Supporting Information

Buchon et al. 10.1073/pnas.0901924106

SI Text

Production of Recombinant ModSP, GNBP1 and Grass Proteins. ModSP was expressed with the *Bombyx mori* protein disulfide isomerase (PDI) as a fusion partner in Sf9 cell as previously described (1). The ER retention signal (KDEL) and signal sequence of PDI cDNA were removed to make secreted PDI. DNA fragments encoding PDI deleted of the KDEL fragment and ModSP were generated by PCR using cDNA templates and cloned into the pFASTBac vector (Invitrogen). A *Tobacco* etch virus (TEV) cleavage site was designed between PDI and ModSP. Sf9 cells were subsequently transfected with the pFASTBac PDI-TEV-ModSP plasmid and the viruses were harvested and amplified as described in the Bac-to-Bac baculovirus expression system manual (Invitrogen). For protein expression, 1-L cultures of sf9 cells were infected by recombinant baculovirus containing PDI-TEV-ModSP, and supernatant was collected 48 h after infection. After centrifugation, proteins of the supernatants were concentrated on a Vivaflow 200 system (Satorius) with buffer exchange to binding buffer (150 mM NaCl and 20 mM Tris·HCl, pH 8.0). His-tagged recombinant proteins were purified by affinity column chromatography using TARON Metal Resin (Clontech). N-terminal amino acid sequencing of the catalytic serine protease domain of recombinant ModSP was used to verify the identities of the purified proteins. *Drosophila* recombinant GNBP1 protein was produced as indicated for ModSP except that His-tagged recombinant GNBP1 was purified by affinity column chromatography using Ni-NTA Resin (Qiagen). TEV protease was added to the fraction containing expressed rPDI-fusioned Dm-GNBP1 and incubated for 48 h at 4 °C. For purification of r*Dm*-GNBP1, TEV treated rPDI-fusioned *Dm*-GNBP1 was applied to a Hitrap-Q FPLC column preequilibrated with 20 mM citrate buffer (pH 6.0) and repeatedly washed until the absorbance at 280 nm was zero. The column was eluted with a linear gradient of NaCl from 0 to 1 M >100 min. The purified recombinant *Dm*-GNBP1 was confirmed by N-terminal amino acid sequencing. Recombinant Grass was expressed without PDI fusion partner in Sf9 cells. Recombinant Grass protein was purified using the protocol for Dm-GNBP1 and confirmed by N-terminal amino acid sequencing.

Amino Acid Sequencing. To obtain amino acid sequences, proteins were blotted onto a polyvinylidene difluoride (PVDF) membrane and strained with solution containing 0.1% Coomassie brilliant blue R-250 and 50% methanol. The membrane was destained with 50% methanol containing 10% acetic acid (vol/vol) until the protein bands became visible. The fragments of interest were excised and analyzed on a protein sequencer (Applied Biosystems) using standard techniques.

Antibodies Against N-Terminal Domain and C-Terminal Domain of ModSP. The 2 different polyclonal antibodies against ModSP were raised using keyhole limpet hemocyanin (KLH)-conjugated synthetic peptides: EDDIRQHDRRLQGNC is located at the N terminus of LDL domain *Dm*-ModSP and APGYKGRTENYYQ is located at the C terminus of catalytic serine protease domain of *Dm*-ModSP. KLH-conjugated synthetic peptides with >95% purity were purchased from Peptron . Rabbit was injected s.c. with each 500 μg of KLH-conjugated peptide in Freund's complete adjuvant. A booster dose (250 μg in Freund's incomplete adjuvant) was given after 4 weeks, and the animals were bled 2–3 weeks thereafter. The resulting antibodies were affinity-purified as previously described (2).

