

# Identification and Characterization of a Chitin-binding Protein Purified from Coelomic Fluid of the Lugworm *Arenicola marina* Defining a Novel Protein Sequence Family<sup>\*[5]</sup>

Received for publication, September 19, 2012, and in revised form, October 25, 2012. Published, JBC Papers in Press, October 31, 2012, DOI 10.1074/jbc.M112.420976

Nina Vitashenkova<sup>‡§1</sup>, Jesper Bonnet Moeller<sup>‡1</sup>, Rikke Leth-Larsen<sup>¶1</sup>, Anders Schlosser<sup>‡</sup>, Kit Peiter Lund<sup>‡</sup>, Ida Tornøe<sup>‡</sup>, Lars Vitved<sup>¶1</sup>, Søren Hansen<sup>¶1</sup>, Anthony Willis<sup>||</sup>, Alexandra D. Kharazova<sup>§</sup>, Karsten Skjødt<sup>¶1</sup>, Grith Lykke Sørensen<sup>‡</sup>, and Uffe Holmskov<sup>‡2</sup>

From the <sup>‡</sup>Department of Cardiovascular and Renal Research and <sup>¶1</sup>Department of Cancer and Inflammation, Institute of Molecular Medicine, University of Southern Denmark, J.P. Winsloews Vej 25.3, 5000 Odense C, Odense, Denmark, the <sup>||</sup>Medical Research Council Immunochemistry Unit, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, United Kingdom, and the <sup>§</sup>Department of Cytology and Histology, Faculty of Biology, St. Petersburg State University, St. Petersburg 199034, Russia

**Background:** The immune system of the lugworm depends solely on innate pattern recognition molecules.

**Results:** AML-1 was isolated from the coelomic fluid, cloned and characterized as a pattern recognition receptor that binds chitin.

**Conclusion:** AML-1 is the first polychaete lectin cloned and characterized.

**Significance:** AML-1 represents a novel protein sequence family that may give rise to a new protein structure.

We have isolated a novel type of lectin named *Arenicola marina* lectin-1 (AML-1) from the lugworm *A. marina*. The lectin was purified from the coelomic fluid by affinity chromatography on a GlcNAc-derivatized column and eluted with GlcNAc. On SDS-PAGE, AML-1 showed an apparent molecular mass of 27 and 31 kDa in the reduced state. The N-terminal amino acid sequences were identical in these two bands. In the unreduced state, a complex band pattern was observed with bands from 35 kDa to more than 200 kDa. Two different full-length clones encoding polypeptides of 241 and 243 amino acids, respectively, were isolated from a coelomocyte cDNA library. The two clones, designated AML-1a and AML-1b, were 92% identical at the protein level and represent a novel type of protein sequence family. Purified AML-1 induced agglutination of rabbit erythrocytes, which could be inhibited by *N*-acetylated saccharides. Recombinant AML-1b showed the same band pattern as the native protein, whereas recombinant AML-1a in the reduced state lacked a 27 kDa band. AML-1b bound GlcNAc-derivatized columns and chitin, whereas AML-1a did not bind to these matrices. Immunohistochemical analysis revealed that AML-1 is expressed by coelomocytes in the nephridium and in round cells in the epidermis and in eggs. Moreover, AML-1 expression was up-regulated in response to a parasitic infection. We conclude that AML-1 purified from coelomic fluid is encoded by AML-1b and represents a novel type of protein family that binds acetylated components.

Members of different protein families have through evolution acquired the ability to bind microbial carbohydrates and

thereby mark pathogens for destruction. Molecules with lectin activity are known in all phyla, and it is believed that these pattern recognition receptors (PRRs)<sup>3</sup> play a crucial role in invertebrates that lack the adaptive immune system (1–6). Some lectins isolated from invertebrates show sequence homology to those isolated from vertebrates. Several lectins with carbohydrate-recognizing domains of C-type (calcium-dependent) lectins have been found in tunicates, echinoderms, and arthropods (7–10). Galectins (calcium-independent galactose-specific lectins) have been detected in annelids (11, 12), nematodes (13–15), and sponges (16, 17). Fibrinogen-related domains (FReDs or FRePs) with lectin activity are found in mammalian lectins, such as the ficolins, and in invertebrates (18–20). FReDs with lectin activity have been purified from the ascidian *Halocynthia roretzi* (9), the horseshoe crab *Tachypleus tridentatus* (21), and the snail *Biomphalaria glabrata* (22, 23). The canonical FReD ligand-binding site (S1) in the horseshoe crab lectin tachylectin 5A, FIBCD1, and M-ficolin binds acetylated components in a calcium- and acetate-dependent manner (18). Contrary to this, L-ficolin does not contain a functional S1 acetyl-binding site, but several different binding sites (S2–S4) which were revealed through structural studies of the fibrinogen domains of L-ficolin in complex with various ligands (24). Invertebrate lectins have diverse functions like agglutination of foreign cells (25–27), inhibition of bacterial growth (28–30), and enhancement of phagocytosis, also referred to as opsonization (30, 31).

The polychaete annelid, *Arenicola marina*, is a typical representative of marine sedentary lugworms (32, 33). The annelids cluster together with the molluscs and form the Lophotrochozoa, which are a major lineage of the protostomes (6, 34). *A. marina* has a well developed body cavity (coelom) filled with

\* This research work was supported by the Benzon Foundation.

[5] This article contains supplemental Table 1.

<sup>1</sup> Both authors contributed equally to this work.

<sup>2</sup> To whom correspondence should be addressed. E-mail: uholmkskov@health.sdu.dk.

<sup>3</sup> The abbreviations used are: PRR, pattern recognition receptors; FReDs, fibrinogen-related domains; MIC, minimum inhibitor concentration; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol.

coelomic fluid. The latter contains free cells (coelomocytes) that phagocyte foreign materials, including Gram-positive or Gram-negative bacteria *in vivo* and *in vitro* (35, 36). Antimicrobial peptide arenicins have been purified from the coelomocytes of *A. marina* and extensively characterized (37, 38), whereas more than 100 Toll-like receptor genes have been identified in the annelids *Capitella* and *Helobdella* (39). Moreover, some humoral factors with agglutinating activity toward vertebrate erythrocytes have been detected in the lugworm (40). At present, no lectins have been found and characterized in the coelomic fluid of *A. marina*.

In the present study, we describe the purification, complete primary structure, immunolocalization, and binding specificity of a lectin isolated from coelomic fluid of *A. marina*. The lectin was named AML-1 (*A. marina* lectin-1), and it was found to be an oligomeric protein that binds specifically to *N*-acetylated carbohydrates, including chitin, independently of calcium. Screening of a cDNA library constructed from coelomocyte mRNA revealed the presence of two isoforms encoding the AML-1 lectin. Sequence analysis showed no homology with any known lectins or other known protein families, and AML-1 defines a new group of invertebrate immune defense molecules.

## EXPERIMENTAL PROCEDURES

**Buffers and Reagents**—The following buffers and reagents were used: Tris-buffered saline (TBS) (140 mM NaCl, 10 mM Tris, 2 mM Na<sub>2</sub>SO<sub>4</sub>, pH 7.4); TBS/Ca (TBS containing 5 mM CaCl<sub>2</sub>); TBS/EDTA (TBS containing 20 mM EDTA); TBS/GlcNAc (TBS containing 100 mM *N*-acetyl-D-glucosamine); and TBS/Tw (TBS containing 0.1% (v/v) Tween 20). StrataClean resin was purchased from Stratagene. Mono- and disaccharides, bovine submaxillary mucin, lipopolysaccharides (LPS) from *Escherichia coli* (strain 0127:B8 and strain 055:B5), and *Salmonella minnesota* were purchased from Sigma. GlcNAc and glucose were coupled to divinylsulfone-activated TSK-GEL Toyopearl resin (Tosoh Corp.) at a concentration 200 mg of sugar/ml of gel. The primers used for PCR were purchased from Invitrogen or GE Healthcare.

