A single modular serine protease integrates signals from pattern-recognition receptors upstream of the *Drosophila* Toll pathway

Nicolas Buchon[a,1], Mickael Poidevin[b,1], Hyun-Mi Kwon[c], Aurélien Guillou[a], Valentin Sottas[a], Bok-Luel Lee[c], and Bruno Lemaitre[a,b,2]

aGlobal Health Institute, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland; bCentre de Génétique Moléculaire (CGM), Centre National de la Recherche Scientifique, 91198 Gif-sur-Yvette, France; and cNational Research Laboratory of Defense Proteins, College of Pharmacy, Pusan National University, Kumjeong Ku, Busan 609-735, Korea

The *Drosophila* Toll receptor does not interact directly with microbial determinants, but is instead activated by a cleaved form of the cytokine-like molecule Spätzle. During the immune response, Spätzle is processed by complex cascades of serine proteases, which are activated by secreted pattern-recognition receptors. Here, we demonstrate the essential role of ModSP, a modular serine protease, in the activation of the Toll pathway by Gram-positive bacteria and fungi. Our analysis shows that ModSP integrates signals originating from the circulating recognition molecules GNBP3 and PGRP-SA and connects them to the Grass-SPE-Spätzle extracellular pathway upstream of the Toll receptor. It also reveals the conserved role of modular serine proteases in the activation of insect immune reactions.

innate immunity | proteolytic cascade | insect immunity | antifungal | antimicrobial peptides

Sequential activation of extracellular serine protease (SP) cascades regulates important innate immune reactions, like blood clotting in arthropods and complement activation in vertebrates. In *Drosophila*, proteolytic cascades are also involved in the regulation of the Toll pathway, which mediates resistance to Gram-positive bacteria and fungi. Unlike mammalian Toll-like receptors, the *Drosophila* Toll receptor does not interact directly with microbial determinants and is instead activated by a cleaved form of the secreted cytokine-like molecule Spätzle (Spz) (1). The immune-induced cleavage of Spz is triggered by proteolytic cascades, conceptually similar to vertebrate blood coagulation or complement activation cascades. These proteolytic cascades have a functional core consisting of several SP that undergo zymogen activation, upon cleavage by an upstream protease (2, 3). Spätzle processing enzyme (SPE), an immune-regulated SP with a Clip-domain, has been identified as the terminal SP that maturates Spz (4, 5).

Genetic analysis supports the existence of several complex cascades of SP that link microbial recognition to activation of SPE. Pattern-recognition receptors (PRRs) are thought to be present in the hemolymph where they sense microbial-derived molecules (6). Detection of Gram-positive bacteria is mediated through the recognition of peptidoglycan by the peptidoglycan-recognition protein-SA (PGRP-SA) with the help of Gram-negative binding protein1 (GNBP1) (7–9). Activation of the Toll pathway by fungi is in part mediated by GNBP3 through the sensing of β-1,3-glucan (10). RNAi and loss of function analyses have shown that activation of SPE by either PGRP-SA/GNBP1 or GNBP3 requires Grass, a Clip-domain SP (4, 11).

Spores of entomopathogenic fungi such as *Beauveria bassiana* have the capacity to germinate on the fly cuticle and generate hyphae, which can penetrate the cuticle of insects and reach the hemolymph. It has been proposed that the presence of *B. bassiana* is detected, in a GNBP3-independent manner, through direct activation of the Toll pathway by a fungal protease PR1 (10). PR1 would directly cleave the host SP, Persephone (Psh), which triggers Toll pathway activation (10). Recently, this mode of activation was

extended to the sensing of proteases produced by various bacteria (11). Surprisingly, tracheal melanization in mutant larvae lacking the serpin Spn77Ba also activates the Toll pathway in a Psh-dependent manner (12). This suggests that Psh-dependent Toll pathway activation is induced by a host factor derived from melanization. This also points to a possible cross-talk between the proteolytic cascades that regulate the Toll pathway and those regulating the melanization reaction.

Despite these recent studies, the extracellular events that lead to the activation of the Toll pathway remain poorly characterized. For instance, the apical protease linking recognition by PRRs to the cleavage of Spz is not known, leaving an important gap in our knowledge of Toll pathway activation. The in vitro reconstruction of a similar cascade with purified proteases in the beetle *Tenebrio molitor*, suggests an important role for a modular SP (*Tm*-MSP). In this insect, binding of peptidoglycan to the PGRP-SA/GNBP1 complex induces activation of the *Tm*-MSP zymogen that in turn activates another SP (*Tm*-SAE) to cleave *Tm*-SPE, the *Tenebrio* homolog of SPE, resulting in both activation of Spz and melanization (13, 14).

This result prompted us to investigate the role of the *Drosophila* homolog of *Tm*-MSP, which is encoded by the *CG31217* gene. We generated a null mutation in *CG31217* by homologous recombination and demonstrate its essential role in the activation of Toll by secreted PRRs.

Results

Generation of a *modSP* Mutant by Homologous Recombination. The CG31217 serine protease, named here ModSP (for modular serine protease), contains 4 low density lipoprotein-receptor class A (LDLa) domains and 1 complement control protein (CCP) module at its N terminus (Fig. 1*A*). ModSP shares 47% amino acid identity with the *Tm*-MSP (Fig. S1). Unlike several other *Drosophila* SPs involved in the immune response, *modSP* is not transcriptionally induced upon microbial infection and does not contain a Clip-domain (15, 16).

To analyze the function of ModSP, homologous recombination was used to replace the entire *modSP* locus with a copy of the *white* gene (Fig. 1*B*). Two independent fly lines, *modSP^1* and *modSP^2*, were obtained, and the absence of *modSP* transcript in homozygous mutants was confirmed by RT-qPCR (Fig. S2). We verified that the

Fig. 1. Structure of ModSP and generation of *modSP* mutants. (*A*) Domain organization of the ModSP protein. Symbols indicate the domains of LDLa, CCP, and SP respectively (adapted from ref. 14). The LDL receptor class A domain contains 6 disulphide-bound cysteines and a cluster of negatively charged amino acids that form the LDL receptor binding sites for LDL and calcium. The complement control protein (CCP) module, also known as Sushi domain, is found in a wide variety of complement like molecules and adhesion proteins. (*B*) Schematic representation of the *modSP* deletion. The gene map was adapted from FlyBase and includes *modSP* and the neighboring genes. The deleted segment replaced by the *w* gene (black box) and the flanking sequences used for recombination (dotted lines) are indicated.

modSP deletion did not affect the expression of flanking genes. Both *modSP¹* and *modSP²* mutants were perfectly viable and exhibited no visible developmental defects. Here, we only report data obtained with the *modSP¹* mutant, because both alleles produced identical results.