Measurement of the Lys-Type Peptidoglycan-Dependent PO Activity. Hemolymph was collected as previously described (2). An assay of PO activity was carried out according to our previously published method (3). Briefly, to measure PO activity, 30 μL crude hemolymph (350 μg proteins) were preincubated in 70 μL 20 mM Tris·HCl buffer (pH 8.0) containing 10 μL serially diluted soluble Lys-PG for 10 min at 30 °C, and then 400 μL of substrate solution (1 mM 4-methyl-catechol, 2 mM 4-hydroxyproline ethylester in 20 mM Tris·HCl buffer, pH 8.0, containing 5 mM $CaCl_2$) were added to the reaction mixture. After incubation at 30 °C for 20 min, the increase in absorbance at 520 nm was measured using a Shimadzu spectrophotometer. One unit of PO activity was defined as the amount of enzyme causing an increase in absorbance of 1 at 520 nm per 20 min incubation (A_{520}/20 min).

Live Imaging. Flies were dissected in PBS and directly mounted in an antifading solution (Citifluor). Samples were observed with an Axioplot (Zeiss), and photos were taken using the deconvolution properties of an apotome grid (Zeiss).

Staining of Lipid Vesicles. The hemolymph of 20 flies expressing *UAS-ModSP-GFP* under the control of the *c564-gal4* driver was extracted. Whole hemolymph was immediately fixed on slides with PBS:paraformaldehyde 4%, for 2 min, rinsed with PBS, and stained for 5 min in a solution of oil red O. After extensive rinsing, slides were mounted in PBS/glycerol and observed using an Axioplot microscope (Zeiss).

1. Goo TW, et al. (2008) *Bombyx mori* protein disulfide isomerase enhances the production of nuecin, an antibacterial protein. *BMB Rep* 41:400–403.
2. Zhang R, et al. (2003) Characterization and properties of a 1,3-beta-D-glucan pattern recognition protein of *Tenebrio molitor* larvae that is specifically degraded by serine protease during prophenoloxidase activation. *J Biol Chem* 278:42072–42079.
3. Park JW, et al. (2007) Clustering of peptidoglycan recognition protein-SA is required for sensing lysine-type peptidoglycan in insects. *Proc Natl Acad Sci USA* 104:6602–6607.
4. Ji C, Wang Y, Guo X, Hartson S, Jiang H (2004) A pattern recognition serine proteinase triggers the prophenoloxidase activation cascade in the tobacco hornworm, *Manduca sexta*. *J Biol Chem* 279:34101–34106.
5. Kim CH, et al. (2008) A three-step proteolytic cascade mediates the activation of the peptidoglycan-induced toll pathway in an insect. *J Biol Chem* 283:7599–7607.
6. Giebel B, Stuttem I, Hinz U, Campos-Ortega JA (1997) Lethal of scute requires overexpression of daughterless to elicit ectopic neuronal development during embryogenesis in *Drosophila*. *Mech Dev* 63:75–87.
7. Kambris Z, et al. (2006) *Drosophila* immunity: A large-scale in vivo RNAi screen identifies five serine proteases required for Toll activation. *Curr Biol* 16:808–813.