**Worms**—Lugworms (*A. marina*) were purchased from a local fishing shop. The worms were kept at +4 °C in aired seawater during experiments. Coelomic fluid was collected by carefully reaching the body cavity with a needle and allowing the fluid to drip by gravity. Samples of coelomic fluid were centrifuged for 10 min at 800 × *g* to separate cells, and the supernatant was frozen and stored at −80 °C until use. The pellet containing coelomocytes was washed once with TBS, snap frozen, and stored in liquid nitrogen.

**Affinity Chromatography**—Coelomic fluid was mixed with an equal volume of TBS/Ca and applied to a GlcNAc-TSK column. The column was washed extensively, and bound proteins were eluted initially with TBS/EDTA and then with TBS/GlcNAc. The eluted fractions were analyzed by SDS-PAGE, and those of interest were pooled, quantified by measurement of A<sub>280</sub>, and kept at +4 °C.

**Preparation of Antibodies**—Rabbits were immunized subcutaneously with 25 μg of purified AML-1 in Freund's complete adjuvants. The following monthly boosts were done with the same antigen amount in Freund's incomplete adjuvants, and

antisera were collected 2 weeks after a boost. Antibodies were purified using a 1-ml Protein G column (GE Healthcare).

**Western Blotting**—Electrophoresis was performed on 4–12% polyacrylamide gradient gels using a discontinuous buffer system (Bio-Rad). Samples were reduced by heating at 100 °C for 1 min in sample buffer (1.5% (w/v) SDS, 5% (v/v) glycerol, 0.1 M Tris, pH 8.0) with 60 mM dithiothreitol and alkylated by the addition of iodoacetamide to a final concentration of 140 mM. Non-reduced samples were heated for 1 min in sample buffer containing 2 mM iodoacetamide, and then additional iodoacetamide was added to a final concentration of 140 mM. Protein bands were directly visualized by silver staining or Coomassie Brilliant Blue staining or blotted onto an Amersham Biosciences Hybond™-P polyvinylidene difluoride membrane according to the recommendations of the manufacturer (GE Healthcare). The membranes were blocked for 2 h at room temperature in TBS/Tw containing 5% (w/v) nonfat dry milk (Bio-Rad), followed by overnight incubation at 4 °C with primary antibody (10 μg/ml rabbit anti-AML-1 IgG) in TBS/Tw containing 2.5% (w/v) nonfat dry milk. After three washes in TBS/Tw, the membranes were incubated for 1 h at room temperature in TBS/Tw containing either alkaline phosphatase-coupled goat anti-rabbit IgG (D0314, Dako) diluted 1:2,000 or HRP-coupled goat anti-rabbit IgG (P0448, Dako) diluted 1:10,000. The membranes were washed extensively in TBS/Tw and developed as described by Schlosser *et al.* (41) or by the enhanced chemiluminescence ECL Plus™ kit, as described by the manufacturer (GE Healthcare).

**Two-dimensional SDS-PAGE**—GlcNAc eluate with proteins of interest was mixed with StrataClean resin (10:1, v/v) and centrifuged at 12,000 × *g* for 5 min. Pellets containing StrataClean resin beads with adsorbed proteins were dissolved in sample buffer and subjected to first dimensional electrophoresis on a 4–20% polyacrylamide gradient gel under non-reducing conditions. The lane was excised from the gel and placed in sample buffer containing 60 mM dithiothreitol for 2 h before the reduced protein lane was mounted on top of a 4–20% polyacrylamide gradient gel, and a second dimensional electrophoresis was performed.

**Amino Acid Sequencing**—The AML-1-containing protein fraction was reduced and subjected to SDS-PAGE on a NOVEX precast system using 10% NuPage BisTris gel. Protein bands were blotted onto polyvinylidene difluoride membrane (Immobilon P, Millipore) and sequenced directly or digested with trypsin at 37 °C overnight. Resulting peptides were separated by reverse phase HPLC using a Brownlee C18 column (Applied Biosystems) and subjected to amino acid sequencing. The protein sequencer was an Applied Biosystems Procise 494A protein sequencer.

**Construction of cDNA Library of *A. marina* Coelomocytes**—A pool of coelomocytes from six worms was used to construct a coelomocyte cDNA library. Total RNA and poly(A) RNA were isolated by the use of an Rneasy Midi kit (Qiagen) and Oligotex mRNA kit (Qiagen), respectively. The cDNA was produced with the λ-ZAP Express cDNA synthesis kit (Stratagene) according to the manufacturer's instructions and cloned into λ-ZAP Express vector.

## Characterization of a Lugworm Chitin-binding Protein

**Production of Oligonucleotide Probe**—First strand cDNA synthesis on poly(A) RNA was performed using SuperScript II RNase H and random hexamer primers according to the manufacturer's protocol (Invitrogen). The resulting single-stranded cDNA was amplified by PCR with different combinations of six 17-base degenerated primers (forward-reverse) (supplemental Table 1) designed on the basis of the obtained partial amino acid sequence. The PCR products were separated by agarose gel electrophoresis, purified from the gel by the QIAquick gel extraction kit (Qiagen), and sequenced. The deduced protein sequence from an amplified 150-bp PCR product was found to correspond to the protein sequence found in four AML-1-derived peptides. This PCR product was then labeled with [ $\alpha$ - $^{32}$ P]dCTP by use of the Oligolabeling kit (GE Healthcare) and used as a probe for the screening of the cDNA library.

**Screening of cDNA Library**—Approximately  $3 \times 10^5$  plaques of the constructed cDNA library were plated with *E. coli* strain XL-Blue MRF'. The plates were incubated overnight at 37 °C, and duplicate nitrocellulose lifts were performed. The filters were prehybridized in a solution containing  $6 \times$  SSC,  $5 \times$  Denhardt's solution, 0.5% (w/v) SDS, and 100 mg of denatured salmon sperm DNA/ml overnight at 68 °C and then hybridized overnight at 68 °C with the labeled oligonucleotide probe in the same buffer. The filters were washed for 10 min in  $2 \times$  SSC with 0.1% (w/v) SDS at room temperature, followed by two washes for 30 min each in  $2 \times$  SSC with 0.1% SDS at 68 °C. The air-dried filters were exposed to Kodak Scientific imaging film (Biomax MS) overnight at  $-80$  °C with an intensifying screen. Positive plaques were harvested, and DNA inserts were excised from ZAP Express vector as pBK-CMV phagemid vector and introduced into XOLR strain *E. coli* by *in vivo* excision according to the ZAP Express cDNA kit instructions. The obtained single colonies were isolated, and the plasmids were purified by the Quantum Prep plasmid miniprep kit (Qiagen). The inserts were sequenced in both directions with sequence specific primers (supplemental Table 1).

**Expression of AML-1a and AML-1b in CHO Cells**—Full-length AML-1a and AML-1b cDNA including their respective signal sequences were synthesized and obtained from Genescript. The cDNAs were cloned into the expression vector pcDNA5/FRT/V5-His-TOPO® TA vector according to the manufacturer's recommendations (Invitrogen) and sequenced in their entirety. The produced pcDNA5/FRT vectors were transfected into Flp-In CHO cells (Invitrogen) using jetPEI™ (Polyplus transfection). Stably transfected clones were selected with 800  $\mu$ g/ml hygromycin B (Invitrogen) and cultured in Ham's F-12 with Glutamax™ (Invitrogen) supplemented with 10% (v/v) fetal calf serum before protein expression was verified by SDS-PAGE and Western blotting of the cell culture supernatants.

**Chitin Binding Assay**—Chitin beads (Sigma), GlcNAc-TSK, or glucose-TSK beads were washed three times in TBS/Tw containing 5 mM CaCl<sub>2</sub> and pelleted by centrifugation. 500  $\mu$ l of supernatant from CHO cells expressing either recombinant AML-1a or AML-1b were mixed 1:1 (v/v) with TBS/Tw containing 10 mM CaCl<sub>2</sub> and added to 100  $\mu$ l of pelleted beads in the presence or absence of the potential inhibitors; 10 mM EDTA, 50 mM acetate, 50 mM GlcNAc, 50 mM glucosamine, 50

mM *N*-acetylated alanine or 50 mM alanine. After overnight incubation at 4 °C the pelleted beads were washed three times in TBS/Tw containing 5 mM CaCl<sub>2</sub>. Bound protein was eluted from the beads by boiling the samples in sample buffer, and samples were subsequently analyzed by SDS-PAGE and Western blotting.