ModSP Mediates Toll Activation by Gram-Positive Bacteria. When challenged by a range of Gram-positive bacteria (*Enterococcus faecalis, Listeria monocytogenes*), *modSP¹*-deficient flies rapidly succumbed to infection (Fig. 2*A* and *B*). The observed phenotypes were similar, albeit slightly weaker, to those generated by a null mutation in *spz* (*spz^rm7^*). The *modSP¹* mutation did not cause a
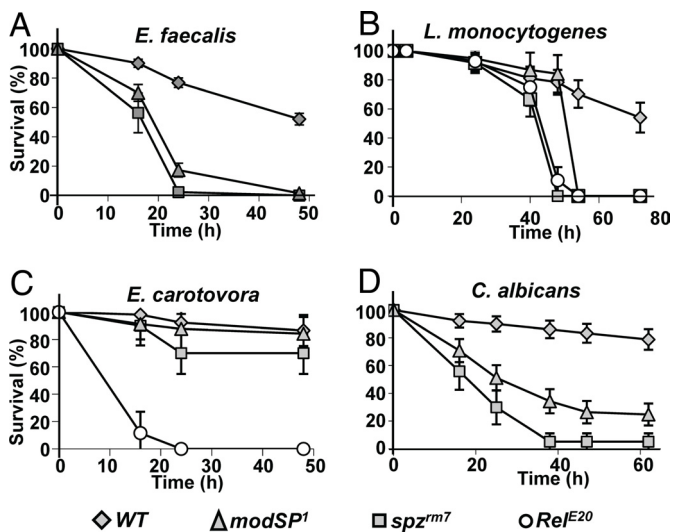
Fig. 2. ModSP is required for resistance to Gram-positive bacteria and fungal infections. The survival rates (%) of *modSP¹* flies infected with the Gram-positive bacteria *E. faecalis* (*A*), *L. monocytogenes* (*B*), the Gram-negative bacterium *E. carotovora 15* (*C*), and the yeast *C. albicans* (*D*) were compared with wild-type flies and flies mutant in either the Toll pathway (*spz^rm7^*) or the Imd pathway (*Rel^E20^*). Flies were infected by septic injury with a needle dipped in a concentrated pellet of each microbe (OD = 200, OD = 10 for *E. faecalis*). Bars represent confidence intervals at 5%.
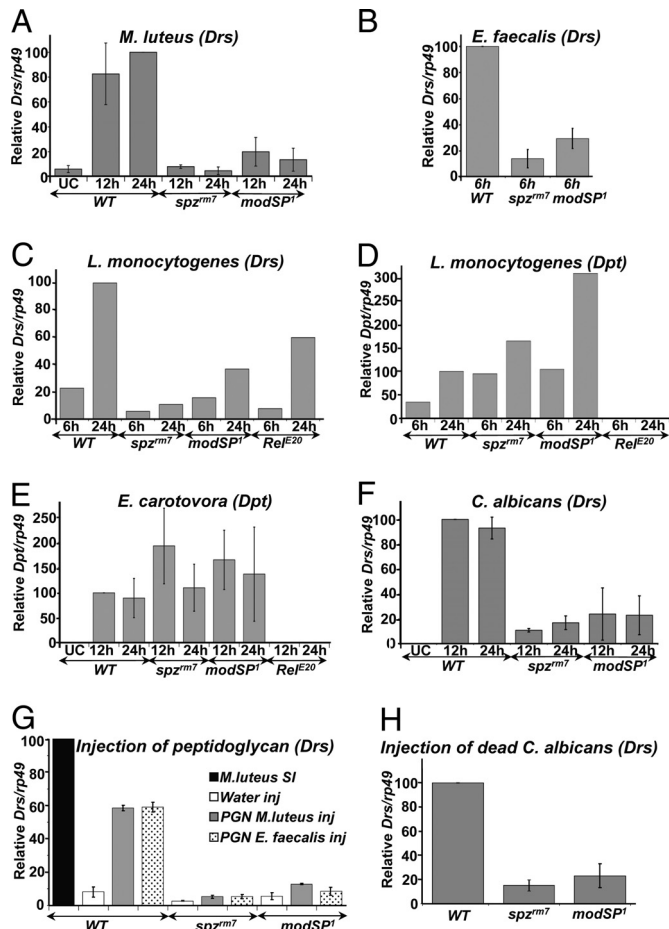
Fig. 3. Analysis of antimicrobial peptide gene expression in *modSP* mutant. *Drs* (*A–C* and *F–H*) or *Dpt* (*D* and *E*) gene expression was monitored by RT-qPCR with total RNA extracted from wild-type, *spz^rm7^*, *Rel^E20^*, and *modSP¹* adults collected at different time points after septic injury with *M. luteus* (*A*), *E. faecalis* (*B*), *L. monocytogenes* (*C* and *D*), *E. carotovora* (*E*), and *C. albicans* (*F*). (*G* and *H*) RT-qPCR analysis was performed with total RNA extracts from wild-type and mutant females collected 24 h after injection with 9 nL *M. luteus* peptidoglycan, *E. faecalis* PGN ([PGN] = 5 mg/mL) (*G*) or dead yeast (OD$_{600}$ = 300) (*H*). The mean and standard deviation of 2 to 3 independent experiments is shown, except *L. monocytogenes*, for which 1 representative experiment among 2 is shown.

general immune deficiency, because *modSP¹* flies exhibited a wild-type survival to a septic injury with the Gram-negative bacterium *Erwinia carotovora*, which is known to activate the Imd pathway (Fig. 2*C*).

To test whether this immune susceptibility results from improper Toll activity, we monitored the expression of the antifungal peptide gene *Drosomycin* (*Drs*), a target of the Toll pathway. *modSP¹* mutants weakly induced *Drs* after infection with the Gram-positive bacteria *M. luteus* and *E. faecalis* (Fig. 3 *A* and *B*). Lysine-type peptidoglycan, a peptidoglycan form found in most Gram-positive bacteria including *M. luteus* and *E. faecalis*, is a potent inducer of the PGRP-SA-Toll pathway (17). The *modSP¹* mutation strongly reduced the induction of *Drs* in response to purified peptidoglycan extracted from *M. luteus* and *E. faecalis* as observed in *spz^rm7^* mutants (Fig. 3*G*). This experiment demonstrates that ModSP is required for Toll pathway activity in a step downstream of the recognition of lysine-type peptidoglycan.

L. monocytogenes is a Gram-positive bacterium with diaminopimelic (DAP)-type peptidoglycan that can activate both Toll and Imd pathways (18). A strong reduction of *Drs* expression was

IMMUNOLOGY

observed in *modSP[1]* flies upon infection with *L. monocytogenes*, whereas the level of *Diptericin* (*Dpt*, regulated by the Imd pathway) was similar or even higher to that observed in wild-type flies (Fig. 3 *C* and *D*). In addition, *modSP[1]* did not affect the expression of the antibacterial peptide *Dpt* upon infection with the Gram-negative bacterium *E. carotovora* (Fig. 3*E*). The levels observed were similar or even higher than those observed in wild-type flies. In conclusion, our survival and antimicrobial peptide gene expression analyses strongly suggest a role of ModSP in the activation of the Toll pathway by Gram-positive bacteria.

ModSP Overexpression Activates the Toll Pathway. We next performed a series of experiments to further assess that the *modSP* gene is responsible for the immune phenotype observed in *modSP[1]* flies. First, similar phenotypes, both in terms of immune susceptibility and *Drs* expression, were observed in flies carrying the *modSP[1]* mutation over deficiencies [*Df(3R)P10, Df(3R)Spf*] uncovering the *modSP* locus. This reduces the possibility that the phenotype is because of an associated mutation and indicates that *modSP[1]* behaves as an amorphic mutation. We also generated fly lines overexpressing a full-length *modSP* cDNA under the control of an UAS element. Expression of *modSP* in whole organisms using the ubiquitous driver *daughterless* (*da-Gal4*) led to larval lethality. Importantly, a high level of *Drs* expression was observed in flies overexpressing *modSP* in fat body and hemocytes (genotype: *c564-Gal4, UAS-modSP*). This level reached 60% of that observed in wild-type flies collected 16 h after infection with *M. luteus* (Fig. S3*A*), indicating that overexpression of *modSP* is sufficient to activate the Toll pathway. Finally, 2 *modSP* RNAi fly lines from the Vienna *Drosophila* RNAi Center were used to specifically knockdown *modSP* expression in vivo. The levels of *modSP* transcripts were depleted to 30%–40% of the wild-type level in *modSP-RNAi* flies (genotype: *da-Gal4, UAS-modSP-IR*) (Fig. S2). Importantly, these flies were more susceptible upon septic injury with *E. faecalis* and had reduced *Drs* expression upon *M. luteus* infection (Fig. S3 *B* and *C*). Together, these experiments demonstrate that ModSP is essential for Toll activation by Gram-positive bacteria and that overexpression of full-length ModSP is sufficient to activate the Toll pathway.

ModSP Functions Between PGRP-SA and Grass in Activating the Toll Pathway. We next performed epistasis experiments to determine the position of ModSP in the cascade leading to Toll activation by Gram-positive bacteria. Detection of Gram-positive bacteria is mediated via the recognition of lysine-type peptidoglycan by PGRP-SA and GNBP1 (7–9). Simultaneous overexpression of *GNBP1* and *PGRP-SA* triggers the Toll pathway, resulting in a constitutive expression of *Drs* in the absence of an immune challenge (8). The *Drs* expression induced by the overexpression of both *GNBP1* and *PGRP-SA* was blocked by the *modSP[1]* mutation (Fig. 4*A*). Conversely, the *PGRP-SA[seml]* mutation did not block the high expression of *Drs* induced by the overexpression of *modSP* (Fig. 4*D*). Together these experiments indicate that ModSP functions downstream of PGRP-SA. We further observed that *modSP[1]* mutation did not block the induction of *Drs* expression upon overexpression of a matured and constitutively active form of Spz (*UAS-spz**), or SPE (*UAS-SPE**), which is in agreement with a function upstream of Spz and SPE (Fig. 4 *B* and *C*). Importantly, the *Herrade* mutation that affects the *grass* SP gene (11) fully blocked *Drs* expression induced by overexpression of *modSP* (Fig. 4*E*). Altogether, our epistatic analysis indicates that ModSP functions downstream of PGRP-SA and GNBP1 and upstream of Grass in the pathway that links Gram-positive bacterial recognition to Toll activation.