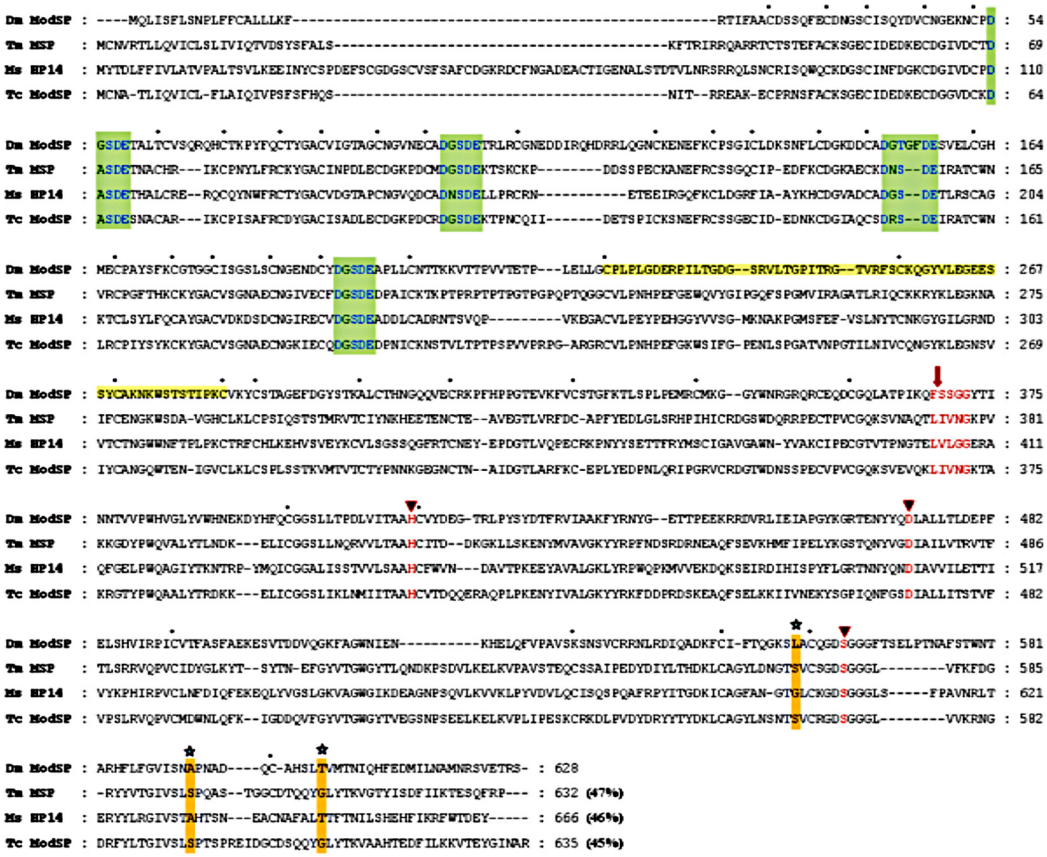

Fig. S1. Sequence comparison between ModSP and other insect homologs. ModSP protein sequence was compared with the sequence of *T. molitor* MSP, *M. sexta* hemolymph protease 14 (*Ms*-HP14) (4, 5), and *Tribolium castaneum* MSP (*Tc*-MSP, XP_967486); closed circles and red-green closed reverse triangles indicate the conserved cysteine and catalytic triad residues of the *Dm*-ModSP. Four green open boxes and the yellow highlighted box show the specific residues of LDLa and CCP domains, respectively. The red arrow indicates the site where the *Dm*-ModSP is cleaved during activation. The blue stars indicate the substrate-specificity residues of *Dm*-ModSP. Gaps were introduced to obtain maximal sequence similarity. Percent values indicate the sequence identities between *Dm* ModSP and other proteins.

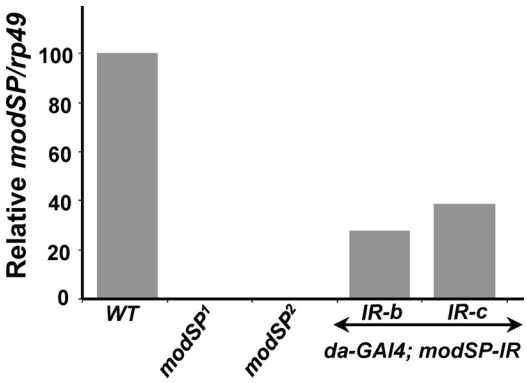

Fig. S2. Relative *modSP* expression after RNAi using the *da-GAL4* driver in combination with *modSP-IR* constructs and in *modSP* mutants. Relative *modSP* expression is shown as a ratio over the expression of the housekeeping gene *rp49*.