**Hemagglutinating Activity**—The hemagglutination assay was performed in round-bottomed 96-well microtiter plates using rabbit, rat, guinea pig, and sheep erythrocytes. 2-Fold serial dilutions of the lectin purified from coelomic fluid in TBS (25  $\mu$ l) were mixed with an equal volume of a 4% suspension of erythrocytes in TBS. After incubation for 1 h at room temperature, the extent of agglutination was examined visually. The assay was also performed in the presence of 5 mM CaCl<sub>2</sub> or 20 mM EDTA.

**Hemagglutination Inhibition Assay**—The highest dilution of the lectin in TBS that still gave visible hemagglutination was used in an inhibition assay. All inhibitors were dissolved in TBS at concentrations 200 mM for mono- and disaccharides and 1 mg/ml for LPS and bovine submaxillary mucin. 2-Fold serial dilutions of inhibitor (25  $\mu$ l) were placed in the wells and mixed with the same volume of the lectin samples. Mixtures were left for 10 min at room temperature before 25  $\mu$ l of 4% rabbit erythrocytes suspension were added to each well. Minimum concentrations of the tested substances that still inhibited hemagglutination were defined as minimum inhibitor concentration (MIC).

**Immunohistochemistry**—4- $\mu$ m sections were cut from neutral-buffered formaldehyde-fixed paraffin-embedded tissue blocks. Sections were mounted at ChemMate Capillary Gap Slides (Dako) dried at 60 °C, deparaffinized, and rehydrated. Antigen retrieval was performed using microwave heating in Target Retrieval Solution (Dako). Antigen retrieval was followed by blocking of endogenous biotin using the Dako biotin-blocking system. The incubation with rabbit anti-AML-1 antibody diluted 1:600 was done for 25 min at room temperature. Immunostaining was automated using the ChemMate HRP/DAB detection kit (K5001, Dako) on a TechMate 1000 instrument (Dako). Immunostaining was followed by brief nuclear counterstaining in Mayer's hematoxylin.

## RESULTS

**Purification of Native AML-1**—Different mono- and disaccharides were coupled to TSK, and batch experiments revealed that GlcNAc-TSK bound several putative lectins from the coelomic fluid both calcium-dependently and calcium-independently. Coelomic fluid diluted with CaCl<sub>2</sub>-containing buffer was applied to GlcNAc-TSK column, and bound proteins were eluted first with EDTA and then with GlcNAc-containing buffers and finally analyzed by SDS-PAGE (data not shown).

In the GlcNAc eluate, two bands with molecular masses of 31 and 27 kDa were seen in the reduced state. In the non-reduced state, two prominent bands with molecular masses of 160 and 108 kDa were seen together with minor bands of 80, 54, and 35 kDa (Fig. 1A). When more protein was loaded onto the gel (in the case of two-dimensional SDS-PAGE), the minor bands became clearly visible, and some additional bands of more than 200 kDa appeared (Fig. 1B).

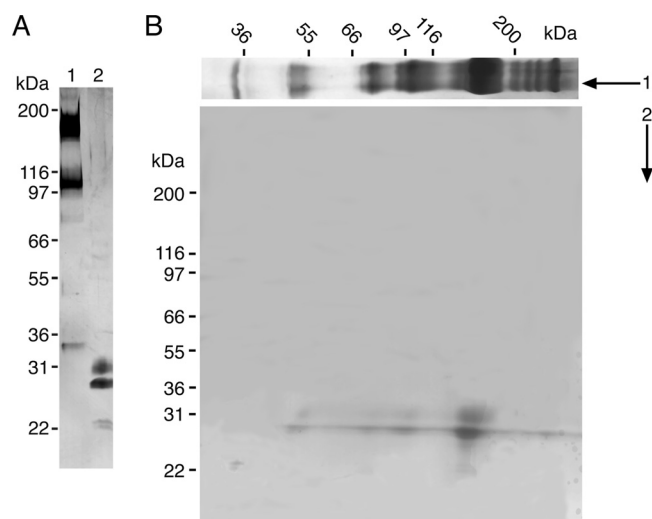


FIGURE 1. **SDS-PAGE of purified AML-1.** *A*, lanes 1 and 2 show AML-1 in non-reduced and reduced state, respectively. *B*, two-dimensional SDS-PAGE of purified AML-1. The first dimension was analyzed under non-reducing conditions. The lane was cut out, incubated in sample buffer containing DTT, reduced, and placed on top of another SDS-polyacrylamide gel and electrophoresed. Both gels were silver-stained.

**TABLE 1**

**Summary of peptide sequences derived by trypsin digestion of AML-1**

Lowercase letters represent amino acids for which determination is not sure. Underlined residues indicate corresponding nucleotide sequence used as a PCR primer.

Peptide	Sequence
Td 1a	ENTLNLQYVYTVEL
Td 2	ASVFNVDVNLSEVNAHFPT <sub>r</sub>
Td 3a, -b, and -c	GGTFK and SQTGIIVADGYR and NSALTK
Td 4	LAVLENTVLLLAEK
Td 5	APPVGPACQPCDGLG <sub>r</sub>
Td 6a and -b	SQTGIIVA and DNLSTV <u>NMCEN</u> V <sub>ear</sub>
Td 7	<u>CEAMQTGSDY</u> TTLDATG <sub>r</sub>
Td 8	RCEAMQTGSDYTTLDATg

All bands were reduced to 31 and 27 kDa with the exception of a 35 kDa band that appeared as a 23 kDa band. It appeared subsequently that the 35 kDa band was a product of proteolysis because its amount markedly increased with time of storage.

**Amino Acid Sequencing of AML-1**—The N-terminal amino acid sequence of the 27 and 31 kDa bands were determined. Both proteins showed the same N-terminal sequence over the first 12 amino acid residues. The protein was eluted from the 27 kDa band and digested with trypsin. Resulting peptides were analyzed using protein sequencing. Twelve readable sequences were obtained with a total of 109 amino acids (Table 1).

**cDNA Cloning and Sequencing**—A  $\lambda$ -ZAP cDNA library of *A. marina* coelomocytes with insert sizes ranging from 0.5 to 1.8 kb was constructed. By screening the library ( $3 \times 10^5$  plaques) with the AML-1-specific probe, 0.5% of all obtained clones were positive. 14 of those were isolated and sequenced. All 14 clones were full-length. 12 of the clones showed identical nucleotide sequences, and this cDNA was named AML-1a. The last two were also identical but differed from AML-1a at 31 positions in the coding region, and this cDNA was named AML-1b. AML-1a was encoded by 994 nucleotides with an open reading frame (ORF) of 723 nucleotides (accession number HQ009860; Fig. 2A). AML-1b was encoded by 1000 nucleotides with an ORF of 729 nucleotides (accession number

HQ009861; Fig. 2B). Except for one insertion/deletion of six nucleotides, all differences were single nucleotide substitutions, and 17 of these were non-synonymous substitutions.

The two cDNAs encode polypeptides of 242 and 244 amino acid residues with calculated molecular masses of 26,259 Da (AML-1a) and 26,469 Da (AML-1b). These masses are in agreement with those obtained by SDS-PAGE (27 kDa protein band). Moreover, they demonstrated 92% identity at the protein level with each other and were in agreement with peptide sequences obtained by direct protein sequencing of the native proteins (Fig. 3). Two potential N-linked glycosylation sites were identified in AML-1b (at Asn-55 and Asn-208), the last one being lost in AML-1a. The isoelectric points of both proteins calculated from amino acid composition were slightly different: 6.80 for AML-1a and 6.59 for AML-1b. Both deduced polypeptides contained hydrophobic leader peptides of 19 amino acid residues and seven cysteine residues.

