Available online at www.sciencedirect.com
SCIENCE DIRECT
Insect Biochemistry and Molecular Biology 36 (2006) 291–299
Insect Biochemistry and Molecular Biology
www.elsevier.com/locate/ibmb

The tick plasma lectin, Dorin M, is a fibrinogen-related molecule ☆

Ryan O.M. Rego[a], Vojtěch Kovář[a], Petr Kopáček[a], Christoph Weise[b], Petr Man[c], Ivo Šauman[d], Libor Grubhoffer[a,*]

[a]*Institute of Parasitology, Academy of Sciences of the Czech Republic and Faculty of Biological Sciences, University of South Bohemia, Branišovská 31, České Budějovice 370 05, Czech Republic*
[b]*Institute of Chemistry–Biochemistry, Free University Berlin, Thielallee 63, D-14 195 Berlin, Germany*
[c]*Institute of Microbiology, Academy of Sciences of the Czech Republic, Vídeňská 1083, 142 20 Prague 4, Czech Republic*
[d]*Institute of Entomology, Academy of Sciences of the Czech Republic, Branišovská 31, České Budějovice 370 05, Czech Republic*

Abstract

A lectin, named Dorin M, previously isolated and characterized from the hemolymph plasma of the soft tick, *Ornithodoros moubata*, was cloned and sequenced. The immunofluorescence using confocal microscopy revealed that Dorin M is produced in the tick hemocytes. A tryptic cleavage of Dorin M was performed and the resulting peptide fragments were sequenced by Edman degradation and/or mass spectrometry. Two of three internal peptide sequences displayed a significant similarity to the family of fibrinogen-related molecules. Degenerate primers were designed and used for PCR with hemocyte cDNA as a template. The sequence of the whole Dorin M cDNA was completed by the method of RACE. The tissue-specific expression investigated by RT-PCR revealed that Dorin M, in addition to hemocytes, is significantly expressed in salivary glands. The derived amino-acid sequence clearly shows that Dorin M has a fibrinogen-like domain, and exhibited the most significant similarity with tachylectins 5A and 5B from a horseshoe crab, *Tachypleus tridentatus*. In addition, other protein and binding characteristics suggest that Dorin M is closely related to tachylectins-5. Since these lectins have been reported to function as non-self recognizing molecules, we believe that Dorin M may play a similar role in an innate immunity of the tick and, possibly, also in pathogen transmission by this vector.

Keywords: Dorin M; Soft tick; *Ornithodoros moubata*; Hemocyte; Lectin; Fibrinogen-related molecule; Non-self recognition; Arthropod immunity

1. Introduction

Lectins have been shown to be involved in the innate immunity of invertebrates and vertebrates, since they possess the primitive functions of aggregation and micro-organism phagocytosis through opsonisation (Ratcliffe and Rowley, 1987; Vasta, 1991). Evolutionary pressure has allowed lectins to obtain more sophisticated functions in the elimination of pathogens, which is evident from their connection to the complement system (Iwanaga, 2002; Matsushita, 1996; Matsushita et al., 2001; Vasta et al., 1999). It has been reported that the lectin pathway of the complement system can be activated not only by a member of the collectins family, the mannan-binding lectin (Gadjeva et al., 2001), but by a group of proteins, named ficolins as well (Matsushita et al., 2001). A number of proteins belonging to the ficolin family have been identified at the cDNA and/or protein level in both vertebrates and invertebrates. Collectins and ficolins are characterized by the presence of the short N-terminal segment, and both collagen-like and C-terminal carbohydrate-recognition domains (CRD) (Lu and Le, 1998). This carbohydrate recognition domain in the ficolin family is similar to the fibrinogen C-terminal half of the β and γ chains of fibrinogen, hence the term fibrinogen-like domain (Lu, 1997; Matsushita and Fujita, 2001). Although the fibrinogen-like domain is unrelated to CRD (Lu, 1997), human L-ficolin/P35 has been shown, like collectins, to bind to *N*-acetyl-D-glucosamine (GlcNAc) and similar sugar

Abbreviations: OMFREP, *O. moubata* fibrinogen related protein; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; RT, reverse transcription

☆ *Note*: The nucleotide sequence of Dorin M reported in this paper has been deposited in the GenBank database with accession no. AY333989.

*Corresponding author. Tel.: +420 387 775 456; fax: +420 385 310 388.
E-mail address: liborex@paru.cas.cz (L. Grubhoffer).

0965-1748/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ibmb.2006.01.008

structures expressed on bacterial surfaces (Le et al., 1998; Matsushita et al., 1996).

The molecular factors involved in invertebrate innate immunity, and their evolutionary development in vertebrate innate immunity, have not yet been studied in full. However, protein sequence similarities that have been revealed point to a significant evolutionary conservation within their effector proteins including lectins. In the solitary ascidian *Halocynthia roretzi*, a protochordate that may be considered a phylogenetic intermediate between vertebrates and invertebrates, molecules associated with the lectin complement pathway have been cloned and sequenced. This includes a group of four ficolins that exhibit high homology to mammalian ficolins (Kenjo et al., 2001).

Tachylectins 5A and 5B (Tls-5), isolated and cloned from different tissues of the horseshoe crab, *Tachypleus tridentatus*, seem to play a role in its innate immunity. Tls-5 have a fibrinogen-like domain but lack a collagen-like domain (Gokudan et al., 1999). They show significant sequence similarity to mammalian ficolins at the protein level and, therefore, their participation in the activation of the horseshoe crab complement-like system is assumed (Iwanaga, 2002).

The soft tick, *Ornithodoros moubata*, is recognized as a species of great importance in transmitting *Borrelia duttoni*, a causative agent of relapsing fever, and African swine fever virus. It is generally accepted that lectins are involved in self/non-self recognition and, therefore, might be involved in the transmission of pathogens by the vector (Grubhoffer et al., 1997; Maudlin and Welburn, 1988). In our laboratory, we have purified from this species a sialic-acid binding lectin named Dorin M which is abundantly present in the hemolymph of the tick (Grubhoffer and Kovář, 1998; Kovář et al., 2000). In this study, we report the primary structure of this molecule, revealing that Dorin M is a fibrinogen-related lectin likely playing a role as a pattern recognition molecule.

2. Materials and methods

2.1. Animals

O. moubata ticks were obtained from the laboratory colonies maintained at the Institute of Parasitology ASCR. Nymphs and adults were fed through a Parafilm® membrane on whole citrated bovine blood until repletion. Fed ticks were dried on filter paper until coxal fluid excretion was completed, and maintained in polyethylene cages as separate instars in a climatized room (25–27 °C, 80–90% RH).