ModSP Mediates Toll Activation by GNBP3 upon Fungal Infection. Detection of yeast is mediated via the recognition of β-glucans by GNBP3. A loss-of-function mutation in *GNBP3*, *GNBP3[hades]*, induces a phenotype of reduced survival to the yeast *Candida albicans* and deficient Toll activation. Conversely, overexpression of *GNBP3*
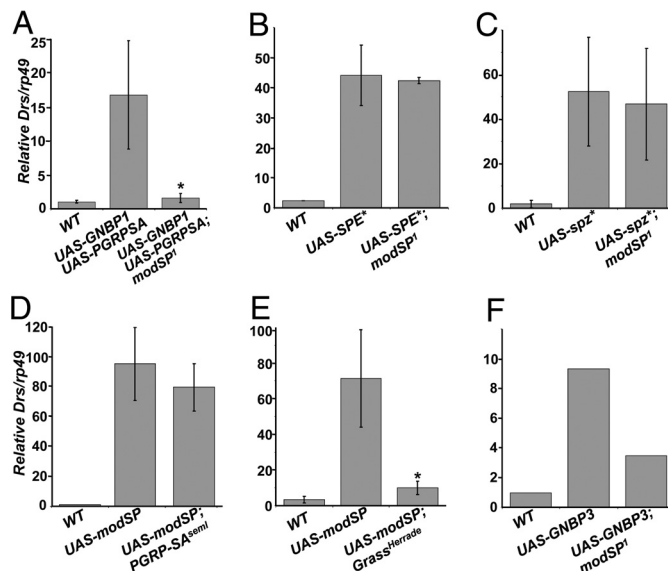
Fig. 4. ModSP functions downstream of PRRs and upstream of Grass. (*A*) *Drs* expression induced by overexpressing *GNBP1* and *PGRP-SA* was reduced by the *modSP[1]* mutation. (*B* and *C*) Overexpression of *UAS-SPE** (*B*), *UAS-spz** (*C*), induced constitutive *Drs* expression that was independent of *modSP*. (*D* and *E*) Overexpression of *UAS-modSP* induced constitutive *Drs* expression that was independent of *PGRP-SA[seml]* (*D*) but dependent on *Grass* (*E*). (*F*) *Drs* expression induced by overexpressing *GNBP3* was reduced by the *modSP[1]* mutation. *Drs* was measured in the absence of immune challenge. To indicate the level of *Drs* obtained in the different conditions, the *Drs/rp49* ratio was normalized to the levels observed in flies collected 16 h after septic injury with *M. luteus* (fixed to 100%). Results are representative of 3 experiments, except panel *F* that could be obtained only once. Values significantly different in a *t* test for $P < 0.05$ are denoted by *. The *C564-Gal4* driver was used to drive UAS transgenes in these experiments.

triggers the Toll pathway, resulting in a constitutive expression of *Drs* in the absence of an immune challenge (10). We thus investigated whether ModSP could also function downstream of GNPB3 in the recognition of yeast. In support of this hypothesis, we observed that *modSP[1]* deficient flies rapidly succumbed to *C. albicans* infection (Fig. 2*D*). In addition, *modSP[1]* flies displayed reduced *Drs* induction upon septic injury with *C. albicans* (Fig. 3*F*). The *modSP[1]* phenotype was similar, albeit weaker, to that generated by a null mutation in *spz*. The relevance of GNBP3 to Toll activation is particularly prominent when flies are injected with dead *C. albicans*, a condition limiting the activation of the Toll pathway through Psh (11). *Drs* expression was also strongly reduced by the *modSP[1]* mutation upon injection of dead *C. albicans* (Fig. 3*H*).

We next investigated the interaction between GNBP3 and ModSP. Unfortunately, we were not able to reproducibly induce *Drs* by overexpression of GNBP3 as previously described (10). Nevertheless, we performed 1 successful experiment in which *Drs* was induced upon GNBP3 overexpression to 10% of the level reached in wild-type flies collected 16 h after infection with *M. luteus*. In this experiment, we observed a reduction of the GNBP3 mediated *Drs* induction by the *modSP[1]* mutation (Fig. 4*F*).

Collectively, these experiments demonstrate a requirement of ModSP in the activation of the Toll pathway by both yeast and Gram-positive bacteria.

ModSP Is Not Involved in the Sensing of Protease Activity by the Psh Pathway. In *Drosophila*, injections of proteases of fungal or bacterial origin activate the Toll pathway in a Psh-dependent manner (10, 11). To investigate whether ModSP interferes with Psh-dependent Toll pathway activation, we first monitored the level of *Drs* expres-
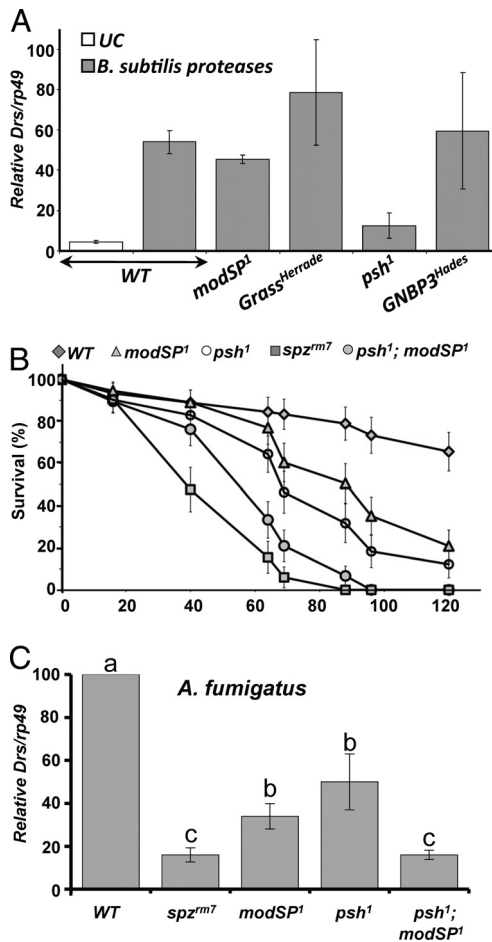
Fig. 5. ModSP and Psh define 2 distinct pathways that cooperatively activate Toll. (*A*) *Drs* gene expression was monitored by RT-qPCR with total RNA extracted from wild-type, *GNBP3^Hades^*, *Grass^Herrade^*, *psh^1^*, and *modSP^1^* females collected 16 h after injection with *B. subtilis* derived proteases. Each graph represents the mean of 4 independent experiments with standard deviation. UC, unchallenged flies. (*B* and *C*) The survival rates (*B*) and the level of *Drs* expression (*C*) of *modSP^1^* and *modSP^1^*; *psh^1^* double mutants infected with *A. fumigatus* were compared with wild-type, *psh^1^*, and *spz^rm7^* flies. Data were analyzed by ANOVA and means were separated for significance according to Fisher's protected LSD at $P < 0.05$; ($F = 63.25$, df = 4, $P < 0.0001$).

sion upon injection of proteases derived from *Bacillus subtilis* or *Aspergillus oryzae*. Unlike the *psh^1^* mutation, *modSP^1^* did not impair Toll activation in these experimental conditions (Fig. 5*A* and Fig. S4).

We next compared the response of *modSP^1^* and *psh^1^* flies to the injection of spores from *A. fumigatus*, a filamentous fungus known to produce proteases. Both *modSP^1^* and *psh^1^* flies exhibited a moderate increase of susceptibility to *A. fumigatus* infection compared with wild-type (Fig. 5*B*). In addition, the level of Toll activity in *modSP^1^* flies and to a lesser extent in *psh^1^* flies was reduced, although generally higher than levels observed in *spz^rm7^* mutants (Fig. 5*C*). Importantly, in flies double mutant for *psh^1^* and *modSP^1^*, the Toll pathway activation by *A. fumigatus* was reduced to a level comparable to the *spz^rm7^* mutant. Correlatively, *psh^1^*; *modSP^1^* double mutants exhibited an increased susceptibility to *A. fumigatus* compared with single mutant. These data indicate that the ModSP and the Psh branches of the Toll pathway contribute independently to the resistance to this fungus. We conclude that, like Grass (11), ModSP does not interfere with the Psh mediated activation of the Toll pathway.