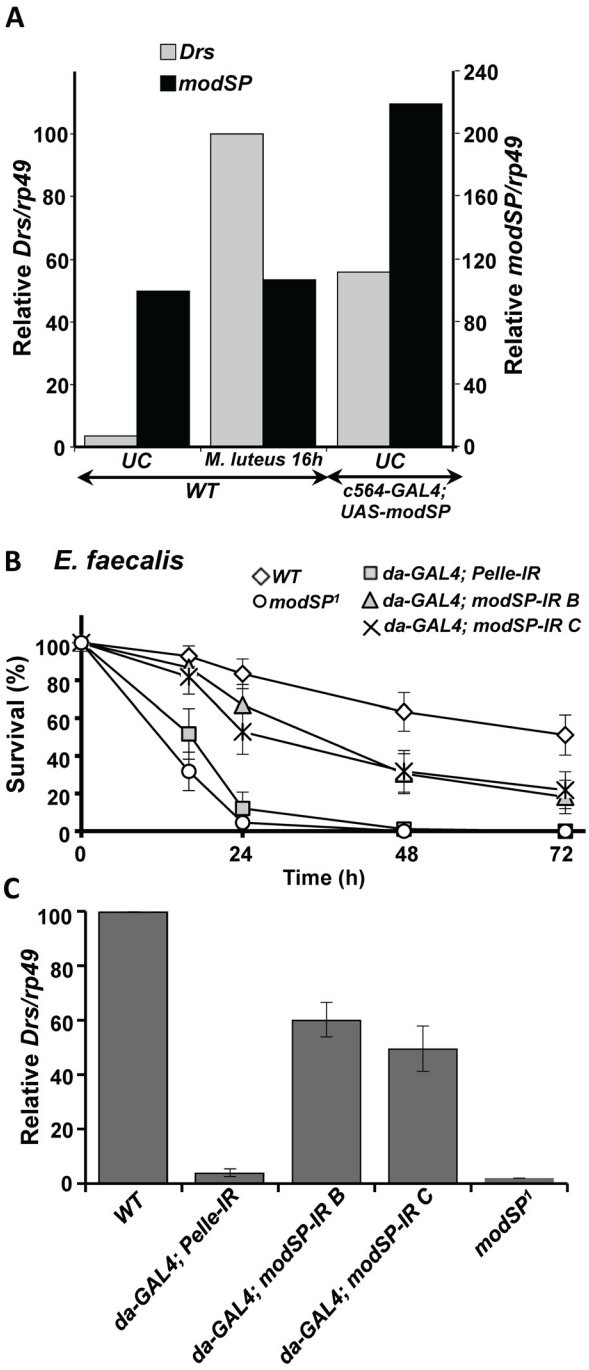

Fig. S3. Silencing of *modSP* by RNAi reduces Toll pathway activity. Overexpression of *modSP* with the fat body Gal4 driver *c564-Gal4* induces a high level of *Drs* expression in absence of challenge. Levels of *Drs* and *modSP* expression were monitored by RT-qPCR and normalized on the expression level of *rp49* (*A*). The survival rates (*B*) and the level of *Drs* expression (*C*) of *modSP-RNAi* flies infected with *E. faecalis* and *M. luteus*, respectively, were compared with wild-type flies, *pelle-RNAi*, and *modSP*^{1} flies. In agreement with the partial extinction of the *modSP* gene expression in the *modSP RNAi* flies (Fig. S2), the phenotype was weaker than that observed in *modSP*^{1} deficient flies. Two independent insertions of a *UAS-modSP-IR* construct (*B* and *C*) were used in combination with the ubiquitous GAL4 driver *da-GAL4* (6). The *UAS-modSP-IR* RB and RC were obtained from the Vienna RNAi stock center to produce dsRNA targeting the region 222–582 of the mRNA. A *UAS-pelle-IR* (National Institute of Genetics, Mishima, Japan) construct targeting the intracellular component of the Toll pathway Pelle was used as a positive control (4).