2.2. Immunocytochemistry

Hemolymph was obtained from ten ticks after excising the legs at the coxae level using a glass micropipette. It was then placed on SuperFrost Plus microscope slides, and hemocytes were fixed in 4% paraformaldehyde in 0.1 M sodium phosphate buffer, pH 7.5 for 10 min at room temperature (RT). Following washing in distilled water and phosphate-buffered saline supplemented with 0.3% Tween 20 (PBS-Tw), the slides were blocked with 10% normal goat serum diluted in PBS-Tw for 30 min at RT and incubated with rabbit anti-Dorin M antibody (Kovář et al., 2000) (diluted 1:500 in PBS-Tw) in a humidified chamber overnight at 4 °C. After thorough rinsing with PBS-Tw (three times for 10 min at RT), the hemocytes were incubated with goat anti-rabbit IgG-Cy3-conjugated secondary antibody (Jackson ImmunoResearch; 1:800 in PBS-Tw; 1 h at RT). After rinsing with PBS-Tw (three times for 10 min at RT) the slides were dehydrated and mounted in DPX mounting medium (Fluka). The mounted specimens were viewed and photographed under an Axioplan II microscope (Carl Zeiss) equipped with epifluorescence and CCD camera.

2.3. Fragmentation and sequencing of Dorin M

Dorin M was purified from hemolymph plasma using chromatography on Blue-Sepharose followed by anion exchange FPLC on a MonoQ column, as described previously (Kovář et al., 2000). Pure lectin (10 μg) was dissolved in 100 μl of 0.1 M NH_4CO_3 and digested with 2 μg of trypsin, modified, sequencing grade (Boehringer-Mannheim) at 37 °C for 18 h. The resulting fragments were separated by reverse-phase HPLC using 4.6 × 250 mm Vydac 300, C18 column (Grom, Germany) with a Waters 626 pump and Waters 490E detector. The solvent system used was 0.1% (v/v) trifluoroacetic acid in water (aqueous phase) and 0.1% trifluoroacetic acid in acetonitrile:water, 80:20 (organic phase-B). After an initial lag phase of 10 min at 2% B, an 80 min linear gradient of 2–60% B was run at a flow rate of 0.5 ml/min. Fractions were collected manually according to their absorption at 215 nm, and their purity and mass were determined using the electrospray mass spectrometer, model LCQ (Finnigan-MAT). Selected peptides were sequenced using a type 473 A protein sequencer (Applied Biosystems), and their mass and sequence were confirmed using a Bruker Reflex mass spectrometer (MALDI-MS with sinapic acid as matrix).

2.4. In gel digestion and peptide mass fingerprinting

Coomassie stained/destained band of Dorin M was excised from the polyacrylamide gel, cut to small pieces and washed extensively in distilled water. The "in gel" digestion was carried out with trypsin (Sequencing grade, Roche) in a cleavage buffer consisting of 0.05 M 4-ethylmorpholine acetate pH 8.2 and 10% acetonitrile, overnight at 37 °C. Mass spectra were measured on a matrix-assisted laser desorption/ionization reflectron time-of-flight mass spectrometer (BIFLEX II, Bruker–Franzen) equipped with a nitrogen laser (337 nm) and griddless delayed-extraction ion source. A saturated solution of α-cyano-4-hydroxy-

cinnamic acid in 50% acetonitrile/0.2% TFA was used as a matrix. Theoretical masses of proteolytic fragments were calculated from the sequence using PeptideMass at http://www.expasy.ch/tools/, and the recorded masses were matched with these values.

2.5. Total RNA extraction and first strand cDNA synthesis

Total RNA was extracted using 0.5 ml Trizol® Reagent (Gibco BRL) from hemocytes obtained from 0.5 ml of *O. moubata* hemolymph by centrifugation (2000 rpm, 10 min, 4 °C). The sample was kept overnight at −75 °C, thawed, homogenized in the microtube and, in further, the total RNA isolation followed the manufacturer's instruction. First strand cDNA was synthesized using a SuperScript™ -II reverse transcriptase (Gibco BRL) and oligonucleotide $(T)_{17}$ following the manufacturer's protocol. Of the resulting cDNA, 1/2 μL was used as a template for PCR reactions.

2.6. PCR and rapid amplification of cDNA ends (5′- and 3′-RACE)

Two Dorin M-specific 312 and 714 bp PCR fragments were amplified using a combination of one degenerated sense primer, 5′-GAG GC(T/C/A/G) TT(T/C) GCA AA(T/C) GT(T/C/A/G) GA(A/G) TGG-3′ and two degenerate anti-sense primers 5′-CAC (A/G)TT ATT TCC TA(A/G) CCA (A/G)TA (T/C)TC-3′ , 5′-CTT CAT CTC AAC (A/G)TT (T/C/A/G)GG (A/G)TA-3′ derived from amino-acid sequences of internal tryptic peptides EAFANVEW, EYWLGNNV and YPNVEMK, respectively (see below). PCR was performed with a Thermocycler T3 (Biometra) using TaKaRa Ex Taq™ HS polymerase (TaKaRa BIOMEDICALS). The PCR program included an initial denaturation step of 2 min at 94 °C, and then 25 cycles were run as follows: 92 °C for 20 s, 55 °C for 30 s (this annealing temperature was decreased by 0.5 °C in each subsequent cycle) and 74 °C for 30 s. Additional three amplification cycles with denaturation at 92 °C for 20 s, annealing at 48 °C for 30 s, extension at 74 °C for 30 s and final extension step at 74 °C for 2 min.

Using the obtained sequence, 3′ and 5′ RACE-PCR were performed to complete the cDNA sequence. The sequences of 3′- and 5′ end portions of the Dorin M gene were determined by a modified protocol for SMART™ cDNA Library Construction Kit (CLONTECH Laboratories, Inc.). Briefly, 5 μg of total RNA isolated from hemocytes were used for synthesis of the first strand cDNA using a universal adaptor primer and PowerScript™ reverse transcriptase, according to the manufacturer's instruction.

Two 3′-RACE products of 481 and 333 bp spanning across the unknown 3′ end were amplified using PCR reaction performed with Advantage® 2 PCR Kit CDS III/3′ PCR Primer (Clontech) and gene-specific sense primers 5′-CAA AGT TGC GAG CGA AGA AGA ATA-3′ and 5′-GAA TTT AAA GGC GCC TGG TGG TA-3′, respectively. The PCR program included an initial denaturation step of 3 min at 94 °C, and then 30 cycles were run as follows: 94 °C for 45 s, 52 ± 5 °C for 45 s and 72 °C for 1 min, with a final extension of 7 min at 72 °C.