**Sequence Similarity to Other Proteins**—Nucleotide sequences of AML-1a and AML-1b were subjected to a BLAST search. A predicted protein (XM-001626230.1) identified from the sequencing of the sea anemone *Nematostella vectensis* (42) was found to contain four tandem-arranged domains with homology to AML-1a and AML-1b. Three 10-amino acid serine/proline-rich linker domains separate the binding domains in XM-001626230.1. An alignment of the AML-1a and AML-1b and the four domains (a–d), where we have omitted the serine/proline-rich linker, are shown in Fig. 4. The alignment allowed a prediction of the minimal sequence requirement for the binding domain. The domain starts corresponding to the third cysteine at position 69 in AML-1a and AML-1b, and the minimum domain required for pattern recognition includes 171–175 residues. The domain includes the last four conserved cysteine residues, and the homology between the four sea anemone protein domains and the two AML domains varies from 34 to 41% identity. It is likely that the three first cysteines in AML-1a and AML-1b account for the asymmetrical interchain disulfide bridging of the proteins. Apart from the XM-001626230.1 sequence, no other sequences were identified with homology to AML.

**Hemagglutination and Inhibition Assays**—AML-1 was found to agglutinate rabbit erythrocytes most effectively among various erythrocytes tested. The minimum lectin concentration required for agglutination under these conditions was  $\sim 250$  ng/ml. The ability of AML-1 to agglutinate rat erythrocytes was 4-fold lower, whereas sheep and guinea pig were not agglutinated at lectin concentrations of 2  $\mu$ g/ml. Neither calcium (5 mM) nor EDTA (20 mM) influenced the hemagglutinating activity. We then used rabbit erythrocytes as a model in inhibition assays to reveal the binding specificity of AML-1. Table 2 shows that the most potent inhibitors were N-acetylated monosaccharides, such as N-acetyl-D-mannosamine (MIC = 0.02 mM), N-Acetyl-D-glucosamine, and N-acetyl-D-galactosamine (MIC = 0.19 mM) and N-acetyl-neuraminic acid (MIC = 1.5 mM). No mono- or disaccharides lacking the N-acetyl group showed inhibition potency, even at a concentration of 100 mM. Bovine submaxillary mucin and LPS from *E. coli* strain O127-B4 and *S. minnesota* also demonstrated inhibitory potency.

# Characterization of a Lugworm Chitin-binding Protein

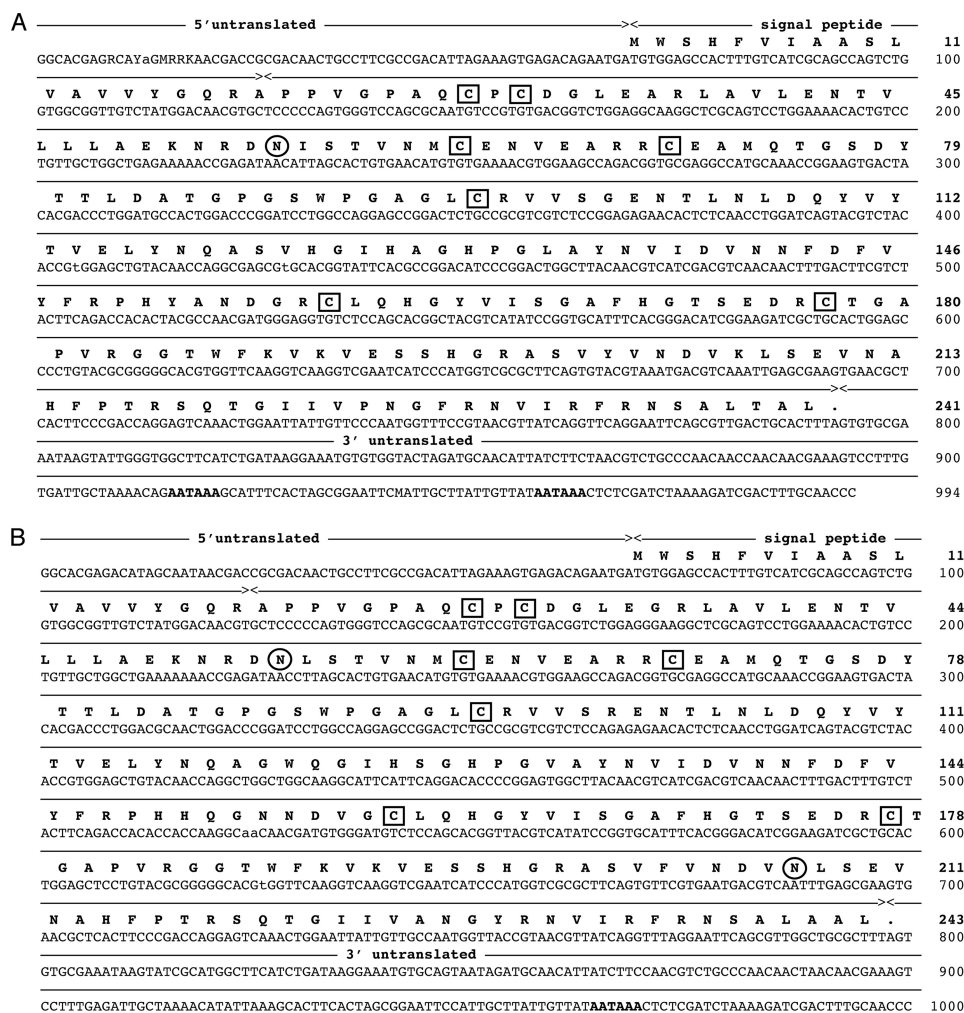


FIGURE 2. Nucleotide sequence and deduced amino acid sequence of AML-1a (A) and AML-1b (B). The initiation methionine is marked as +1. Nucleotides are numbered from the start of the cDNA clones. Potential glycosylation sites are shown in circles, and cysteine residues are shown in boxes. The polyadenylation sites are underlined.

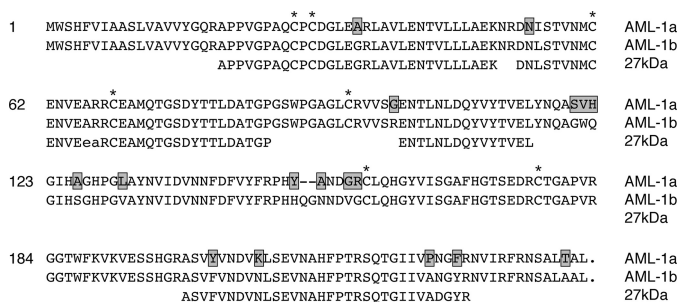


FIGURE 3. Alignment of AML-1a and AML-1b with the peptide sequence obtained from the 27 kDa band from the GlcNAc eluate from the GlcNAc-TSK co1umn. Non-identical residues are boxed.

**Recombinant AML-1a and AML1b**—We expressed full-length recombinant AML-1a and AML-1b in CHO cells (Fig. 5A). The cell supernatants were analyzed by Western blotting using the polyclonal rabbit anti-AML-1 antibody. AML-1b showed two bands with molecular masses of 27 and 31 kDa in the reduced state. No or only faint bands were seen for AML-1a in the reduced state. However, prolonged development revealed a single 31 kDa band in the reduced state (data not shown). In the non-reduced state, three prominent bands with

molecular masses of 160, 108, and 40 kDa were seen for AML-1b, whereas bands of 160 and 75 kDa were seen for AML-1a.