ModSP Exhibits a High Level of Autoproteolysis and Does Not Cleave Grass. To gain insight into the activation mechanisms that regulate ModSP, we produced recombinant His-tagged ModSP and Grass proteins using the insect Sf9 cells. When we purified the recombinant ModSP-containing fractions and analyzed them by SDS-PAGE, we observed 4 major bands corresponding to ModSP. N-terminal sequencing showed that sequence of bands 1 and 2 matched the N terminus of ModSP, whereas the sequence of band 3 and 4 corresponded to the catalytic domain of ModSP (Fig. 6*A*). Western blot analysis using a polyclonal antibody against the N-terminal domain or the central domain of the protein confirmed that band 1 corresponded to full-length zymogen and that the other bands corresponded to either the N-terminal fragment (band 2) or the C-terminal part (band 3 and 4) resulting from the cleavage of ModSP. These results suggest that, as soon as it is expressed, ModSP is autoactivated. Of note, under the same conditions, *T. molitor* MSP did not show any autoactivation. Our results also reveal that *Drosophila* ModSP and *Tm*-MSP share the same cleavage site (Fig. 6*A* and Fig. S1). The presence of 2 processed C-terminal fragments of ModSP could either be because of the presence of a posttranslational modification or an additional cleavage at the C terminus of ModSP. We next investigated whether ModSP could cleave Grass in vitro. Fig. 6*B* shows that Grass was not cleaved when incubated with recombinant ModSP and $CaCl_2$. In the same condition, *Tm*-SAE was cleaved by *Tm*-MSP. However, in absence of a known ModSP substrate, we cannot be fully sure that the 28 kDa form of ModSP we detected is active. Nevertheless, this result suggests that ModSP does not cleave the Grass zymogen. Additionally, the direct activation of Grass by ModSP is unlikely since the predicted cleavage of Grass requires a trypsin-like enzyme although ModSP exhibits a chymotrypsin-like specificity (Fig. S5).

Discussion

In this study, we have reported the generation of a null mutation of the gene encoding the SP ModSP and demonstrated its essential role in the activation of the Toll pathway by PRRs. This was illustrated by the observation that *modSP* deficient flies fail to activate the Toll pathway in response to either Gram-positive bacteria or yeast. Epistatic analyses demonstrated that ModSP acts downstream of PGRP-SA, but upstream of the SP Grass. Importantly, we showed that ModSP does not participate in the Psh-dependent branch of the Toll pathway as shown by the wild-type activation of this pathway in *modSP^1^* flies upon injection of bacterial proteases. The *modSP^1^* phenotype was very similar to that of a loss of function *Grass* mutant, suggesting that these SP act in a linear pathway connecting microbe recognition by PRR to the activation of Spz by SPE. The analysis of *psh^1^*; *modSP^1^* double mutant flies indicates a synergistic action of the ModSP and Psh pathways in the response against filamentous fungi, which might be detected through both host PRRs and their virulence factors. Collectively, our results confirm the model of an activation of Toll by 2 extracellular pathways: A PRR-dependent pathway and a Psh-dependent pathway (11). We extend this model by showing that Grass and ModSP function in a common SP cascade. In addition, the apical position of ModSP suggests a direct branching of signals from secreted PRRs to the ModSP-Grass pathway. Biochemical analyses in *T. molitor* indicate that *Tm*-MSP interacts directly with the PRR complexes involved in the sensing of peptidoglycan (14) or glucan (30). Although it was not formally demonstrated in the present study, it seems likely that *Drosophila* ModSP acts as the most upstream SP directly activated by secreted PRRs.

ModSP has a number of structural features that make it unique among *Drosophila* SP and suggest its critical role in the initial events leading to the activation of the proteolytic cascade upstream of Toll. First, ModSP does not contain a Clip-domain in contrast to Grass and SPE. The function of the Clip-domain is still unknown but its presence in many SP that function in signaling cascades suggests a regulatory role (19). This indicates that ModSP activation differs

from the chain reaction activating Grass or SPE. Second, ModSP contains additional domains such as the CCP and LDLa motifs in its N-terminal extremity. Consistent with the presence of the LDLa domain, a ModSP-GFP fusion protein was secreted and found bound to the surface of lipid vesicles that circulate in the hemolymph (Fig. S6). We speculate that the association of ModSP to vesicles is important to nucleate the activation of downstream SP.

In the beetle, *T. molitor*, *Tm*-MSP is activated upon formation of a tri-molecular complex composed of PGRP-SA, GNBP1, and lysine-type peptidoglycan (14). Although the possibility of a SP other than ModSP functioning as the initial protease in the *Drosophila* Toll pathway cannot be ruled out, our epistatic analysis suggests an apical role of ModSP. We show that overexpression of a full-length version of ModSP is sufficient to reach a high level of Toll activation, in contrast to other SP that generally require the overexpression of a preactivated form to fully induce the cascade (5, 20). These observations favor a model in which ModSP directly interacts with GNBP3 or GNBP1/PGRP-SA recognition complexes. Recruitment of ModSP by PRRs would increase its local concentration, a situation sufficient for its autoproteolysis. In agreement with this hypothesis, a recombinant form of ModSP produced in baculovirus appears to be unstable as a zymogen, presumably because of a high level of autoactivation (Fig. 6*A*). Unfortunately, this high level of autoproteolysis did not permit in vitro reconstitution experiments using ModSP, GNBP1, and PGRP-SA as performed with *Tm*-MSP (14). Nevertheless, we observed a higher level of ModSP activation when the protein was incubated with GNBP1 and PGRP-SA. The most parsimonious model is that ModSP would interact with GNBP1 or GNBP3, the common protein family members found in the 2 recognition complexes. The exact contribution of GNBP1 to sensing Gram-positive bacteria is currently debated. It was recently proposed that GNBP1 functions upstream of PGRP-SA by cleaving peptidoglycan, a step required for an optimal binding of PGRP-SA to peptidoglycan (21). In contrast, *T. molitor* GNBP1 is recruited subsequent to the binding of PGRP-SA to peptidoglycan and is required for the interaction with *Tm*-MSP (22). The implication of ModSP in sensing of fungi through GNBP3 and peptidoglycan through GNBP1/PGRP-SA suggests a similar mechanism of activation of this SP and *Tm*-MSP. This would favor a role of GNBP1 as a linker between PGRP-SA and ModSP. In accordance with this model, recombinant full-length *Drosophila and Tenebrio* GNBP1 did not exhibit any enzymatic activity toward peptidoglycan in vitro (Fig. S7).