The 5′ part of Dorin M was amplified using Advantage® 2 PCR Kit 5′ PCR Primer (Clontech) and two anti-sense gene-specific primers: 5′-CTC CGA AAC CGT GTG CAT AGT CAG-3′and 5′-GAA CTG TCC CCG CCT CTG AAA T-3′ derived from the internal sequence of Dorin M. Two 5′ RACE PCR products of 446 and 388 bp, respectively, were obtained under the above-mentioned conditions.

2.7. Cloning and sequencing

The obtained DNA fragments were ligated into a pCR® 2.1-TOPO® Vector (Invitrogen). Following transformation into One Shot® TOP10 chemically competent *E. coli* cells, recombinants were identified and plasmid DNA recovered using the Plasmid Miniprep Kit (QIAGEN). The inserts were sequenced in both directions using M13 Forward and M13 Reverse primers. The nucleotide sequence of this fragment was analyzed using an automated DNA sequencer model CEQ 200 (Beckman Coulter) and CEQ Dye Terminator Sequencing Kit (Beckman Coulter). At least three different clones of each product were sequenced to assure high fidelity of the obtained sequence data.

2.8. Database analysis

All sequences obtained were checked for accuracy using the SeqMan program (DNASTAR). The complete stretch of cDNA sequence was obtained by overlapping the sequences of the 3′ and 5′ RACE-PCR products using this program. The sequences obtained, both nucleotide and deduced amino-acid, were compared with those in the GenBank database using the NCBI BLAST search program (http://www.ncbi.nlm.nih.gov/blast/) to search for similarity. Conserved domains were box-shaded with the help of BOXSHADE program available at http://www.ch.embnet.org. The theoretical molecular mass and isoelectric point was counted using ExPASy pI/Mw tool (http://www.expasy.org/tools/pi_tool.html), and the putative signal sequence was predicted using the SignalP V1.1 server (http://www.cbs.dtu.dk/services/SignalP/) (Nielsen et al., 1997).

2.9. RT-PCR analysis of Dorin M expression

Total RNA was extracted from salivary glands (3.5 μg), midgut (2 μg), malphigian tubules (4 μg), ovary (4 μg) and hemocytes (2 μg), and reverse transcribed into first strand cDNA, as described above. The presence of Dorin M was determined by PCR amplification of 504 bp using gene-specific sense and anti-sense primers: 5′ AAG GAC TCC CTA TTA CC 3′ and 5′ ATA CGA GGG ATG TTC G 3′,

respectively. PCR was performed using HotStar Taq polymerase (QIAGEN), and included an initial denaturation step of 15 min at 95 °C followed by 30 cycles of 94 °C for 1 min, 52.5 °C for 1 min and 72 °C for 2 min, with a final extension of 8 min at 72 °C. Mitochondrial cytochrome b (GenBank AB073679) was used as an internal standard, and a product of 416 bp was amplified using sense and antisense primers: 5′ TGA GGG GCA ACA GTA AT 3′ and 5′ GCA TAA GCA AAT AAA AAG TA 3′, respectively, using the identical protocol except for the annealing temperature being 47.5 °C.

3. Results

In our preliminary immunofluorescence experiments, rabbit polyclonal antibodies raised against the isolated Dorin M (Kovář et al., 2000) recognized this molecule to be associated with tick hemocytes. It was not clear, however, if the molecule is just bound to the surface of the hemocytes or if the plasma lectin is expressed in the hemocytes. Immunoflourescence confocal microscopy (Figs. 1(A) and (B)) clearly demonstrated that Dorin M is present in the cytoplasm of the hemocytes. This result also determined the tick hemocytes to be used as the starting material for RNA isolation and cDNA synthesis.

The N-terminal sequence of Dorin M was apparently blocked (Kovář et al., 2000). In order to obtain any sequence information from the isolated molecule, 10 μg of Dorin M was cleaved by trypsin and separated using a C18-reverse phase HPLC (Fig. 2). Selected fractions were submitted to Edman sequencing and/or de novo sequence analysis by mass spectrometry. Three internal amino-acid sequences, viz. **GGSTGLYYYSYPNVEMK, XXXXX$_{(540)}$VGEAFANVEWSVAELKR and EYWLGNNVLHALTSDK** were determined for fractions 22, 27 and 29, respectively. The Blast search revealed that the peptides 22 and 29 display a significant homology to conserved motifs present in fibrinogen-related lectins, especially tachylectins 5A and 5B from the horseshoe crab (Gokudan et al., 1999). Based on these fragments, and assuming that the peptide sequence 27 is located within a non-conserved N-terminal part of the molecule, the degenerate primers were designed and used for PCR amplification as outlined in Material and methods.

The entire cDNA sequence (Fig. 3) was obtained by overlapping six PCR and RACE PCR products. Dorin M cDNA consists of 1055 bases with an ORF of 837 bases encoding a 279 amino-acid protein (submitted to GenBank accession no. AY333989). The signal sequence (shaded in Fig. 3) was identified with the putative cleavage in between Ala$_{18}$ and Gln$_{19}$. This result was also confirmed experimentally. The mass of 5 amino-acid residues of fragment 27 was determined to be 540 Da which matches exactly the deduced amino-acid sequence (QDPTD) in which the N-terminal glutamine is modified to pyroglutamate. This was most likely the reason why the N-terminus of the isolated Dorin-M was blocked for Edman degradation (Kovář

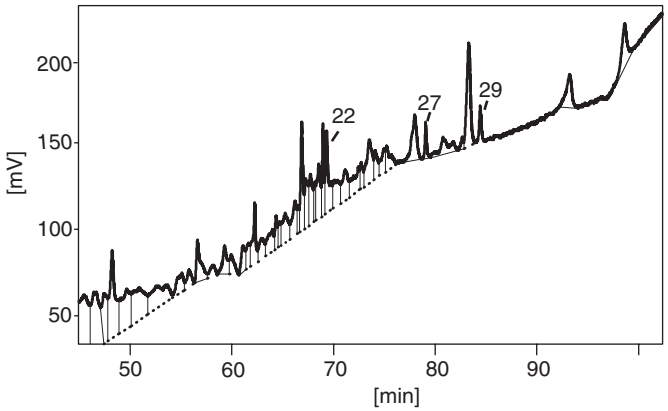

Fig. 2. Reverse phase HPLC separation of tryptic peptides. Isolated Dorin M (10 μg) was digested with trypsin, and the resulting peptides were separated by C_{18} reverse-phase HPLC. The amino-acid sequences of fragments 22, 27 and 29 were determined using Edman degradation and/or mass spectrometry.