**Binding of Recombinant AML-1a and AML-1b to GlcNAc-TSK and Chitin**—The high similarity between AML-1a and AML-1b could indicate that the two proteins also shared overlapping binding specificity. Pull-down with GlcNAc-TSK and chitin beads revealed, however, a clear difference between the two proteins. Recombinant AML-1b bound efficiently to GlcNAc-TSK and chitin beads, whereas AML-1a failed to bind (Fig. 5, A and B). Because GlcNAc-TSK was used in the initial purification, it is very likely that the initial purification from coelomic fluid only included AML-1b and excluded AML-1a. Chitin is a linear homopolymer of  $\beta$ -1,4-linked *N*-GlcNAc and is an important structural component in the cell wall of most fungi and in the eggshell of parasitic nematodes (43). The binding between AML-1b and chitin was calcium-independent but could be inhibited completely by GlcNAc and *N*-acetylated alanine but not with glucosamine or alanine (Fig. 6). Interestingly sodium acetate partially inhibited the binding.

**Expression and Localization of AML-1**—The specificity of the rabbit anti-AML-1 antibody was analyzed by Western blot-

## Characterization of a Lugworm Chitin-binding Protein



FIGURE 4. Alignment of the predicted AML-1a and AML-1b binding domains with the four homologous domains of the predicted protein XM-001626230.1 from the sea anemone. The alignment was prepared using ClustalW from the DNA Star package.

**TABLE 2**

**Inhibitory effect of several substances on AML-1-induced hemagglutination of rabbit erythrocytes**

Substance	Minimum inhibitory concentration	
	mm	μg/ml
<i>N</i> -Acetyl-D-glucosamine	0.19	
<i>N</i> -Acetyl-D-galactosamine	0.19	
<i>N</i> -Acetyl-D-mannosamine	0.024	
<i>N</i> -Acetyl-D-neuraminic acid	1.5	
D-Glucose	NI <sup>a</sup>	
D-Galactose	NI	
D-Mannose	NI	
L-Fucose	NI	
D-Fucose	NI	
D-Glucose-6-phosphate	NI	
Methyl- $\alpha$ -D-mannopyranoside	NI	
D-Glucosamine	NI	
D-Galactosamine	NI	
D-Mannosamine	NI	
Maltose	NI	
Lactose	NI	
Sucrose	NI	
Bovine submaxillary mucin		7.8
LPS, <i>E. coli</i> 0127:B8		15.6
<i>S. minnesota</i>		62.5
<i>E. coli</i> 055:B5		NI

NI, no inhibitory effect at 100 mm or 250 μg/ml.

ting of purified AML-1 obtained from, coelomic fluid, Triton X-100 extract of coelomocytes, and eggs (Fig. 7). The electrophoretic mobility of AML-1 visualized by Western blotting in these tissues and body fluids showed the identical band pattern. Omitting primary antibody or using IgG from preimmune rabbits gave no specific band. The polyclonal antibody was subsequently applied for immunohistochemical analyses of AML-1 in lugworm tissues. Immunoreactivity was seen in the nephridium, the staining being clearly granular and localized to the apical side of the cells (Fig. 8, A and B). Positive cells with round morphology were scattered throughout the basal part of epidermis (Fig. 8C), and the number of these cells was markedly increased in relation to parasitic infections (Fig. 8D). Coelomocytes of different morphology were clearly positive (Fig. 8, E and F), as were cells in tissues as well as cells found in vessels (Fig. 8G). Eggs showed strong immunoreactivity (Fig. 8, H and I). No

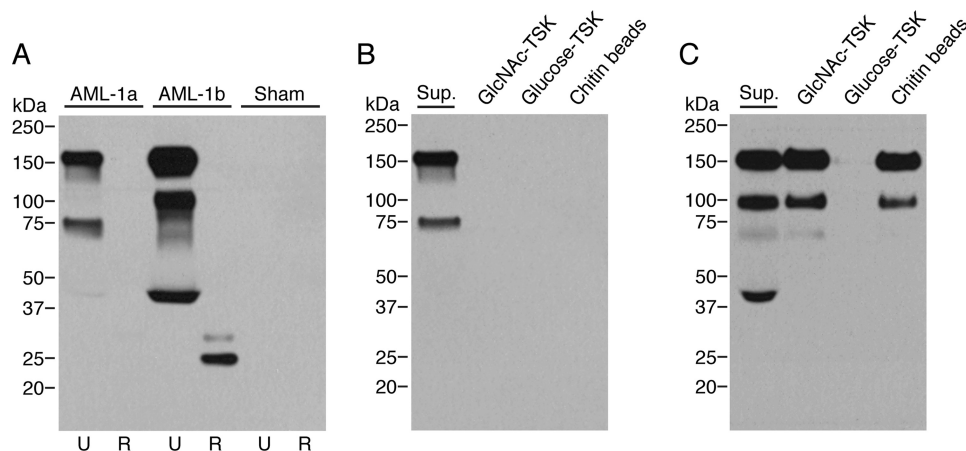
staining was seen when preimmune rabbit IgG was used instead of specific antibody (data not shown).

## DISCUSSION

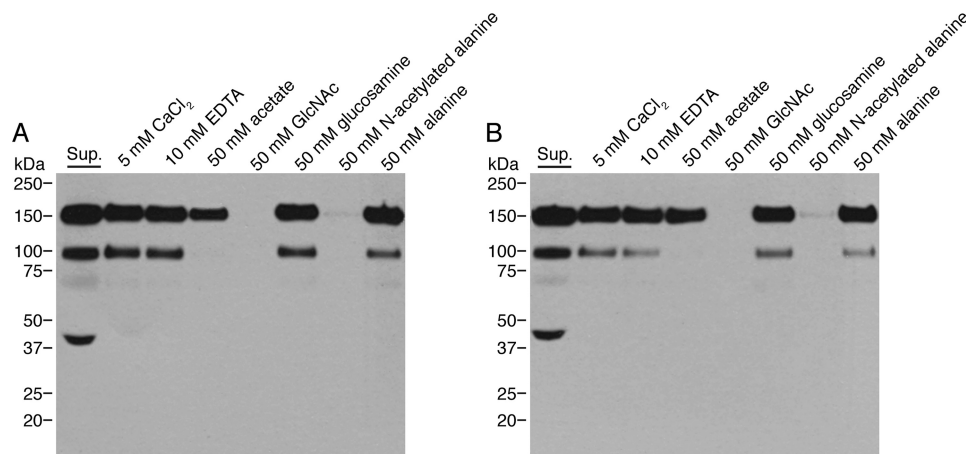
The immune system of invertebrates is exclusively based on PRRs recognizing microbial-derived molecular patterns that mark the microorganisms for elimination (25). Different lectin families act as PRRs in vertebrates and in invertebrates (2, 4, 6, 44, 45). In the present study, we have identified and characterized a lectin from the lugworm *A. marina* that binds acetylated carbohydrates and named it AML-1. AML-1 was purified from coelomic fluid by a single-step affinity chromatography on a GlcNAc-TSK column. SDS-PAGE analysis showed two bands with molecular masses of 27 and 31 kDa in the reduced state and multiple bands in the non-reduced state, indicating that AML-1 might exist as di-, tri-, tetra-, and hexameric molecules and also as more complex structures. Oligomerization like this is a hallmark of PRRs because this increases the binding strength to microorganisms. It may also arrange the PRR so that an appropriate number of binding sites match the spatial arrangement of microbial molecular patterns, leaving endogenous ligands unbound due to alternative spacing.

N-terminal amino acid sequencing of the 27 and 31 kDa bands revealed the same sequence for the first 12 residues. Additional protein sequences were obtained by trypsin digestion and peptide sequencing of peptides from the dominant 27 kDa protein band. Degenerated primers were then synthesized based on the protein sequence, and the resulting PCR products were used to screen a coelomocyte cDNA library. Two different full-length cDNA clones were obtained and designated AML-1a and AML-1b. AML-1a encoded a protein of 242 amino acids, with one *N*-glycosylation site and a calculated molecular mass of 26.3 kDa. AML-1b encoded a protein of 244 amino acids with two *N*-glycosylation sites and a calculated molecular mass of 26.5 kDa. All of the protein sequences obtained from the 27 kDa band were included in the protein encoded by AML-1b, whereas AML-1a differed at six positions. The encoded protein of AML-1a and AML-1b showed 92% identity at the protein level.