The participation of ModSP and SPE in an extracellular pathway linking PRR recognition to Spz activation in both *T. molitor* (Coleoptera) and *D. melanogaster* (Diptera), which diverged ≈250 million years ago, demonstrates the conservation of this extracellular signaling module in insects. In the lepidopteran *Manduca sexta*, the hemolymph protein 14 (*Ms*-HP14) contains a domain arrangement very similar to that of ModSP (Fig. S1) and regulates the melanization cascade in response to microbial infection (23, 24). Collectively, biochemical studies performed in *T. molitor* and *M. sexta*, and genetic analysis in *D. melanogaster*, reveal striking similarities in the mechanisms underlying SP activation by PRRs (Fig. 6*C*). All involve the sequential activation of an apical modular SP (that displays a certain level of autoactivation) and clip-domain proteases. Interestingly, a similar organization is also observed in the proteolytic cascade that regulates Toll during dorso-ventral patterning of the embryo. *Gastrulation Defective*, the apical SP, contains von Willebrand domains and is also thought to be autoactivated, although the precise mechanism that triggers its activation has not been determined (25, 26). These features are also reminiscent of the complement activation by the lectin pathway in mammals in which the recognition of carbohydrate by the mannose binding lectin (MBL) leads to the autoactivation of MBL-associated serine proteases (MASPs). MASPs are also modular proteases with CUB, CCP, and EGF domains in their N terminus (27). Thus,
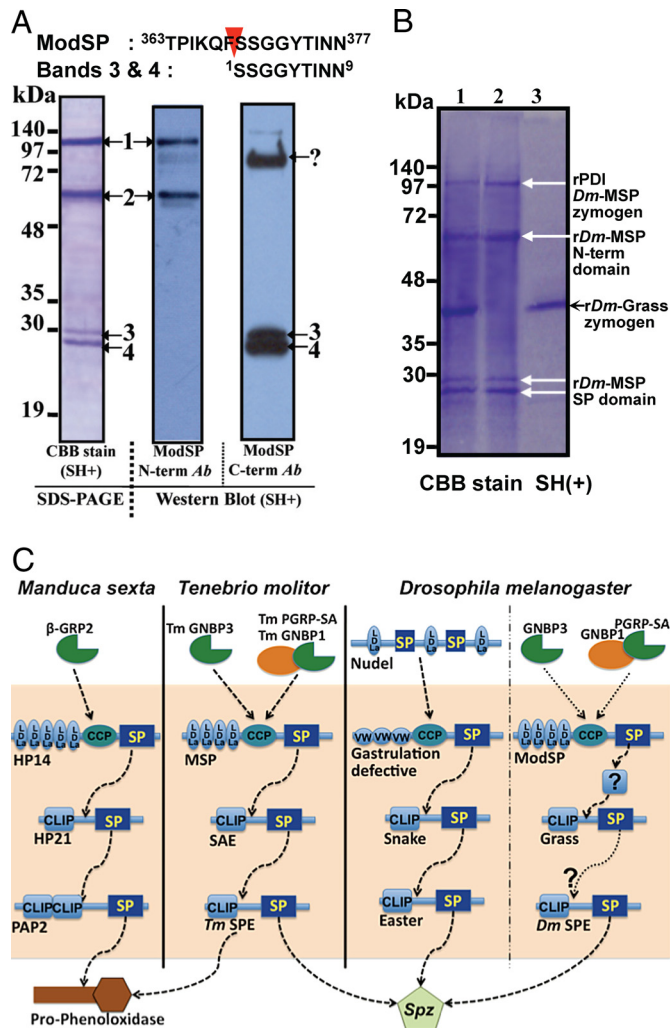
Fig. 6. ModSP exhibits a high level of autoproteolysis. (*A*) SDS-PAGE gel using purified recombinant rPDI-fusioned ModSP-containing fractions reveals 4 bands corresponding to ModSP as indicated by N-terminal sequencing (*Top*). Bands 1 and 2 matched the N terminus of ModSP, whereas the sequence of band 3 and 4 corresponded to the catalytic domain of ModSP. Western blot analysis using a polyclonal antibody against the N-terminal domain (*Middle*) or the central domain (*Right*) of the protein confirmed that band 1 corresponded to full-length zymogen and that the other bands corresponded to either the N-terminal fragment (band 2) or the C-terminal part (band 3 and 4) resulting from the cleavage of ModSP. Red arrow-head indicates the cleavage site of catalytic serine protease domain of ModSP (*Top*). The question mark (?) indicates a nonspecific band. (*B*) rModSP does not cleave rGrass in vitro. 1, rDm-Grass zymogen was incubated with rDm-MSP and CaCl₂ at 30 °C for 1 h. 2, 3 indicate the purities of rPDI-fusioned ModSP, a mixture of active and zymogen forms, and rGrass zymogen under reducing conditions, respectively. The gel was stained by Coomassie brilliant blue R-250 (CBB). (*C*) Extracellular cascades of serine proteases in insects. Biochemical studies performed in *T. molitor* and *M. sexta*, and genetic analysis in *D. melanogaster*, reveal striking similarities in the mechanisms underlying serine protease activation by PRRs. In *M. sexta*, HP-14 is converted to a 2-chain active form in the presence of β-1,3-glucan and the β-1,3-glucan recognition protein 2 (βGRP2, a protein related to GNBP3) (23, 24). Binding of β-1,3-glucan and βGRP2 triggers the autoactivation of proHP14 into HP14 that then processes HP21. In turn, HP21 cleaves proPAP-2 into PAP2 that finally activates proPO. Hence, the mechanism of activation of HP14 in response to glucan is highly reminiscent of the mode of activation of *Tm*-MSP, with the exception that HP14 can be activated upon direct binding to peptidoglycan during bacterial infection (24). This suggests that the PGRP-SA/GNBP1 complex involved in sensing Gram-positive bacteria has been added on to a more ancient cascade core during evolution. Interestingly, a similar organization is also observed in the proteolytic cascade that regulates Toll during dorso-ventral patterning of the embryo.

Buchon et al.

genetic and biochemical analyses now reveal a similar level of organization for various proteolytic cascades in insects and increase our understanding as to how these cascades have evolved to fulfill diverse developmental or immune functions.

We are still far from a complete understanding of the precise sequence of events leading to Toll pathway activation, as the task is complicated by the high number of SP encoded in the *Drosophila* genome (16). Our analysis strongly suggests the existence of an intermediate SP acting between ModSP and Grass. Future works should identify this SP and determine the interaction between Grass and SPE. To date, no serpin regulating SP involved in the PRR-dependent pathways has been identified despite the critical role of this family in the negative control of proteolytic cascades. Further experiments combining genetics, biochemistry, and cell biology are required to identify additional components of this cascade and to clarify in vivo how and where proteolytic cascades downstream of PGRP-SA or GNBP3 are activated in the hemolymph compartment.

Materials and Methods

Fly Stocks. A deletion of the modSP locus was obtained by homologous recombination (28). DNA sequences (of 4.2 kb and 4.8 kb) flanking the 5′ and 3′ parts, respectively, of the modSP locus (Fig. 1*B*) were cloned in the p(W25) vector (28). modSP sequence (3 kb; 3R: from nucleotide 12478339 to 12481351) was replaced by the white+ gene.

spz^{rm7}, rel^{E20}, psh^1, $GNBP3^{hades}$, $PGRP-SA^{seml}$, and $Grass^{Herrade}$ null alleles are described elsewhere (4, 7, 10, 11, 20). The *c564* line specifically expresses GAL4 in adult fat body and hemocytes. Oregon^R flies were used as wild-type controls, if not indicated differently. Epistatic analyses were performed by standard genetic crosses. The *UAS-GNBP3*, *UAS-PGRP-SA*, *UAS-GNBP1*, *UAS-SPE**, *UAS-spz**, are described elsewhere (4, 5, 7, 8, 10, 11, 20); an asterisk (*) indicates that the constitutively activated form was used. A full-length cDNA of *modSP* (using the *CG31217*_cDNA gold LD43740 from DGRC) was placed downstream of UAS sequence using the pUASt vector. The *UAS-ModSP-GFP* transgene was obtained by a fusion of GFP to the C terminus of ModSP inserted in the pDONR221 Gateway entry clone (Invitrogen) and finally subcloned in the pTWG transgenesis vector. *Drosophila* stocks and crosses were maintained at 25 °C using standard fly medium. F1 progeny carrying both the *UAS-construct* and the *GAL4* driver were transferred to 29 °C at late larval stage for optimal efficiency of the UAS/GAL4 system except for crosses involving *UAS-SPE** or *UAS-modSP* that were performed at 20 °C.

Infection and Survival Experiments. Bacterial and fungal infections were performed by pricking adults in the thorax with a thin needle previously dipped into a concentrated pellet of a microbial culture (OD = 200). Bacterial and fungal strains have been described in ref. 9. Flies were counted at different time points to monitor survival. Each experiment was performed with ≈80 flies for each genotype. Injection with *M. luteus* and *E. faecalis* peptidoglycan and its production were previously described (17). Injection of proteases from *B. subtilis* (>16 U/g, P5985; Sigma–Aldrich) or *A. oryzea* (>500 U/g, P6110; Sigma–Aldrich) was performed with a Nanoject apparatus using a concentration of 1/1500 as described in ref. 11.

RT-qPCR. *Drs*, *Dpt*, *ModSP*, *ModSP* flanking genes and *rp49* mRNA quantification by RT-qPCR was performed as described in ref. 9. Primers used to monitor mRNA expression can be obtained upon request.