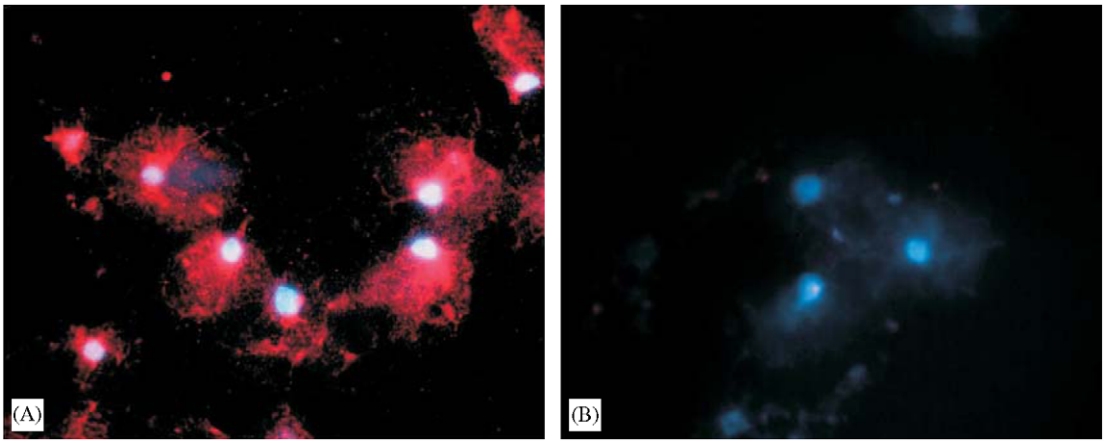

Fig. 1. Immunofluorescent detection of Dorin M in the hemocytes of *O. moubata* using confocal microscopy. Panel A—polyclonal rabbit anti-Dorin M antibody was used to detect the molecule (red); the hemocytes were counterstained with DAPI to visualize the location of nuclei (blue). Panel B—negative control staining in which the primary antibody was replaced with normal goat serum. DAPI staining depicts the hemocyte nuclei (blue).

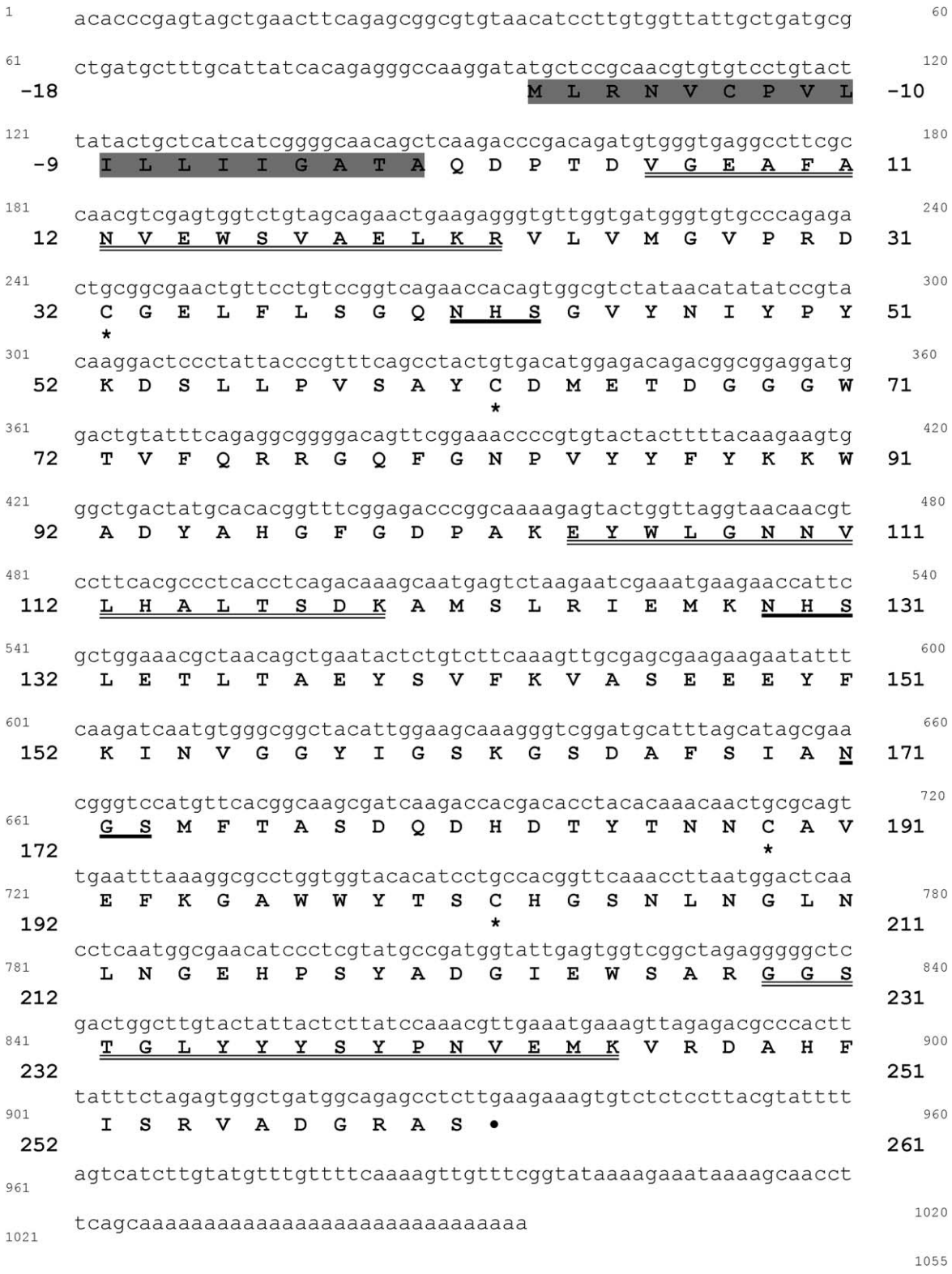

Fig. 3. Nucleotide and deduced amino-acid sequence of Dorin M. Dorin M consists of 1055 bases with an ORF of 837 bases encoding a 279 amino-acid protein. The shaded area symbolizes the signal sequence with cleavage in between Ala_{18} and Gln_{19}. The three amino-acid sequences of internal tryptic fragments are double-underlined. Bold-underlined are the three putative glycosylation sites. Asterisks indicate the cysteins that are likely involved in the intrachain disulfide bonds as predicted according to the homology with tachylectins 5A, 5B [14]. The stop codon is indicated by the dot (●).

et al., 2000). The calculated mass of the mature protein is 29,114.16 Da, and the theoretical isoelectric point is 5.13. Three putative glycosylation sites (underlined in Fig. 3) were identified within the amino-acid sequence of the mature Dorin M (N_{41}–H_{42}–S_{43}, N_{129}–H_{130}–S_{131} and N_{171}–G_{172}–S_{173}).