## Characterization of a Lugworm Chitin-binding Protein



**FIGURE 5. Western blotting analysis of recombinant AML-1a and AML-1b and their binding to GlcNAc-TSK, glucose-TSK, and chitin beads.** *A*, analysis of supernatants from CHO cells transfected with AML-1a or -1b cDNA. *Lanes 1 and 2*, recombinant AML-1a in the reduced and non-reduced state; *lanes 3 and 4*, recombinant AML-1b in the reduced and non-reduced state; *lanes 5 and 6*, sham-transfected CHO cell supernatant in the reduced and non-reduced state. *B* and *C*, pull-down analysis of supernatants (*Sup.*) containing the recombinant forms of AML-1a and AML-1b, respectively. 100  $\mu$ l of pelleted GlcNAc-TSK, glucose-TSK, or chitin beads were incubated with 500  $\mu$ l of supernatant. After overnight incubation at 4 °C and extensive washing, the bound proteins were eluted by boiling samples in sample buffer and subsequently analyzed by SDS-PAGE (unreduced) and Western blotting using the polyclonal rabbit anti-AML-1 antibody.



**FIGURE 6. Inhibition of binding of recombinant AML-1b to chitin and GlcNAc-TSK beads.** Pull-down analysis of supernatants (*Sup.*) containing the recombinant AML-1b and GlcNAc-TSK (*A*) or chitin (*B*) beads. 500  $\mu$ l of supernatant from CHO cells expressing recombinant AML-1b was mixed 1:1 (v/v) with TBS/Tw containing 10 mM  $\text{CaCl}_2$ , 20 mM EDTA, 100 mM acetate, 100 mM GlcNAc, 100 mM glucosamine, 100 mM *N*-acetylated alanine, or 100 mM alanine and added to 100  $\mu$ l of pelleted beads. After overnight incubation at 4 °C and extensive washing, the bound proteins were eluted by boiling samples in sample buffer and subsequently analyzed by SDS-PAGE (unreduced) and Western blotting using the polyclonal rabbit anti-AML-1 antibody.

No orthologous proteins were found in the genes of other animals, but a BLAST search using the AML-1a and AML-1b protein sequences as bait identified a predicted protein from the sea anemone *N. vectensis* (42) that contained four tandem repeat domains, all showing homology to AML-1a and AML-1b. Thus, the domain of AML-1 can be identified in specimens belonging to a Cnidaria lineage and in one of the major Protostome lineages, Annelida, whereas it appears to be lost during evolution in other Protostome lineages, such as the Nematodes (*Caenorhabditis elegans*) and Arthropods (*Drosophila*) as well as in all Deuterostomes.

Both of the predicted AML-1 proteins contain a 19-amino acid leader peptide, indicating that they can be transported out of the cell. Seven cysteine residues found in both sequences propose intra- and/or interchain disulfide bonds, as seen on SDS-PAGE in the unreduced state. The difference observed in molecular mass between the 27 and 31 kDa bands of AML-1 could be due to differential glycosylation. However, treatment

of the purified lectins with *N*-glycosidase did not reveal any shift in electrophoresis mobility (data not shown), and the difference in predicted molecular mass of the two cDNA clones coding for AML-1 does not explain the observed difference in molecular mass on SDS-PAGE. Different isoforms could exist in different animals because the cDNA library was constructed using mRNA from six lugworms, and the protein was also purified from pools of coelomic fluid. To gain further insights into the structure and functions of AML-1, we expressed recombinant AML-1a and AML-1b in CHO cells. The band pattern of AML-1b was identical to the native AML-1 purified from coelomic fluid. The band pattern of AML-1a indicated that AML-1a also is an oligomeric protein, but it differed from AML-1b both in the reduced and non-reduced state. Due to low expression efficacy, no AML-1a band was initially observed in the reduced state, but prolonged development of the Western blot revealed a single band of 31 kDa corresponding to the upper band seen in AML-1b in the reduced state. These data

show that the band pattern of AML-1b can account for the band pattern observed in AML-1 purified from coelomic fluid.

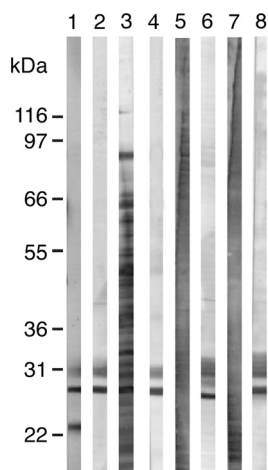
The binding specificity of AML-1 purified from coelomic fluid was examined by inhibition of the AML-1-mediated agglutination of rabbit erythrocytes. All *N*-acetylated carbohydrates were potent inhibitors of agglutination, whereas the corresponding non-acetylated carbohydrates fail to inhibit. Bovine submaxillary mucin, which contains a complex carbohydrate structure, including many acetylated sugars, also inhibited the agglutination (46). Finally, LPS from *E. coli* 0127.B8 and *S. min-*

*nesota* inhibited the hemagglutination activity, whereas LPS from another smooth strain, *E. coli* 055:B5, had no effect on the hemagglutination. Together these findings demonstrated that AML-1 is an acetyl group-specific PRR. C-type lectins like the RegIIIg (or HIP/PAP) molecule (47) or fibrinogen-like molecules like the tachylectins (48) and FIBCD1 (43, 49) exploit acetyl -group recognition for targeting pathogen-associated molecular patterns. The acetyl group recognition of these molecules is often dependent on the presence of calcium ions, but this was not the case for AML-1.

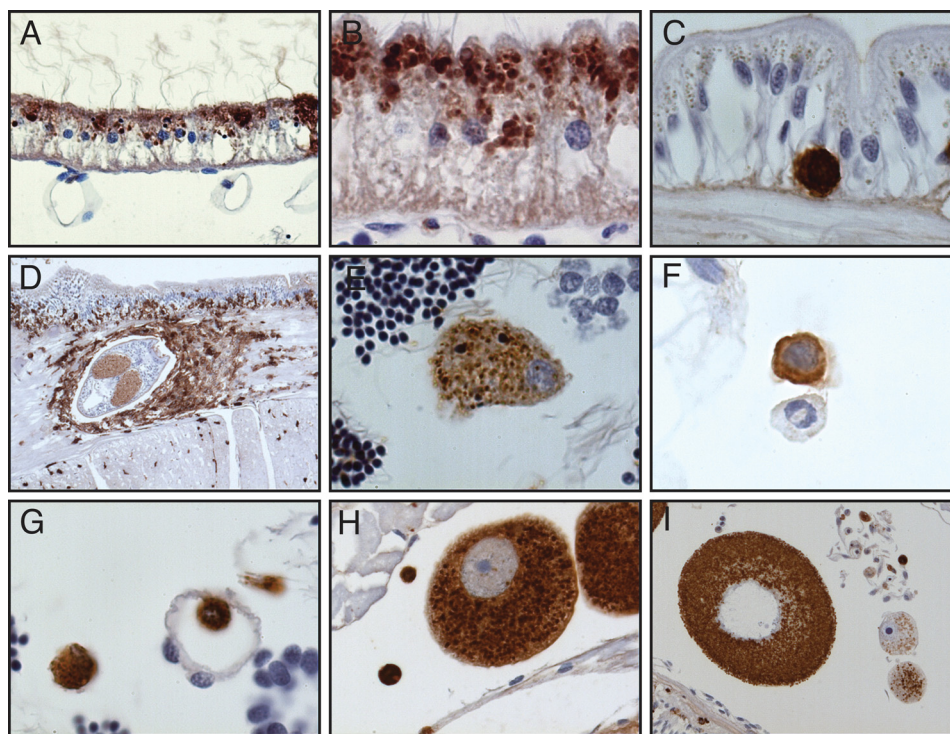
We then compared the binding specificity of recombinant AML-1a and AML-1b using a pull-down assay. AML-1b bound efficiently to GlcNAc-TSK, whereas this was not the case for AML-1a. This further suggests that AML-1 purified by GlcNAc affinity chromatography from coelomic fluid is identical to AML-1b.