Production of Recombinant ModSP Protein. ModSP was expressed in Sf9 cells as a fusion partner with the *Bombyx mori* protein disulfide isomerase (bPDI) as previously described (29). Two rabbit polyclonal antibodies were raised using keyhole limpet hemocyanin (KLH)-conjugated synthetic peptide against the N-terminal and the C-terminal parts of ModSP respectively. Additional information is provided in the *SI Text*.

1. Lemaitre B, Hoffmann J (2007) The host defense of *Drosophila melanogaster. Annu Rev Immunol* 25:697–743.
2. LeMosy EK, Hong CC, Hashimoto C (1999) Signal transduction by a protease cascade. *Trends Cell Biol* 9:102–107.
3. Jang IH, Nam HJ, Lee WJ (2008) CLIP-domain serine proteases in *Drosophila* innate immunity. *BMB Rep* 41:102–107.
4. 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.
5. Jang IW, et al. (2006) A Spätzle-processing enzyme required for toll signaling activation in *Drosophila* innate immunity. *Dev Cell* 10:45–55.
6. Charroux B, Rival T, Narbonne-Reveau K, Royet J (2009) Bacterial detection by *Drosophila* peptidoglycan recognition proteins. *Microbes Infect* 2009 Apr 1. [Epub ahead of print]
7. Michel T, Reichhart JM, Hoffmann JA, Royet J (2001) *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. *Nature* 414:756–759.
8. Gobert V, et al. (2003) Dual activation of the *Drosophila* toll pathway by two pattern recognition receptors. *Science* 302:2126–2130.
9. Pili-Floury S, et al. (2004) In vivo RNAi analysis reveals an unexpected role for GNBP1 in the defense against Gram-positive bacterial infection in *Drosophila* adults. *J Biol Chem* 279:12848–12853.
10. Gottar M, et al. (2007) Dual detection of fungal infections in *Drosophila* by recognition of glucans and sensing of virulence factors. *Cell* 127:1425–1437.
11. El Chamy L, Leclerc V, Caldelari I, Reichhart JM (2008) Sensing of 'danger signals' and pathogen-associated molecular patterns defines binary signaling pathways 'upstream' of Toll. *Nat Immunol* 9:1165–1170.
12. Tang H, Kambris Z, Lemaitre B, Hashimoto C (2008) A serpin that regulates immune melanization in the respiratory system of *Drosophila*. *Dev Cell* 15:617–626.
13. Kan H, et al. (2008) Molecular control of phenoloxidase-induced melanin synthesis in an insect. *J Biol Chem* 283:25316–25323.
14. 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.
15. De Gregorio E, Spellman PT, Tzou P, Rubin GM, Lemaitre B (2002) The Toll and Imd pathways are the major regulators of the immune response in *Drosophila*. *EMBO J* 21:2568–2579.
16. Ross J, Jiang H, Kanost MR, Wang Y (2003) Serine proteases and their homologs in the *Drosophila melanogaster* genome: An initial analysis of sequence conservation and phylogenetic relationships. *Gene* 304:117–131.
17. Leulier F, et al. (2003) The *Drosophila* immune system detects bacteria through specific peptidoglycan recognition. *Nat Immunol* 4:478–484.
18. Mansfield BE, Dionne MS, Schneider DS, Freitag NE (2003) Exploration of host-pathogen interactions using *Listeria monocytogenes* and *Drosophila melanogaster*. *Cell Microbiol* 5:901–911.
19. Piao S, et al. (2005) Crystal structure of a clip-domain serine protease and functional roles of the clip domains. *EMBO J* 24:4404–4414.
20. Ligoxygakis P, Pelte N, Hoffmann JA, Reichhart JM (2002) Activation of *Drosophila* Toll during fungal infection by a blood serine protease. *Science* 297:114–116.
21. Wang L, et al. (2006) Sensing of Gram-positive bacteria in *Drosophila*: GNBP1 is needed to process and present peptidoglycan to PGRP-SA. *EMBO J* 25:5005–5014.
22. 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.
23. Wang Y, Jiang H (2006) Interaction of beta-1,3-glucan with its recognition protein activates hemolymph proteinase 14, an initiation enzyme of the prophenoloxidase activation system in *Manduca sexta*. *J Biol Chem* 281:9271–9278.
24. Wang Y, Jiang H (2007) Reconstitution of a branch of the *Manduca sexta* prophenoloxidase activation cascade in vitro: Snake-like hemolymph proteinase 21 (HP21) cleaved by HP14 activates prophenoloxidase-activating proteinase-2 precursor. *Insect Biochem Mol Biol* 37:1015–1025.
25. Han JH, Lee SH, Tan YQ, LeMosy EK, Hashimoto C (2000) Gastrulation defective is a serine protease involved in activating the receptor toll to polarize the *Drosophila* embryo. *Proc Natl Acad Sci USA* 97:9093–9097.
26. Dissing M, Giordano H, DeLotto R (2001) Autoproteolysis and feedback in a protease cascade directing *Drosophila* dorsal-ventral cell fate. *EMBO J* 20:2387–2393.
27. Duncan RC, Wijeyewickrema LC, Pike RN (2008) The initiating proteases of the complement system: Controlling the cleavage. *Biochimie* 90:387–395.
28. Gong WJ, Golic KG (2004) Genomic deletions of the *Drosophila melanogaster* Hsp70 genes. *Genetics* 168:1467–1476.
29. Goo TW, et al. (2008) *Bombyx mori* protein disulfide isomerase enhances the production of nuecin, an antibacterial protein. *BMB Rep* 41:400–403.
30. Roh KB, et al. (2009) Proteolytic cascade for the activation of the insect Toll pathway induced by the fungal cell wall component. *J Biol Chem*, 10.10741/jbc.M109.007419.