The same approach led to the cloning and sequencing of a similar but clearly different fibrinogen-related molecule from a tick hemocyte cDNA further referred to as OMFREP (for *O. moubata* fibrinogen-related protein, accession no. AF527411, unpublished). The multiple sequence alignment of Dorin M, OMFREP, horseshoe crab tachylectin 5A and 5B and human ficolin-1 (Fig. 4) showed that the sequence identity between Dorin M and OMFREP is about 65%, whereas the similarity to other molecules is significantly lower: Tl-5A (40%), Tl-5B (33%) human ficolin-1 (33%). Like the tachylectins, Dorin M and OMFREP lack the glycine-rich, collagen-like domain in the N-terminal portion of the ficolins (Fig. 4).

In order to ascertain if Dorin M and OMFREP are both present in the lectin isolated from the tick plasma, the mass-fingerprint analysis of three different lectin preparations (one from the affinity purification and two from FPLC-chromatography procedure (Kovář et al., 2000)) was performed. In all cases only the peptide masses corresponding to the theoretical fragments of Dorin M were detected in the mass-spectra of the tryptic digest (Fig. 5). These data clearly indicate that Dorin M is the only and/or pre-dominant form of the isolated plasma lectin.

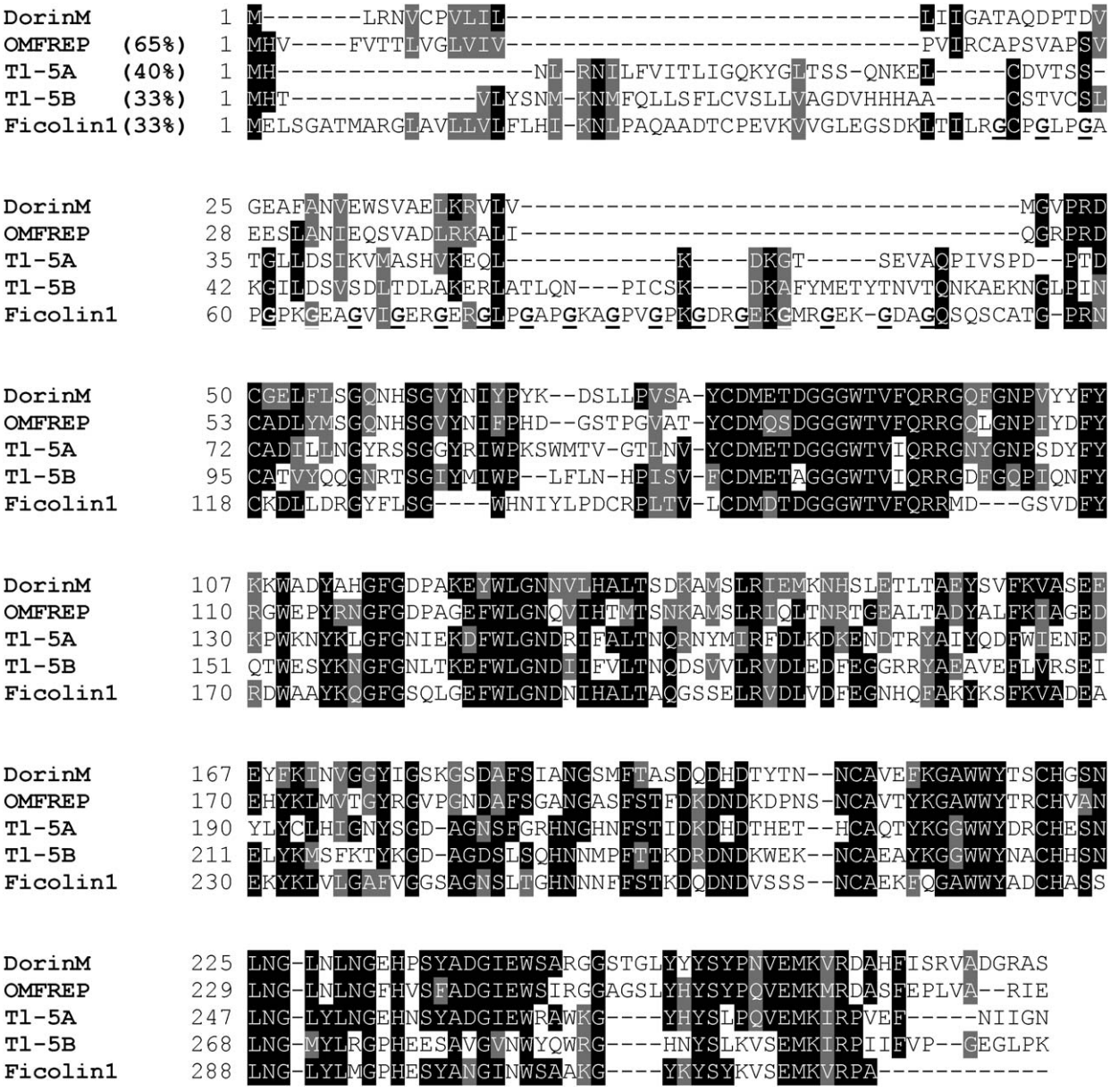

Fig. 4. Multiple sequence alignment of the Dorin M precursor with OMFREP, TLs-5 and human ficolin 1. The numbers given in parentheses show the percentage of residues identical to Dorin M. The bold-underlined glycine residues in the human ficolin 1 sequence indicate the collagen-like domain missing in Dorin M, OMFREP and Tl-5A, 5B. Dorin M—*Ornithodoros moubata* plasma lectin, this work (GenBank accession no. AY333989); OMFREP—*O. moubata* fibrinogen-related protein (GenBank accession no. AF527411); Tl-5A, 5B—*Tachypleus tridentatus* tachylectin 5A, 5B (GenBank accession nos. AB024737, AB0247380), respectively; Ficolin1—*Homo sapiens* ficolin 1 (GenBank accession no. NM_002003).

