>T</b> <b>F</b> <b>N</b> <b>G</b> <b>F</b> <b>C</b> <b>D</b> <b>W</b> <b>N</b> <b>K</b> <b>I</b> <b>K</b> <b>T</b> <b>Y</b> <b>S</b> <b>D</b> <b>I</b> <b>V</b> <b>I</b> <b>N</b> <b>S</b> <b>L</b> <b>T</b> <b>G</b> <b>P</b> <b>I</b> <b>I</b> <b>I</b> <b>K</b>			

FIG. 6. Alignment of deduced amino acid sequences of TPL-2 related genes found in *T. tridentatus*. The alignment was performed using PILEUP program of GCG package. Divergent residues are shown in white boxes.

spectrometric analysis of TPL-2 strongly suggest that the 17,954.4-Da species represents the monomer, the 35,879.4-Da species the dimer, and the 72,088.0-Da species the tetramer of TPL-2. TPL-2 contains 7 Cys, with a free Cys that could form intermolecular disulfide bond. TPL-2 purified by affinity column exists mainly as a dimer even under denaturing and reducing condition (Fig. 1B, lane 2). In the plasma, TPL-2 and its isoform exist mainly as oligomers of even higher molecular masses (Fig. 1D, lanes 1, 2, 5, and 6).

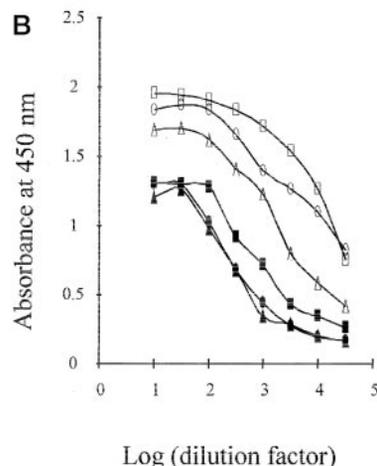
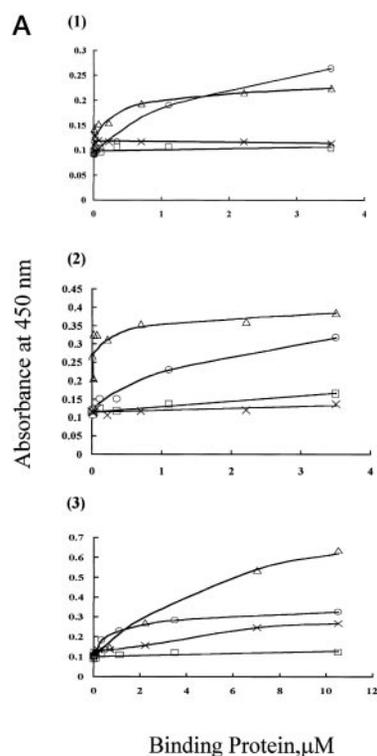


FIG. 7. Binding of TPL-1/TPL-2 and plasma to bacteria. A, TPL-1/TPL-2 incubated with immobilized bacteria: *S. pneumoniae* R36A (1), *V. parahaemolyticus* (2), and *E. coli* Bos-12 (3). Open triangle, TPL-1; open circle, TPL-1; cross, TPL-2 + 0.4 M GlcNAc + 2 M KCl; open square, TPL-1 + 0.4 M GlcNAc. B, pre- or post-column plasma incubated with immobilized bacteria: *S. pneumoniae* R36A (open triangle, pre-column; closed triangle, post-column), *V. parahaemolyticus* (open square, pre-column; closed square, post-column), and *E. coli* Bos12 (open circle, pre-column; closed circle, post-column).

Using primers based on consensus nucleotide sequence between TPL-2 and tachylectin-3, PCR was performed to examine the possible existence of multiple genes for TPL-2-like molecules. Of the 23 clones identified, 5 new genes were found to code for proteins with similar but not identical amino acid sequence to TPL-2 and tachylectin-3 (Fig. 6). The results indicate that a homologous family of multiple genes code for TPL-2 and its isoforms.

TPL-2 binds to LPS from *E. coli*. This binding was the basis for the purification procedure employing LPS-affinity chromatography. The binding of TPL-2 to LPS is apparently independent of  $Ca^{2+}$  ion, since sodium citrate or EDTA was not able to elute TPL-2 from the LPS-affinity matrix. In this respect,

TPL-2 differs from the 12-kDa *Limulus* LPS-binding protein (8) and tachylectin-1 (25).

Among the LPS-binding proteins, the site of interaction with LPS has not been identified, although significant sequence homologies were observed among a number of these proteins (26). TPL-2 does not share any homology with other LPS-binding proteins, including the 12-kDa *Limulus* LPS-binding protein purified by a procedure utilizing LPS-affinity chromatography (8), except TL-3 and TL-P. TPL-2, isolated in this study, differs from most other LPS-binding proteins with a near neutral isoelectric point ( $pI = 7.65$ ), instead of a higher  $pI$  value. The basic nature of these proteins has been considered to be an important factor in their interaction with the negatively charged LPS molecule (27). It could be argued, however, that the proper positioning of the basic amino acid in the three-dimensional structure of the protein is more important than the overall basic nature of the protein for the binding to LPS. In this respect, it is noted that there are three clusters of basic amino acids in the TPL-2 sequence that might be critical for its binding to LPS.