We further showed that AML-1b binds to chitin beads. The binding was inhibited by GlcNAc and by other acetylated carbohydrates and by acetylated alanine but not by the corresponding non-acetylated components. In contrast to what is seen for the binding between FReDs and acetylated sugars, the binding was independent of calcium ions, and it was only weakly inhibited by sodium acetate. This indicates that the binding between AML-1b and the acetylated component is dependent on the acetyl group but less dispensable with free acetate as compared with the acetyl group-binding FReDs.

The high homology between the binding domains of AML-1a and AML-1b and the striking difference in binding specificity between the two proteins naturally narrows down the residues involved in the binding activities of AML-1b. The



**FIGURE 7. Analysis of native AML-1 by SDS-PAGE and Western blotting.** Lanes 1 and 2, AML-1 purified on a GlcNAc-TSK column; lanes 3 and 4, coelomic fluid; lanes 5 and 6, Triton X-100 extract of coelomocytes; lanes 7 and 8, Triton X-100 extract of *A. marina* egg. All samples were analyzed in the reduced state. Lanes 1, 3, 5, and 7 are stained with colloidal gold; lanes 2, 4, 6, and 8 were developed using rabbit anti-AML-1 antibody.



**FIGURE 8. Immunohistochemical localization of AML-1 in healthy and infected lugworm tissues.** A and B, nephridium; C, epidermis; D, epidermis infected with a parasite; E and F, coelomocytes; G, tissue with a vessel; H and I, lugworm eggs. The tissues were stained by an indirect immunoperoxidase technique and counterstained with Mayer's hematoxylin as described under "Experimental Procedures." Original magnification was  $\times 10$  (D),  $\times 100$  (A, E, F, G, H, and I), and  $\times 200$  (B and C).



## Characterization of a Lugworm Chitin-binding Protein

second glycosylation site at position 208 in AML-1b and also the stretch of variable amino acids from position 150 to 157 in AML-1b could potentially be involved in the binding activity, but future site-directed mutagenesis analysis will reveal the actual residues involved in loss of function.

Strong AML-1 immunoreactivity was seen in the phagocytic coelomocytes, which play an important role in the lugworm's defense reaction against, for example, bacteria (35, 36). The nephridium, the primitive kidney that connects the coelom with the exterior, showed strong apical immunoreactivity for AML-1. The epidermis containing round lymphoid-like cells presumably representing migrating coelomocytes was also positive for AML-1. The lugworm eggs, which are ejected outside the body cavity and have to survive a hostile microbial environment all the way through fertilization and early development, also showed strong AML-1 immunoreactivity. AML-1 is thus expressed in cells and tissues with a large microbial burden, and this is in agreement with a function as an immunoprotective molecule. Such a role is further supported by the up-regulation of AML-1 expression in tissues surrounding a parasitic infection, located in the subepidermal muscular layer, where migrating coelomocytes are trying to encapsulate the parasite-infested area (Fig. 8D). The antibody used for the immunohistochemical localization was raised against AML-1 purified from coelomic fluid, which is probably identical to AML-1b. The antibody, however, clearly cross-reacted with recombinant AML-1a in Western blotting, and we therefore cannot exclude the possibility that the immunoreactivity that we observe stems from both AML-1a and AML-1b.

We have discovered AML-1 as the first lectin in the lugworm *A. marina*, and, to our knowledge, this is the first polychaete lectin cloned and characterized. The ability of AML-1 to recognize bacterial LPS and chitin and the localization at the nephridium, in epidermis, and in coelomocytes support the notion that AML-1 is involved in defense reaction toward microbial pathogens. The sequence showed homology with four tandem repeat domains found in a predicted protein from the sea anemone, and together these six domains form a novel protein sequence family. Future structural studies will reveal if the sequence shows structural homology to other known domains or represents a novel protein fold.

### REFERENCES

- Iwanaga, S., and Lee, B. L. (2005) Recent advances in the innate immunity of invertebrate animals. *J. Biochem. Mol. Biol.* **38**, 128–150
- Holmskov, U., Thiel, S., and Jensenius, J. C. (2003) Collections and ficolins. Humoral lectins of the innate immune defense. *Annu. Rev. Immunol.* **21**, 547–578
- Weis, W. I., Taylor, M. E., and Drickamer, K. (1998) The C-type lectin superfamily in the immune system. *Immunol. Rev.* **163**, 19–34
- Drickamer, K. (1995) Increasing diversity of animal lectin structures. *Curr. Opin. Struct. Biol.* **5**, 612–616
- Vasta, G. R., Ahmed, H., Fink, N. E., Elola, M. T., Marsh, A. G., Snowden, A., and Odom, E. W. (1994) Animal lectins as self/non-self recognition molecules. Biochemical and genetic approaches to understanding their biological roles and evolution. *Ann. N.Y. Acad. Sci.* **712**, 55–73
- Loker, E. S. (2012) Macroevolutionary Immunology. A role for immunity in the diversification of animal life. *Front. Immunol.* **3**, 25
- Vasta, G. R., Ahmed, H., and Odom, E. W. (2004) Structural and functional diversity of lectin repertoires in invertebrates, protochordates and ectothermic vertebrates. *Curr. Opin. Struct. Biol.* **14**, 617–630
- Poget, S. F., Legge, G. B., Proctor, M. R., Butler, P. J., Bycroft, M., and Williams, R. L. (1999) The structure of a tunicate C-type lectin from *Polyandrocarpa misakiensis* complexed with D-galactose. *J. Mol. Biol.* **290**, 867–879
- Abe, Y., Tokuda, M., Ishimoto, R., Azumi, K., and Yokosawa, H. (1999) A unique primary structure, cDNA cloning and function of a galactose-specific lectin from ascidian plasma. *Eur. J. Biochem.* **261**, 33–39
- Nair, S. V., Pearce, S., Green, P. L., Mahajan, D., Newton, R. A., and Raftos, D. A. (2000) A collectin-like protein from tunicates. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **125**, 279–289
- Hirabayashi, J., Dutta, S. K., and Kasai, K. (1998) Novel galactose-binding proteins in Annelida. Characterization of 29-kDa tandem repeat-type lectins from the earthworm *Lumbricus terrestris*. *J. Biol. Chem.* **273**, 14450–14460
- Ozeki, Y., Tazawa, E., and Matsui, T. (1997) D-Galactoside-specific lectins from the body wall of an echiuroid (*Urechis unicinctus*) and two annelids (*Neanthes japonica* and *Marphysa sanguinea*). *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **118**, 1–6
- Hirabayashi, J., Ubukata, T., and Kasai, K. (1996) Purification and molecular characterization of a novel 16-kDa galectin from the nematode *Caenorhabditis elegans*. *J. Biol. Chem.* **271**, 2497–2505
- Hirabayashi, J., Satoh, M., and Kasai, K. (1992) Evidence that *Caenorhabditis elegans* 32-kDa  $\beta$ -galactoside-binding protein is homologous to vertebrate  $\beta$ -galactoside-binding lectins. cDNA cloning and deduced amino acid sequence. *J. Biol. Chem.* **267**, 15485–15490
- Hirabayashi, J., Satoh, M., Ohyama, Y., and Kasai, K. (1992) Purification and characterization of  $\beta$ -galactoside-binding proteins from *Caenorhabditis elegans*. *J. Biochem.* **111**, 553–555
- Cooper, D. N., Boulianne, R. P., Charlton, S., Farrell, E. M., Sucher, A., and Lu, B. C. (1997) Fungal galectins, sequence and specificity of two isolectins from *Coprinus cinereus*. *J. Biol. Chem.* **272**, 1514–1521
- Hirabayashi, J., and Kasai, K. (1993) The family of metazoan metal-independent  $\beta$ -galactoside-binding lectins. Structure, function and molecular evolution. *Glycobiology* **3**, 297–304
- Thomsen, T., Schlosser, A., Holmskov, U., and Sorensen, G. L. (2011) Ficolins and FIBCD1. Soluble and membrane bound pattern recognition molecules with acetyl group selectivity. *Mol. Immunol.* **48**, 369–381
- Loker, E. S. (2010) Gastropod immunobiology. *Adv. Exp. Med. Biol.* **708**, 17–43
- Lu, J., and Le, Y. (1998) Ficolins and the fibrinogen-like domain. *Immunobiology* **199**, 190–199
- Gokudan, S., Muta, T., Tsuda, R., Koori, K., Kawahara, T., Seki, N., Mizunoe, Y., Wai, S. N., Iwanaga, S., and Kawabata, S. (1999) Horseshoe crab acetyl group-recognizing lectins involved in innate immunity are structurally related to fibrinogen. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10086–10091
- Adema, C. M., Hertel, L. A., and Loker, E. S. (1999) Evidence from two planorbid snails of a complex and dedicated response to digenean (echinostome) infection. *Parasitology* **119**, 395–404
- Adema, C. M., Hertel, L. A., Miller, R. D., and Loker, E. S. (1997) A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8691–8696
- Garlatti, V., Belloy, N., Martin, L., Lacroix, M., Matsushita, M., Endo, Y., Fujita, T., Fontecilla-Camps, J. C., Arlaud, G. J., Thielens, N. M., and Gaboriaud, C. (2007) Structural insights into the innate immune recognition specificities of L- and H-ficolins. *EMBO J.* **26**, 623–633
- Söderhäll, K. (2010) Invertebrate immunity. Preface. *Adv. Exp. Med. Biol.* **708**, vii–ix
- Dam, T. K., Bandyopadhyay, P., Sarkar, M., Ghosal, J., Bhattacharya, A., and Choudhury, A. (1994) Purification and partial characterization of a heparin-binding lectin from the marine clam *Anadara granosa*. *Biochem. Biophys. Res. Commun.* **203**, 36–45
- Saito, T., Kawabata, S., Hirata, M., and Iwanaga, S. (1995) A novel type of limulus lectin-L6. Purification, primary structure, and antibacterial activity. *J. Biol. Chem.* **270**, 14493–14499
- Garte, S. J., and Rissell, C. S. (1976) Isolation and characterization of a hemagglutinin from Amphitrite ornata, a polychaetous annelid. *Biochim. Biophys. Acta* **439**, 368–379