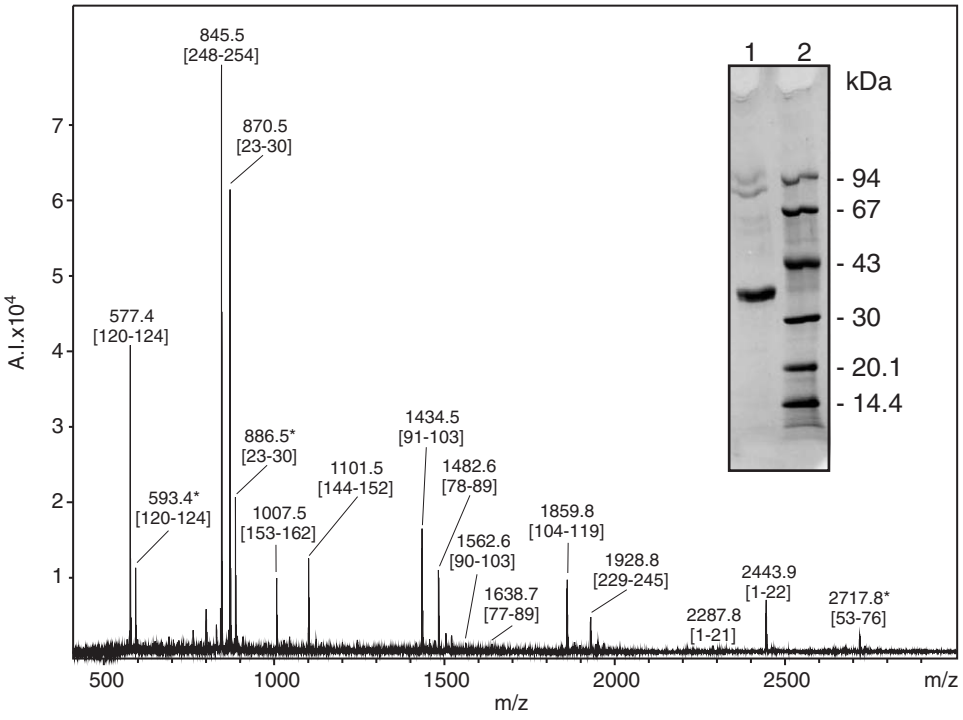

Fig. 5. Peptide mass-fingerprinting analysis of the isolated Dorin M. Isolated Dorin M (~2 μg) was separated from impurities by reducing SDS-PAGE, the gel was Coomassie stained/destained and the band was "in-gel" digested with trypsin. The fragment profile was measured by MALDI-TOF mass spectrometry. The numbers in parentheses below the fragment masses correspond to the amino-acid residue positions in the mature protein (as in Fig. 3). Asterisks by the mass data depict the presence of oxidized methionin in the fragment. Inset: Lane 1—isolated Dorin M; Lane 2—molecular weight markers.

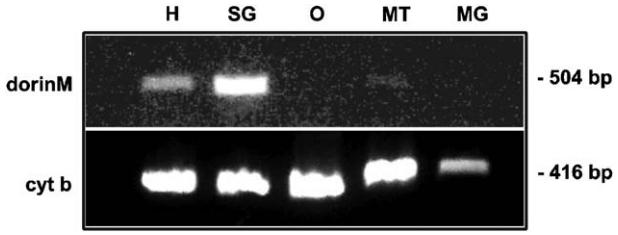

Fig. 6. RT-PCR analysis of Dorin M tissue-specific expression. Total RNA was extracted from various tissues of *O. moubata* and subjected to RT-PCR analysis using Dorin M (dorin M) and cytochrome b (cyt b) gene-specific primers. H—hemocytes; SG—salivary glands; O—ovary; MT—malphigian tubules; MG—midgut. For details see Materials and methods.

The tissue specific expression of Dorin M was investigated by RT-PCR analysis of total RNA isolated from hemocytes, salivary glands, ovary, malphigian tubules and midgut. The mitochondrial gene for cytochrome b (NADH dehydrogenase 6) was used as internal control. As shown in Fig. 6, Dorin M was found to be mainly expressed in salivary glands and hemocytes. Marginal expression was also detected in malphigian tubules, and no expression was observed in ovary and midgut.

4. Discussion

In the present study, we have cloned and sequenced the hemolymph plasma lectin, Dorin M, from the soft tick *O. moubata*. Dorin M is a glycoprotein which, in the native state, forms high molecular weight aggregates of about 640 kDa consisting of non-covalently bound 37 kDa subunits. Upon chemical deglycosylation, the Dorin M subunit mass was reduced of about 5 kDa (Kovář et al., 2000). These previous data have been fully confirmed by the Dorin M primary structure presented herein. The calculated mass of the mature protein is 29,114 Da, and it contains three putative N-linked glycosylation sites in the deduced sequence. The composition and structure of the Dorin M glycans were determined and are to be reported in a separate study (manuscript in preparation).

Overlapping PCR and RACE PCR products completed the entire cDNA sequence. The deduced amino-acid sequence was almost completely confirmed by the peptide mass-fingerprint analysis (Fig. 5) with the expected exception of glycosylated fragments.

Several arthropod lectins have been reported to be of hemocyte origin (Fujita et al., 1998; Iwanaga et al., 1998; Pace et al., 2002; Wang et al., 2004). Immunofluorescent localization of Dorin M in the hemocyte cytoplasm, isolation of its mRNA from hemocytes and the RT-PCR result of tissue-specific expression implies that Dorin M present in tick plasma is synthetized in hemocytes. In addition to it, Dorin M is also significantly expressed in salivary glands which was further confirmed by preliminary immunolocalization experiments (data not shown). The expression pattern of Dorin M is very similar to that of

another plasma protein from *O. moubata*, see the α_2-macroglobulin (Saravanan et al, 2003). The precise immunolocalization of Dorin M in salivary glands and its clear differentiation from a possible presence of OMFREP has to await the results of another study based on monospecific antibodies raised against recombinant N-terminal (low similarity) portions of both molecules.

A meaningful sequence similarity of the C-terminal part of Dorin M with the family of molecules possessing fibrinogen-like domain has been found. Members of this protein family have been identified not only from vertebrates but also from invertebrates (Adema et al., 1997; Dimopoulos et al., 2000; Kenjo et al., 2001; Wang et al., 2004; Xu and Doolittle, 1990). Dorin M confirms the existence of fibrinogen-related molecules similar to vertebrate ficolins in arthropods, in addition to those recognized in the horseshoe crab (Gokudan et al., 1999), mosquitoes (Dimopoulos et al., 2000; Wang et al., 2004) and Drosophila (Baker et al., 1990). The majority of proteins bearing the fibrinogen-like domain have been described as lectins (Doolittle et al., 1997). Moreover, it has been suggested that the fibrinogen-related molecules could have evolved from a lectin-like molecule (Doolittle et al., 1997), which may have functioned as a non-self recognizing protein (Gokudan et al., 1999).