IMMUNOLOGY

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 pFAST-Bac 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.

```
Dm ModSP : ----MQLISFLSNPLFFCALLLKF----------------------------------------------RTIFAACDSSQFECDNGSCISQYDVCNGEKNCPD :  54
Tm MSP   : MCNVRTLLQVICLSLIVIQTVDSYSFALS-----------------------------------------KFTRIRRQARRTCTSTEFACKSGECIDEDKECDGIVDCTD :  69
Ms HP14  : MYTDLFFIVLATVPALTSVLKEEINYCSPDEFSCGDGSCVSFSAFCDGKRDCFNGADEACTIGENALSTDTVLNRSRRQLSNCRISQWQCKDGSCINFDGKCDGIVDCPD : 110
Tc ModSP : MCNA-TLIQVICL-FLAIQIVPSFSFHQS------------------------------------------NIT--RREAK-ECPRNSFACKSGECIDEDKECDGGVDCKD :  64

Dm ModSP : GSDETALTCVSQRQHCTKPYFQCTYGACVIGTAGCNGVNECADGSDETRLRCGNEDDIRQHDRRLQGNCKENEFKCPSGICLDKSNFLCDGKDDCADGTGFDESVELCGH : 164
Tm MSP   : ASDETNACHR---IKCPNYLFRCKYGACINPDLECDGKPDCMDGSDEKTSKCKP--------DDSSPECKANEFRCSSGQCIP-EDFKCDGKAECKDNS--DEIRATCWN : 165
Ms HP14  : ASDETHALCRE--RQCQYNWFRCTYGACVDGTAPCNGVQDCADNSDELLPRCRN-----------ETEEIRGQFKCLDGRFIA-AYKHCDGVADCADGS--DETLRSCAG : 204
Tc ModSP : ASDESNACAR---IKCPISAFRCDYGACISADLECDGKPDCRDGSDEKTPNCQII-------DETSPICKSNEFRCSSGBCID-EDNKCDGIAQCSDRS--DEIRATCWN : 161

Dm ModSP : MECPAYSFKCGTGGCISGSLSCNGENDCYDGSDEAPLLCNTTKKVTTPVVTETP---LELLGCPLPLGDERPILTGDG--SRVLTGPITRG--TVRFSCKQGYVLEGEES : 267
Tm MSP   : VRCPGFTHKCKYGACVSGNAECNGIVECFDGSDEDPAICKTKPTPRPTPTPGTPGPQPTQGGCVLPNHPEFGEWQVYGIPGQFSPGMVIRAGATLRIQCKKRYKLEGKNA : 275
Ms HP14  : KTCLSYLFQCAYGACVDKDSDCNGIRECVDGSDEADDLCADRNTSVQP---------VKEGACVLPEYPEHGGYVVSG-MKNAKPGMSFEF-VSLNYTCNKGYGILGRND : 303
Tc ModSP : LRCPIYSYKCKYGACVSGNAECNGKIECQDGSDEDPNICKNSTVLTPTPSPVVPRPG-ARGRCVLPNHPEFGKWSIFG-PENLSPGATVNPGTILNIVCQNGYKLEGNSV : 269

Dm ModSP : SYCAKNKWSTSTIPKCVKYCSTAGEFDGYSTKALCTHNGQQVECRKPFHPPGTEVKFVCSTGFKTLSPLPEMRCMKG--GYWNRGRQRCEQCQGQLATPIKQFSSGGYTI : 375
Tm MSP   : IFCENGKWSDA-VGHCLKLCPSIQSTSTMRVTCIYNKHEETENCTE--AVEGTLVRFDC-APFYEDLGLSRHPIHICRDGSWDQRRPECTPVCGQKSVNAQTLIVNGKPV : 381
Ms HP14  : VTCTNGWWNFTPLPKCTRFCHLKEHVSVEYKCVLSGSSQGFRTCNEY-EPDGTLVQPECRKPNYYSETTFRYMSCIGAVGAWN-YVAKCIPECGTVTPNGTELVLGGERA : 411
Tc ModSP : IYCANGQWTEN-IGVCLKLCSPLSSTKVMTVTCTYPNNKGBGNCTN--AIDGTLARFKC-EPLYEDPNLQRIPGRVCRDGTWDNSSPECVPVCGQKSVEVQKLIVNGKTA : 375

Dm ModSP : NNTVVPWHVGLYVWHNEKDYHFQCGGSLLTPDLVITAAHCVYDEG-TRLPYSYDTFRVIAAKFYRNYG--ETTPEEKRRDVRLIEIAPGYKGRTENYYQDLALLTLDEPF : 482
Tm MSP   : KKGDYPWQVALYTLNDK---ELICGGSLLNQRVVLTAAHCITD--DKGKLLSKENYMVAVGKYYRPFNDSRDRNEAQFSEVKHMFIPELYKGSTQNYVGDIAILVIRVTF : 486
Ms HP14  : QFGELPWQAGIYTKNTRP-YMQICGGALISSTVVLSAAHCFWVN---DAVTPKEEYAVALGKLYRPWQPKMVVEKDQKSEIRDIHISPYFLGRTNNYQNDIAVVILETTI : 517
Tc ModSP : KRGTYPWQAALYTRDKK---ELICGGSLIKLNMIITAAHCVTDQQERAQPLPKENYIVALGKYYRKFDDPRDSKEAQFSELKKIIVNEKYSGPIQNFGSDIALLITSTVF : 482

Dm ModSP : ELSHVIRPICVTFASFAEKESVTDDVQGKFAGWNIEN----------KHELQFVPAVSKSNSVCRRNLRDIQADKFCI-FTQGKSLACQGDSGGGFTSELPTNAFSTWNT : 581
Tm MSP   : TLSRRVQPVCIDYGLKYT--SYTN-EFGYVTGWGYTLQNDKPSDVLKELKVPAVSTEQCSSAIPEDYDIYLTHDKLCAGYLDNGTSVCSGDSGGGL--------VFKFDG : 585
Ms HP14  : VYKPHIRPVCLNFDIQFEKEQLYVGSLGKVAGWGIKDEAGNPSQVLKVVKLPYVDVLQCISQSPQAFRPYITGDKICAGFAN-GTGLCHGDSGGGLS-----FPAVNRLT : 621
Tc ModSP : VPSLRVQPVCMDWNLQFK--IGDDQVFGYVTGWGYTVEGSNPSEELKELKVPLIPESKCRKDLPVDYDRYYTYDKLCAGYLNSNTSVCRGDSGGGL--------VVKRNG : 582

Dm ModSP : ARHFLFGVISNAPNAD----QC-AHSLTVMTNIQHFEDMILNAMNRSVETRS- : 628
Tm MSP   : -RYYVTGIVSLSPQAS--TGGCDTQQYGLYTKVGTYISDFIIKTESQFRP--- : 632 (47%)
Ms HP14  : ERYYLRGIVSTAHTSN---EACNAFALTTFTNILSHEHFIKRFWTDEY----- : 666 (46%)
Tc ModSP : DRFYLTGIVSLSPTSPREIDGCDSQQYGLYTKVAAHTEDFILKKVTEYGINAR : 635 (45%)
```

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.
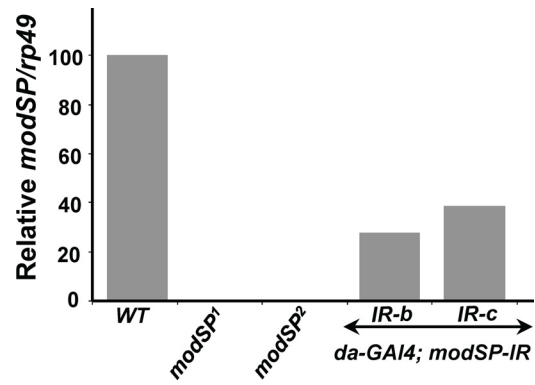
Fig. S2. Relative *modSP* expression after RNAi using the *da-GAL4* driver in combination with *modSP-IR* constructs and in *modSP* mutants. Relative *modSP* expression is shown as a ratio over the expression of the housekeeping gene *rp49*.
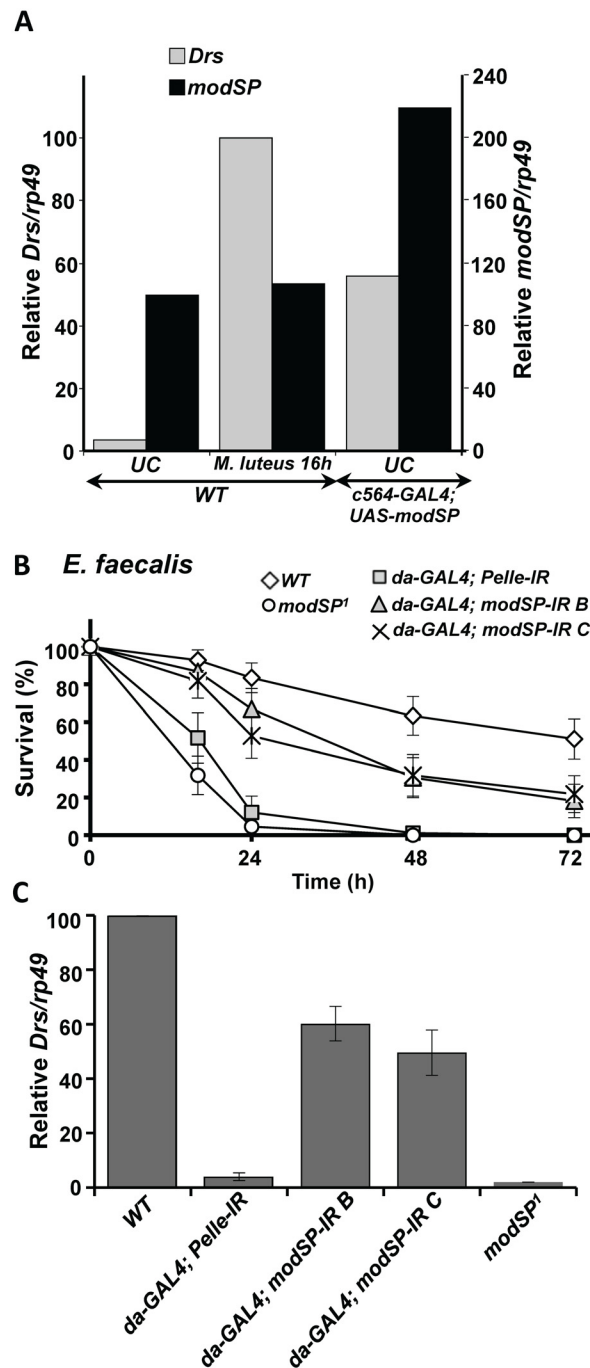
Fig. S3. Silencing of *modSP* by RNAi reduces Toll pathway activity. Overexpression of *modSP* with the fat body Gal4 driver *c564-Gal4* induces a high level of *Drs* expression in absence of challenge. Levels of *Drs* and *modSP* expression were monitored by RT-qPCR and normalized on the expression 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 *modSP¹* 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 *modSP¹* 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).
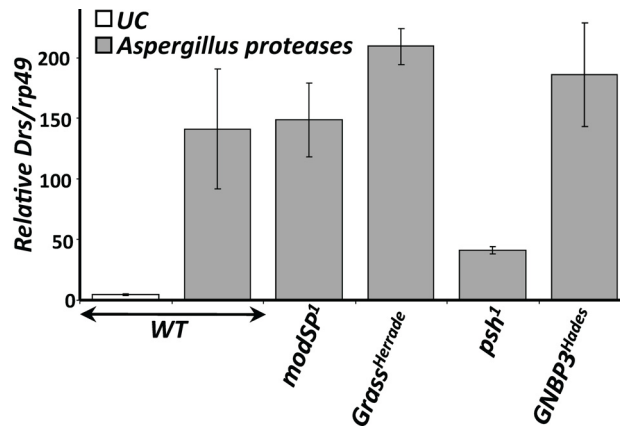
Fig. S4. *modSP¹* does not block Toll activation by proteases derived from *Aspergillus oryzae. Drs* gene expression was monitored by RT-qPCR with total RNA extracted from wild-type, *GNBP3^Hades^*, *Grass^Herrade^*, *psh¹*, and *modSP¹* females collected 16 h after injection with proteases of *Aspergillus*. UC, unchallenged flies.

```
Tm-SAE    : -MLNLNYFTCFVIVLIQLVSSQRFVG---DLCTLESSGAPGVCELFKECKQARDDLQKHQLFPQ-- :  60
Dm-Grass  : MMIASSLAVLYGIAIVSSMGVQSARADYADDCTTPDGDQ-GQCMPFSSCRTIEERLTEAQKAGQKV :  65

Tm-SAE    : -----------QCGYQKNEPIVCCLKKSKRKPGEISLKKCQEYSRLVYEVNRAPVLIINAPNITKN : 115
Dm-Grass  : PADYASYLQKALCGEFNGVRHFCCPSANIQHNSKVMSLFKDENFD-------------------- : 110

Tm-SAE    : ECGHKIIKLIVGGTNATRKEFPHMAVIGFEPQPGDIKWLCGGTVLSKHYILTAAHCLSHQEHGRAR : 181
Dm-Grass  : -CGNFLSQRVSNGYEVKLSSRPWMALLRYQQFGESRFL-CGGAMISERYILTAAHCVHGLQNDLYE : 174

Tm-SAE    : YVRIGVTDLEDTNHRQQLEVEELIPYPEYKSSSHYHDIGLLRLKRSAKLDSFTV------------ : 235
Dm-Grass  : IRLGEHRISTEEDCRQQGRKKKCAPPVVNVGIEKHLIHEKYDARHIMHDIALLKLNRSVPFQKHIK : 240

Tm-SAE    : PACLYRKHDIEAEKAIATGWGHTTWG----GSGSNNLLKVTLDLFDHASCNRSYKNQISRRLKDGI : 297
Dm-Grass  : PICLPITDELKEKAEQISTYFVTGWGTTENGSSSDVLLQANVPLQPRSACSQAYRRAVP------- : 299

Tm-SAE    : IDDIQVCAGSLDDEKDTCQGDSGGPLQIFHESK--DIKCMYDIIGVTSFGKACSG-S-PGVYVRVS : 359
Dm-Grass  : --LSQLCVGGGD-LQDSCKGDSGGPLQAPAQYLGEYAPKMVEFGIVSQGVVTCGQISLPGLYTNVG : 362

Tm-SAE    : QYIGWIEDIVWPENS : 374
Dm-Grass  : EYVQWITDTMASNGL : 377
```