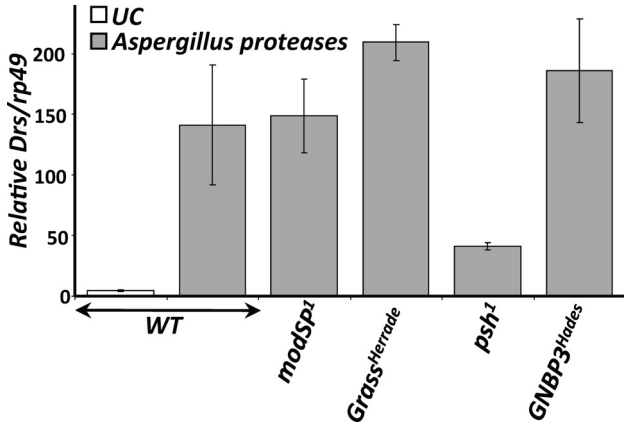

Fig. S4. *modSP*[1] does not block Toll activation by proteases derived from *Aspergillus oryzae*. *Drs* gene expression was monitored by RT-qPCR with total RNA extracted from wild-type, *GNBP3*^{Hades}, *Grass*^{Herrade}, *psh*[1], and *modSP*[1] females collected 16 h after injection with proteases of *Aspergillus*. UC, unchallenged flies.