**Heteromers of TPL-1 and TPL-2**—In addition to forming homo-oligomers, TPL-1 and TPL-2 appear to form heteromers with each other with molecular masses of about 76 kDa (Fig. 1E, lane 2). The absence of this band in the two control samples (Fig. 1E, lanes 1 and 3) assures that the 76-kDa band observed in Fig. 1E (lane 2) represents the TPL-1-TPL-2 heteromer. The band with molecular mass of about 180 kDa observed in lanes 1 and 2, but not in lane 3, most likely originated from the anti-TPL-1 antibodies attached to the agarose gel. During boiling in SDS buffer, TPL-1 antibodies were dissociated from the agarose gel and reacted with anti-TPL-2 antiserum in Western blot. Although the mechanism by which stable homo- and hetero-oligomers of TPL-1 and TPL-2 are formed remains to be clarified, the physiological function of TPL-1 and TPL-2 could be related to their propensity to form clusters of interlocking molecules to immobilize and entrap the invading microorganisms.

**Biological Function of TPL-1 and TPL-2: Binding to Bacteria**—TPL-1 and TPL-2 have been shown to bind to three species of bacteria, *S. pneumoniae* R36A, *V. parahaemolyticus*, and *E. coli* Bost-12 in a dose-dependent and saturable manner (Fig. 7A). The specificity of the binding is demonstrated by inhibition with GlcNAc/GlcNAc plus 2 M KCl, respectively. Although both pre- and post-affinity-column plasma proteins bind to bacteria, significantly more pre-column plasma proteins bind to bacteria (Fig. 7B), indicating the contribution of affinity-column purified TPL-1/TPL-2 in binding to bacteria. One possible binding site of TPL-1 and TPL-2 with these bacteria would be at the NAc moiety of the GlcNAc-MurNAc cell wall peptidoglycan. Likewise, lectins with affinity for the *N*-acetyl-group, tachylectin-2 (23), tachylectin-P (24), and tachylectin-5A and -5B (14), could act as innate defense molecules by binding to the peptidoglycans of bacteria.

**Conclusion**—In contrast to adaptive immune system in having repertoires of specific antigen receptors and antibodies, the phylogenetically ancient innate immune system uses germline-encoded receptors for recognition of common antigens on the surface of microbial pathogens, such as proteases (4), polyphe- nol oxidases (28), pathogen-specific lectins (29), and antibiotic peptides (5, 6). They work in concert for the recognition, immobilization, and elimination of the invading pathogens. The pathogen-specific lectins in the hemolymph are expected to distinguish between self and nonself and to serve as the first line of defense upon the entry of pathogens. The interaction between lectins and pathogens results in the recruitment of other defense mechanisms, which ultimately are responsible

for the immobilization and elimination of the invading pathogens.

C-reactive protein, identified as a pattern-recognition molecule in the hemolymph of American horseshoe crabs, *Limulus polyphemus*, is a polymorphic mixture of closely related proteins (30). Recent study has further shown that a family of genes (31) encodes three main classes of the polymorphic CRPs. One (tCRP-1) binds to phosphocholine (PC)/phosphoethanolamine (PEA) ligand in the presence of  $Ca^{2+}$  but not to sialic acid-ligand, another (tCRP-2) binds to both PC/PEA ligand and to the sialic-ligand, and a third (tCRP-3) binds neither to the PC/PEA ligand nor to the sialic acid ligand (31). Yet, they all share an extensive sequence homology, and a hexameric structure (30, 31).

In this study, two additional pattern recognition molecules, TPL-1 and TPL-2 were isolated from *T. tridentatus*, their cDNA sequence determined, and their biochemical properties investigated. The results obtained suggest that, in general, lectin-like pattern recognition molecules: 1) consist of small molecular weight subunit of protein with mass of 15–25 kDa, which are encoded by families of closely related genes; 2) tend to form homo- or hetero-oligomers; and 3) could assemble to yield a myriad of complex structures with different binding specificity and affinity for ligands, that mimic the diversity of the immunoglobulin system. The innate defense of horseshoe crabs depends on this kind of system to recognize and entrap the varied and everchanging nature of the invading pathogens. The glycostructures of TPL-1 and TPL-2 might be responsible for mediating the formation of stable interlocking cluster of the oligomers, through protein-carbohydrate interactions.

**Acknowledgment**—We thank Dr. Gilbert Jay of the OriGene Technologies Inc., for valuable comments on the biological role of lectins as innate defense molecules and for proofreading the manuscript, and Professor Yuan-Chuan Lee of the Johns Hopkins University for valuable discussion on the possible role of glycostructure in stabilizing the oligomer structure of glycoproteins. We are indebted to Drs. Kay-Hooi Khoo, Po-Huang Liang, Chia-Larn Kwo, and Sheng-Tai Chiou for helpful discussion throughout this study and Bor-Long Huang, Jian-Hong Leu, and Jin-Mei Chen for valuable contributions in the chromatographic procedures and determination of carbohydrate composition.