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.
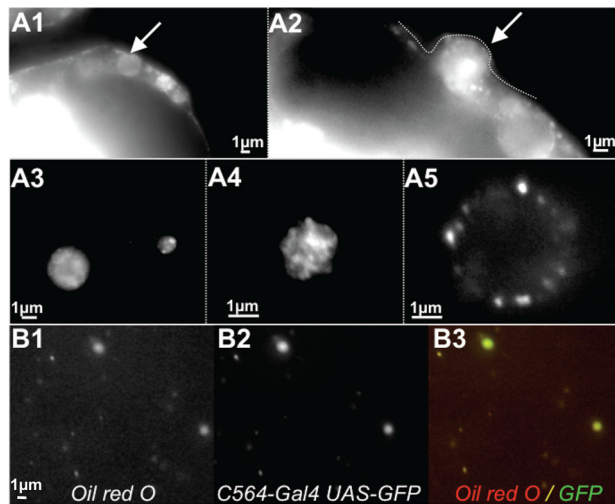
Fig. S6. A ModSP-GFP fusion is secreted in lipid vesicles. Use of flies expressing a GFP tagged version of ModSP (Genotype: *c564-Gal4*; *UAS-ModSP-GFP*) revealed that ModSP-GFP was secreted from the fat body into the hemolymph at the surface of small lipid vesicles of 0.2–5 μm. This unusual mode of 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).
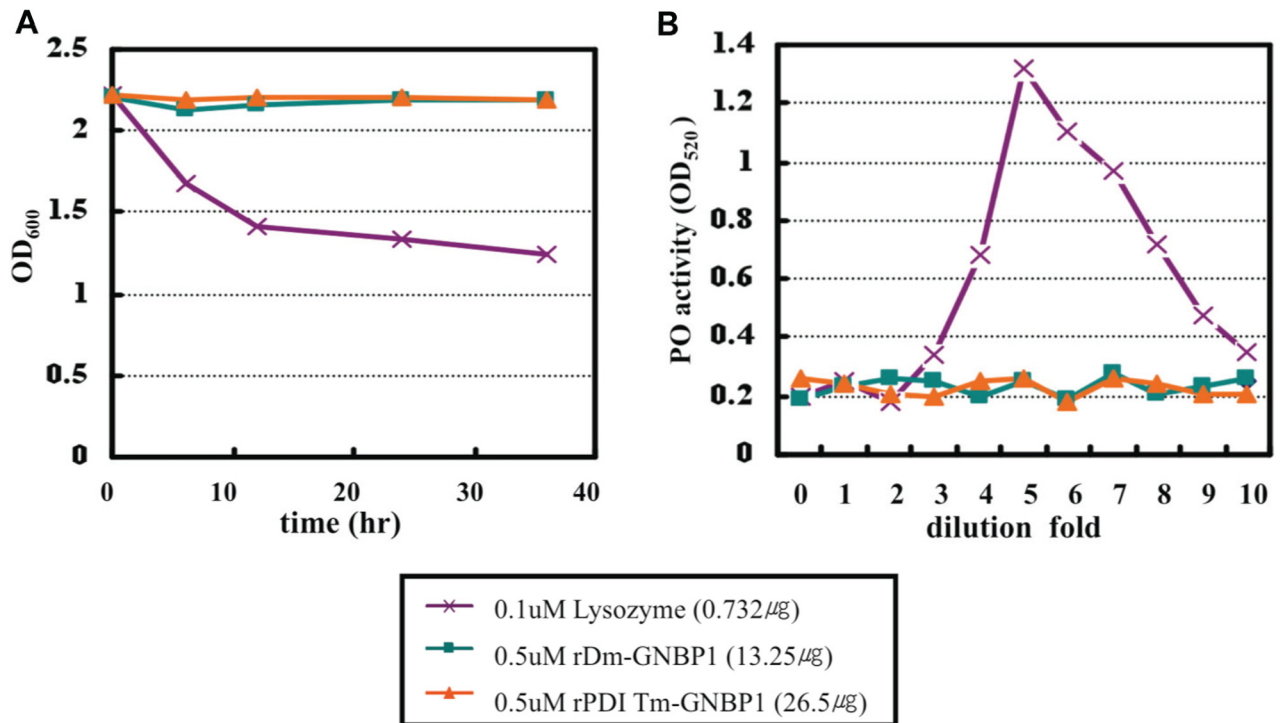
Fig. S7. Recombinant *Tenebrio* and *Drosophila* GNBP1 does not exhibit enzymatic activity against *Micrococcus luteus* lysine-type peptidoglycan. (*A*) Lysine-type peptidoglycan (500 μg) from *M. luteus* was mixed with either rDm-GNBP1, rPDI Tm-GNBP1, or lysozyme (catalog no. L7001; Sigma-Aldrich) in 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.