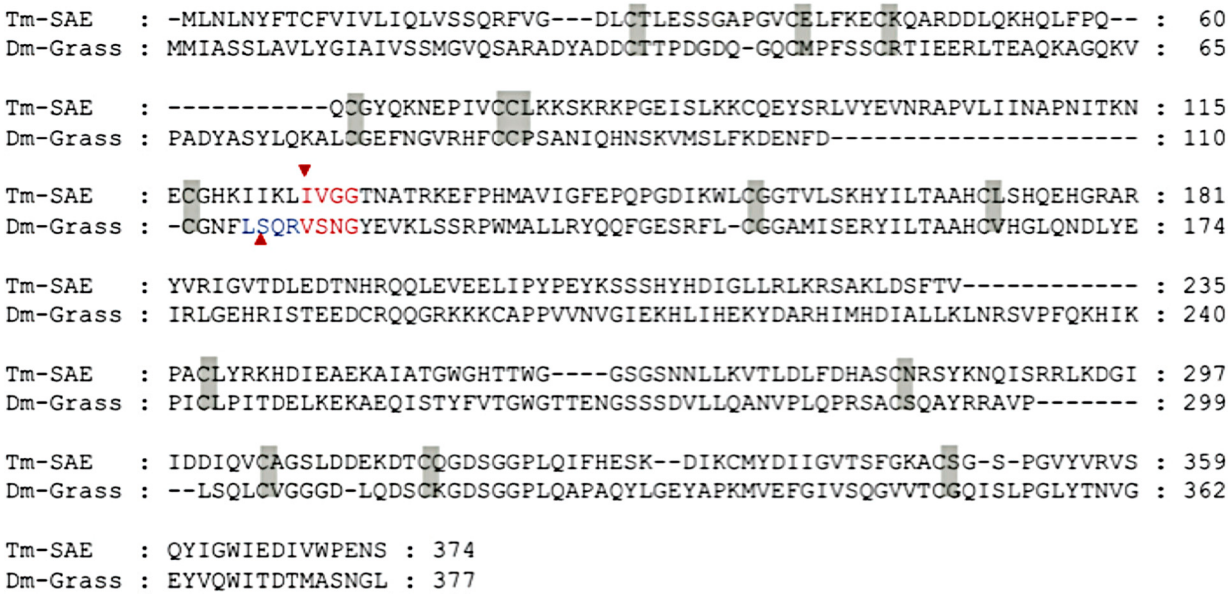

Fig. S5. The Grass protein sequence significantly differs from *T. molitor* SAE (*Tm*-SAE) (5). The predicted cleavage site of both serine proteases is shown. To examine the possibility that ModSP can cleave Grass, we compared cleavage sites between Grass and *Tm*-SAE, which is known to function directly downstream of *Tm*-MSP. The cleavage site of *Tm*-SAE (L↓IVGG, requiring a chymotrypsin-like activity, indicated in red) is not conserved in Grass (R↓VSNG, requiring a trypsin-like enzyme as a direct upstream protease, indicated in red). Additionally, we showed in this study that ModSP and *Tm*-MSP are cleaved at the exact same site, indicating that ModSP has chymotrypsin-like specificity (Fig. 6*B*). However, a direct activation of Grass by ModSP cannot be ruled out since *Dm*-ModSP and *Tm*-MSP display some differences in key amino acid residues of the putative cleavage site (F↓SSGG versus L↓IVNG, respectively) and in the substrate specificity pocket (Leu, Ala, Thr versus Ser, Ser, Gly, respectively). Nevertheless, a putative chymotrypsin cleavage site (L↓SQR, in blue) is observed in the sequence of Grass. Alternatively, an unidentified serine protease with a chymotrypsin-like cleavage site at its N-terminal domain could function between ModSP and Grass. The conserved cysteine residues are indicated in shadow.

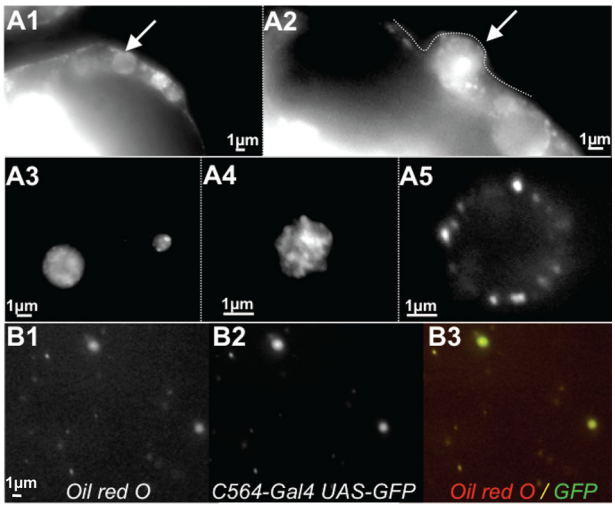

Fig. S6. A ModSP-GFP fusion is secreted in lipid vesicles. Use of flies expressing a GFP tagged version of ModSP (Genotype: *c564-Gal4*; *UAS-ModSP-GFP*) revealed that ModSP-GFP was secreted from the fat body into the hemolymph at the surface of small lipid vesicles of 0.2–5 μm. This unusual mode of protein secretion and localization is consistent with the presence of LDLa repeats that could anchor ModSP to lipoproteins. We used a staining against Caspase-3 activity to demonstrate that the fat body of *c564-Gal4*; *UAS-ModSP-GFP* flies was not apoptotic. This indicates that these vesicles are not because of an abnormal cell death consecutive to the overexpression procedure. (*A*) The study of flies expressing a *UAS-modSP-GFP* under the control of the fat body Gal4 driver (*C564-Gal4*) suggests that ModSP was released from the fat body (*A1*, *A2*) into the hemolymph in lipid vesicles ranging from 0.5 to 5 μm (*A3–A5*). *A2* shows vesicle budding from the fat body (highlighted by the white line). ModSP-GFP localizes at the membrane of vesicles (*A4*: Focus on the surface; *A5*: Section). (*B*) A staining with Oil red O reveals that the vesicles expressing ModSP-GFP are lipidic (*B1*, oil red O, *B2*, GFP, *B3*, merge).