29. Suzuki, T., Takagi, T., Furukohri, T., Kawamura, K., and Nakauchi, M. (1990) A calcium-dependent galactose-binding lectin from the tunicate *Polyandrocarpa misakiensis*. Isolation, characterization, and amino acid sequence. *J. Biol. Chem.* **265**, 1274–1281
30. Tunkijjanukij, S., and Olafsen, J. A. (1998) Sialic acid-binding lectin with antibacterial activity from the horse mussel. Further characterization and immunolocalization. *Dev. Comp. Immunol.* **22**, 139–150
31. Azumi, K., and Yokosawa, H. (1996) Humoral factors and cellular reactions in the biological defence of ascidian *Halocynthia roretzi*. in *New Directions in Invertebrate Immunology* (Soderhall, K., Iwanaga, S., and Vasta, G. R., eds) pp. 43–53, SOS Publications, Fair Haven, NJ
32. Vetvicka, V., Sima, P., Cooper, E. L., Bilej, M., and Roch, P. (1994) *Immunology of Annelids*, CRC Press, Boca Raton, FL
33. Cooper, E. L. (1985) Overview of humoral factors in invertebrates. *Dev. Comp. Immunol.* **9**, 577–583
34. Halanych, K. M., Bacheller, J. D., Aguinaldo, A. M., Liva, S. M., Hillis, D. M., and Lake, J. A. (1995) Evidence from 18S ribosomal DNA that the lophophorates are protostome animals. *Science* **267**, 1641–1643
35. Fitzgerald, S. W., and Ratcliffe, N. A. (1982) Evidence for the presence of subpopulations of *Arenicola marina* coelomocytes identified by their selective response towards Gram+ve and Gram–ve bacteria. *Dev. Comp. Immunol.* **6**, 23–34
36. Fitzgerald, S. W., and Ratcliffe, N. A. (1989) *In vivo* cellular reactions and clearance of bacteria from the coelomic fluid of the marine annelid, *Arenicola marina* l. (Polychaeta). *J. Exp. Zool.* **249**, 293–307
37. Ovchinnikova, T. V., Aleshina, G. M., Balandin, S. V., Krasnosdembkaya, A. D., Markelov, M. L., Frolova, E. I., Leonova, Y. F., Tagaev, A. A., Krasnodembsky, E. G., and Kokryakov, V. N. (2004) Purification and primary structure of two isoforms of arenicin, a novel antimicrobial peptide from marine polychaeta *Arenicola marina*. *FEBS Lett.* **577**, 209–214
38. Andrä, J., Jakovkin, I., Grötzinger, J., Hecht, O., Krasnosdembkaya, A. D., Goldmann, T., Gutschmann, T., and Leippe, M. (2008) Structure and mode of action of the antimicrobial peptide arenicin. *Biochem. J.* **410**, 113–122
39. Davidson, C. R., Best, N. M., Francis, J. W., Cooper, E. L., and Wood, T. C. (2008) Toll-like receptor genes (TLRs) from *Capitella capitata* and *Helobdella robusta* (Annelida). *Dev. Comp. Immunol.* **32**, 608–612
40. Dales, R. P. (1982) Haemoagglutinins and haemolysins in the body fluids of *Neoamphitrite figulus* and *Arenicola marina* (Annelida Polychaeta). *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **73**, 663–667
41. Schlosser, A., Thomsen, T., Shipley, J. M., Hein, P. W., Brasch, F., Tornøe, I., Nielsen, O., Skjødt, K., Palaniyar, N., Steinhilber, W., McCormack, F. X., and Holmskov, U. (2006) Microfibril-associated protein 4 binds to surfactant protein A (SP-A) and colocalizes with SP-A in the extracellular matrix of the lung. *Scand. J. Immunol.* **64**, 104–116
42. Putnam, N. H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., Kapitonov, V. V., Jurka, J., Genikhovich, G., Grigoriev, I. V., Lucas, S. M., Steele, R. E., Finnerty, J. R., Technau, U., Martindale, M. Q., and Rokhsar, D. S. (2007) Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* **317**, 86–94
43. Thomsen, T., Moeller, J. B., Schlosser, A., Sorensen, G. L., Moestrup, S. K., Palaniyar, N., Wallis, R., Mollenhauer, J., and Holmskov, U. (2010) The recognition unit of FIBCD1 organizes into a noncovalently linked tetrameric structure and uses a hydrophobic funnel (S1) for acetyl group recognition. *J. Biol. Chem.* **285**, 1229–1238
44. Messier-Solek, C., Buckley, K. M., and Rast, J. P. (2010) Highly diversified innate receptor systems and new forms of animal immunity. *Semin. Immunol.* **22**, 39–47
45. Vasta, G. R. (2009) Roles of galectins in infection. *Nat. Rev. Microbiol.* **7**, 424–438
46. Tsuji, T., and Osawa, T. (1986) Carbohydrate structures of bovine submaxillary mucin. *Carbohydr. Res.* **151**, 391–402
47. Cash, H. L., Whitham, C. V., Behrendt, C. L., and Hooper, L. V. (2006) Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* **313**, 1126–1130
48. Saito, T., Hatada, M., Iwanaga, S., and Kawabata, S. (1997) A newly identified horseshoe crab lectin with binding specificity to O-antigen of bacterial lipopolysaccharides. *J. Biol. Chem.* **272**, 30703–30708
49. Schlosser, A., Thomsen, T., Moeller, J. B., Nielsen, O., Tornøe, I., Mollenhauer, J., Moestrup, S. K., and Holmskov, U. (2009) Characterization of FIBCD1 as an acetyl group-binding receptor that binds chitin. *J. Immunol.* **183**, 3800–3809