Ficolins are a group of proteins containing both collagen-like and fibrinogen-like domains (Lu and Le, 1998). Non-self recognition is assigned to the carbohydrate-binding activity of fibrinogen-like domain, suggesting that the primary function of ficolins may have been playing a role in non-self recognition (Adema et al., 1997; Matsushita et al., 1996). It looks like ficolins have evolved to recognize the surface carbohydrate characteristic for pathogens and their binding. It has been reported that the collagen-like domains of ficolins are responsible for forming complexes with specific serine proteases, identical to mannose-binding lectins (MBL) that are associated with serine proteases (MASPs) and, thus, activate the complement after pathogen recognition (Matsushita et al., 2001).

A homology search of Dorin M revealed significant sequence similarity with those of Tl-5A, its homolog Tl-5B and vertebrate ficolins (Fig. 4). However, the collagen-like domain found in ficolins is missing in Dorin M and likewise in the corresponding region of Tls-5. In general, the fibrinogen-like domain consists of 220–250 residues, and is characterized by the presence of 24 invariant, mostly hydrophobic residues including four cysteines and 40 highly conserved residues (Holmskov et al., 2003). Four conservative cysteines likely forming two disulfide bridges (Cys_{49}–Cys_{80} and Cys_{206}–Cys_{219}) in native Tl-5A and (Cys_{68}–Cys_{96} and Cys_{223}–Cys_{236}) in native Tl-5B (Kairies et al., 2001) have analogues in the sequence of native Dorin M (Cys_{32}–Cys_{62} and Cys_{189}–Cys_{202}) (indicated by asterisks in Fig. 3). This implies the similar secondary structure. Unlike Tls-5, mature Dorin M has no more cystein residues, which may form interchain disulfide bridge(s). This is consistent with the observation that Dorin M molecules are not covalently linked (Kovář et al., 2000).

Recent characterization of ficolins present in human, mouse and pig serum/plasma, and in the body fluids of ascidians has revealed that they are lectins with common binding specificity for GlcNAc and human serum ficolins which function as opsonins (Matsushita et al., 1996). Monosaccharide containing *N*-acetyl group, such as GlcNAc and GalNAc inhibit the hemagglutination of Tls-5 (Gokudan et al., 1999; Kawabata and Tsuda, 2002). The same inhibitory effect has been seen for Dorin M (Kovář et al., 2000). Both Tls-5 and Dorin M had strong hemagglutinating activity and the best inhibitors examined for them are similar, *N*-acetyl neuraminic acid (Tls-5) and *N*-acetyl-neuraminyl lactose (Dorin M). Thus, the presence of *N*-acetyl group on the surface of pathogens is probably crucial for the recognition by ficolins, Tls-5 as well as Dorin M.

Dorin M hemagglutinating activity is, unlike both Tls-5, not Ca^{2+}-dependent. This seems to be consistent with the differences in the region of Dorin M that aligns with the Tl-5A calcium-binding motif (Kairies et al., 2001). The Ca^{2+}-binding site in Tl-5A is located near the ligand binding pocket. The amino-acids that mediate contact between the Tl-5A and the carbohydrate ligand are four aromatic side chains (Tyr_{210}, Tyr_{236}, Tyr_{248} and His_{220}), which form a funnel with methyl side chain of Ala_{237}, in which the methyl group of GlcNAc fits in the middle (Kairies et al., 2001). Regarding the very similar binding specificity and overall affinity of Dorin M to Tl-5A, we expect parallelism in ligand binding pocket structure. Side chains of aromatic amino-acids Tyr_{199}, Tyr_{219}, $Tyr_{235–237,239}$, His_{203} and Ala_{220} represent the potential funnel segments that mediate contact between the protein and the carbohydrate.

Molecules related to Dorin M, such as Tl-5A and ficolins, have been described as humoral factors of the innate immune systems capable to recognize pathogen-associated molecular patterns. Moreover, it is generally accepted that lectins are involved not only in non-self recognition but also may take part in vector transmission of pathogens (Grubhoffer et al., 1997; Maudlin and Welburn, 1988). Special attention has been focused on lectins that display a certain binding activity for sialic acids, since this carbohydrate moiety has been occasionally identified as a constituent of bacterial cell walls, including the spirochete *Borrelia burgdorferi* (Hulínská et al., 1991). We assume that previously isolated and characterized plasma lectin Dorin M (Kovář et al., 2000) may take part in the mechanism of transmission of another spirochete, *B. duttoni*, the causative agent of relapsing fever transmitted by the tick *O. moubata*.

The protein sequence similarity of Dorin M with invertebrate lectins and mammalian ficolins points out a suggested significance in the evolutionary conservation among the effector proteins in both vertebrates and invertebrates with regards to their innate immunity (Hoffmann et al., 1999). The Dorin M sequence confirmed

the prediction that Dorin M is a member of the fibrinogen-related proteins family (Kawabata and Tsuda, 2002).

Acknowledgments

This work has been supported by grants of the Grant Agency of the Academy of Sciences of the Czech Republic (nos. A6022701 and A6022001/00 to L.G., and no. IAA602307 to P.K.), Grant Agency of the Czech Republic (no. 206/00/0266 to P.K.), Ministry of Education, Youth and Sports of the Czech Republic (no. 123100003) and research project of the Institute of Parasitology, Academy of Sciences of the Czech Republic (no. Z6 022 909). The Edman sequencing and MS analysis was performed in the Neurochemistry Department of the Freie Universität Berlin (F. Hucho) and was supported by the Fonds der Chemischen Industrie.

References

Adema, C.M., Hertel, L.A., Miller, R.D., 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. USA 94, 8691–8696.

Baker, N.E., Mlodzik, M., Rubin, G.M., 1990. Spacing differentiation in the developing *Drosophila* eye: a fibrinogen-related lateral inhibitor encoded by scabrous. Science 250, 1370–1377.

Dimopoulos, G., Casavant, T.L., Chang, S.R., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B., Kafatos, F.C., 2000. *Anopheles gambiae* pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines. Proc. Natl. Acad. Sci. USA 97, 6619–6624.

Doolittle, R.F., Spraggon, G., Everse, S.J., 1997. Evolution of vertebrate fibrin formation and the process of its dissolution. Ciba Found. Sympos. 212, 4–17.

Fujita, Y., Kurata, S., Homma, K., Natori, S., 1998. A novel lectin from *Sarcophaga*. Its purification, characterization, and cDNA cloning. J. Biol. Chem. 273, 9667–9672.