## REFERENCES

1. Fearon, D. T., and Locksley, R. M. (1996) *Science* **272**, 50–54
2. Lemaire, B., Nicolas, E., Michaut, L., Reichhart, J. M., and Hoffmann, J. A. (1996) *Cell* **86**, 973–983
3. Carroll, M. C., and Prodeus, A. P. (1998) *Curr. Opin. Immunol.* **10**, 36–40
4. Medzhitov, R., and Janeway, C. A., Jr. (1997) *Cell* **91**, 295–298
5. Muta, T., and Iwanaga, S. (1996) *Curr. Opin. Immunol.* **8**, 41–47
6. Iwanaga, S., Kawabata, S., and Muta, T. (1998) *J. Biochem. (Tokyo)* **123**, 1–15
7. Fortes-Dias, C. L., Minetti, C. A. S., Lin, Y., and Liu, T.-Y. (1993) *Comp. Biochem. Physiol.* **105**, 79–85
8. Minetti, C. S. A., Lin, Y., Cislo, T., and Liu, T.-Y. (1991) *J. Biol. Chem.* **266**, 20773–20780
9. Volanakis, J. E., and Kaplan, M. H. (1971) *Proc. Exp. Biol. Med.* **136**, 612–614
10. Robey, F. A., and Liu, T.-Y. (1981) *J. Biol. Chem.* **256**, 969–975
11. Enghild, J. J., Thogersen, I. B., Salvesen, G., Fey, G. H., Figler, N. L., Gonias, S. L., and Pizzo, S. V. (1990) *Biochemistry* **29**, 10070–10080
12. Liu, T., Lin, Y., Cislo, T., Minetti, C. A. S. A., Baba, J. M. K., and Liu, T. Y. (1991) *J. Biol. Chem.* **266**, 14813–14821
13. Kawabata, S., and Iwanaga, S. (1999) *Dev. Comp. Immunol.* **23**, 391–400
14. Gokudan, S., Muta, T., Tsuda, R., Koori, K., Kawahara, T., Seki, N., Mizunoe, Y., Wai, S. N., Iwanaga, S., and Kawabata, S. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 10086–10091
15. Chiou, S. T., Chen, Y. W., Chen, S. C., Chao, C. F., and Liu, T. Y. (2000) *J. Biol. Chem.* **275**, 1630–1634
16. Mürer, E. H., Levin, J., and Holmer, R. (1975) *J. Cell. Physiol.* **86**, 533–542
17. Kapitany, R. A., and Zebrowski, E. J. (1973) *Anal. Biochem.* **56**, 361–369
18. Hong, T. H., Chen, S. T., Tang, T. K., Wang, S. C., and Chang, T. H. (1989) *J. Immunol. Methods* **21**, 151–157
19. Freudenberg, M. A., Fomsgaard, A., Mitov, I., and Calanos, C. (1989) *Infection* **17**, 322–328
20. Maniatis, T., Fritsch, E. F., Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, p. 280, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY
21. Inamori, K., Saito, T., Iwaki, D., Nagira, T., Iwanaga, S., Arisaka, F., and Kawabata, S. (1999) *J. Biol. Chem.* **274**, 3227–3278
22. Porath, J., Janson, J.-C., and Laas, T. (1961) *J. Chromatogr.* **60**, 167–177

23. Okino, S., Kawabata, S., Saito, T., Hirata, M., Takagi, T., and Iwanaga, S. (1995) *J. Biol. Chem.* **270**, 31008–31015
24. Nagai, T., Kawabata, S., Shishikura, F., and Sugita, H. (1999) *J. Biol. Chem.* **274**, 37673–37678
25. Saito, T., Kawabata, S., Hirata, M., and Iwanaga, S. (1995) *J. Biol. Chem.* **270**, 14493–14499
26. Huh, C. G., Aldrich, J., Mottahedeh, J., Kwon, H., Johnson, C., and Marsh, R. (1998) *J. Biol. Chem.* **273**, 6565–6574
27. Schumann, R. R., Leong, S. R., Flaggs, G. W., Gray, P. W., Wright, S. D., Mathison, J. C., Tobias, P. S., and Ulevitch, R. J. (1990) *Science* **249**, 1429–1431
28. Hoffmann, J. A., Reichhart, J. M., and Hetru, C. (1996) *Curr. Opin. Immunol.* **8**, 8–13
29. Medzhitov, R., and Janeway, C. A., Jr. (1998) *Curr. Opin. Immunol.* **10**, 12–15
30. Nguyen, N. Y., Suzuki, A., Cheng, S. M., Zon, G., and Liu, T. Y. (1986) *J. Biol. Chem.* **261**, 10450–10455
31. Iwaki, D., Osaki, T., Mizunoe, Y., Wai, S. N., Iwanaga, S., and Kawabata, S. (1999) *Eur. J. Biochem.* **264**, 314–326