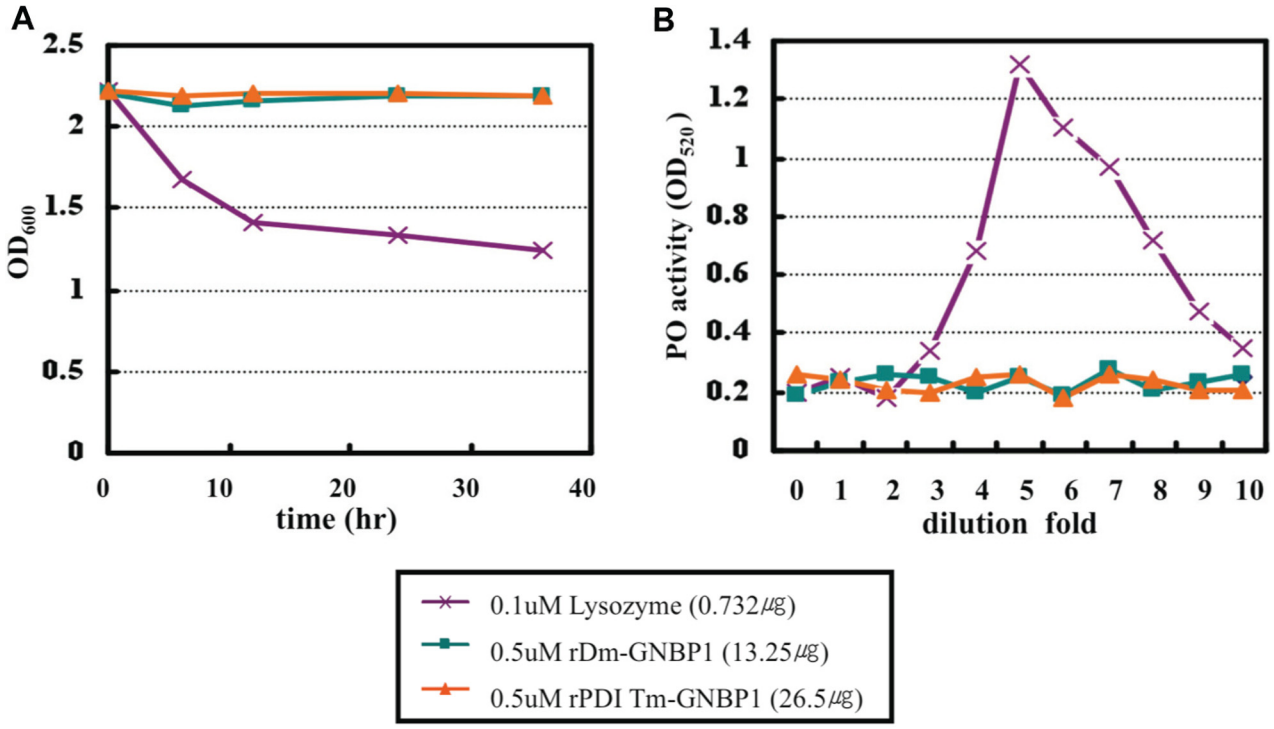

Fig. S7. Recombinant *Tenebrio* and *Drosophila* GNBP1 does not exhibit enzymatic activity against *Micrococcus luteus* lysine-type peptidoglycan. (*A*) Lysine-type peptidoglycan (500 μg) from *M. luteus* was mixed with either rDm-GNBP1, rPDI Tm-GNBP1, or lysozyme (catalog no. L7001; Sigma-Aldrich) in 12.5 mM sodium phosphate buffer (pH 5.5) and incubated at 37 °C under gentle agitation. OD_{600} values taken at 0, 12, 24, and 36 h are depicted. The rapid hydrolysis of peptidoglycan was observed with lysozyme (as observed by a decrease in OD). Full-length *Dm*- and *Tm*-GNBP1 did not show any hydrolytic activity. This experiment was repeated twice and gave similar results. Brown x-mark, green-filled box, and orange triangle indicate lysozyme (0.1 μM, 0.732 μg), rDm-GNBP1 (0.5 μM, 13.25 μg), and rPDI-Tm-GNBP1 (0.5 μM, 26.5 μg), respectively. (*B*) Insoluble polymeric peptidoglycan incubated with either *Tm*-GNBP1 or GNBP1 did not shown any PO activity indicating that neither GNBP1 nor *Tm*-GNBP1 is able to digest peptidoglycan. In contrast, insoluble polymeric peptidoglycan incubated with lysozyme induced PO activity as illustrated by as Bell-shaped curve of PO activity as shown previously (5). Note in contrast of *Tm*-GNBP1 (5), we did not have a functional assay to demonstrate that the *Drosophila* GNBP1 was indeed active.