Gadjeva, M., Thiel, S., Jensenius, J.C., 2001. The mannan-binding lectin pathway of the innate immune response. Curr. Opin. Immunol. 13, 74–78.

Gokudan, S., Muta, T., Tsuda, R., Koori, K., Kawahara, T., Seki, N., Mizunoe, Y., Wai, S.N., Iwanaga, S., Kawabata, S.I., 1999. Horseshoe crab acetyl group recognizing lectins involved in innate immunity are structurally related to fibrinogen. Proc. Natl. Acad. Sci. USA 96, 10086–10091.

Grubhoffer, L., Kovář, V., 1998. Arthropod lectins: affinity approaches in analysis and reparation of carbohydrate binding proteins. In: Wiesner, A., Dunphy, G.B., Marmaras, V.J., Morishima, I., Sugumaran, M., Yamakava, M. (Eds.), Techniques in Insect Immunology FITC-5. SOS Publications, New Haven, pp. 47–57.

Grubhoffer, L., Hypša, V., Volf, P., 1997. Lectins (hemagglutinins) in the gut of the important disease vectors. Parasite 4, 203–216.

Hoffmann, J.A., Kafatos, F.C., Janeway, C.A., Ezekowitz, R.A., 1999. Phylogenetic perspectives in innate immunity. Science 284, 1313–1318.

Holmskov, U., Thiel, S., Jensenius, J.C., 2003. Collections and ficolins: humoral lectins of the innate immune defense. Annu. Rev. Immunol. 21, 547–578.

Hulínská, D., Volf, P., Grubhoffer, L., 1991. Surface carbohydrates of *Borrelia burgdorferi*. In: Dusbábek, F., Bukva, V. (Eds.), Modern Acarology, vol. 2. SPB Academic Publishing bv, The Hague, pp. 73–78.

Iwanaga, S., 2002. The molecular basis of innate immunity in the horseshoe crab. Curr. Opin. Immunol. 14, 87–95.

Iwanaga, S., Kawabata, S., Muta, T., 1998. New types of clotting factors and defense molecules found in horseshoe crab hemolymph: their structures and functions. J. Biochem. (Tokyo) 123, 1–15.

Kairies, N., Beisel, H.G., Fuentes-Prior, P., Tsuda, R., Muta, T., Iwanaga, S., Bode, W., Huber, R., Kawabata, S., 2001. The 2.0-A crystal structure of tachylectin 5A provides evidence for the common origin of the innate immunity and the blood coagulation systems. Proc. Natl. Acad. Sci. USA 98, 13519–13524.

Kawabata, S.I., Tsuda, R., 2002. Molecular basis of non-self recognition by the horseshoe crab tachylectins. Biochim. Biophys. Acta 1572, 414–421.

Kenjo, A., Takahashi, M., Matsushita, M., Endo, Y., Nakata, M., Mizuochi, T., Fujita, T., 2001. Cloning and characterization of novel ficolins from the solitary ascidian, *Halocynthia roretzi*. J. Biol. Chem. 276, 19959–19965.

Kovář, V., Kopáček, P., Grubhoffer, L., 2000. Isolation and characterization of Dorin M, a lectin from plasma of the soft tick *Ornithodoros moubata*. Insect Biochem. Mol. Biol. 30, 195–205.

Le, Y., Lee, S.H., Kon, O.L., Lu, J., 1998. Human L-ficolin: plasma levels, sugar specificity, and assignment of its lectin activity to the fibrinogen-like FBG domain. FEBS Lett. 425, 367–370.

Lu, J., 1997. Collectins: collectors of microorganisms for the innate immune system. Bioassays 19, 509–518.

Lu, J., Le, Y., 1998. Ficolins and the fibrinogen-like domain. Immunobiology 199, 190–199.

Matsushita, M., 1996. The lectin pathway of the complement system. Microbiol. Immunol. 40, 887–893.

Matsushita, M., Fujita, T., 2001. Ficolins and the lectin complement pathway. Immunol. Rev. 180, 78–85.

Matsushita, M., Endo, Y., Taira, S., Sato, Y., Fujita, T., Ichikawa, N., Nakata, M., Mizuochi, T., 1996. A novel human serum lectin with collagen- and fibrinogen-like domains that functions as an opsonin. J. Biol. Chem. 271, 2448–2454.

Matsushita, M., Endo, Y., Nonaka, M., Fujita, T., 2001. Activation of the lectin complement pathway by ficolins. Int. Immunopharmacol. 1, 359–363.

Maudlin, I., Welburn, S.C., 1988. The role of lectins and trypanosome genotype in the maturation of midgut infections in *Glossina morsitans*. Trop. Med. Parasitol. 39, 56–58.

Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 10, 1–6.

Pace, K.E., Lebestky, T., Hummel, T., Arnoux, P., Kwan, K., Baum, L.G., 2002. Characterization of a novel *Drosophila melanogaster* galectin. Expression in developing immune, neural, and muscle tissues. J. Biol. Chem. 277, 13091–13098.

Ratcliffe, N.A., Rowley, A.F., 1987. Insect responses to parasites and other pathogens. In: Soulsby, E.J.L. (Ed.), Immune Responses in Parasitic Infections, vol. 4. CRC Press, Boca Raton, FL, pp. 271–332.

Saravanan, T., Weise, C., Sojka, D., Kopáček, P., 2003. Molecular cloning, structure and bait region splice variants of α_2-macroglobulin from the soft tick *Ornithodoros moubata*. Insect Biochem. Mol. Biol. 33, 841–851.

Vasta, G.R., 1991. The multiple biological roles of invertebrate lectins: their participation in non-self recognition mechanisms. In: Warr, G.W., Cohen, N. (Eds.), Phylogenesis of Immune Functions. CRC Press, Boca Raton, FL, pp. 74–101.

Vasta, G.R., Quesenberry, M.S., Ahmed, H.A., O'Leary, N., 1999. C-type lectins and galectins mediate innate and adaptive immune functions: their roles in the complement activation pathway. Dev. Comput. Immunol. 23, 401–420.

Wang, X., Rocheleau, T.A., Fuchs, J.F., Hillyer, J.F., Chen, C.-C., Christensen, B.M., 2004. A novel lectin with a fibrinogen-like domain and its potential involvement in the innate immune response of *Armigeres subalbatus* against bacteria. Insect Mol. Biol. 13, 273–282.

Xu, X., Doolittle, R.F., 1990. Presence of a vertebrate fibrinogen-like sequence in an echinoderm. Proc. Natl. Acad. Sci. USA 87, 2097–2101.