THE JOURNAL OF IMMUNOLOGY

Tissue- and Ligand-Specific Sensing of Gram-Negative Infection in *Drosophila* by PGRP-LC Isoforms and PGRP-LE

Claudine Neyen, Mickaël Poidevin, Alain Roussel and Bruno Lemaitre

J Immunol 2012; 189:1886-1897; Prepublished online 6 July 2012;
doi: 10.4049/jimmunol.1201022
http://www.jimmunol.org/content/189/4/1886

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
9650 Rockville Pike, Bethesda, MD 20814-3994.

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

Tissue- and Ligand-Specific Sensing of Gram-Negative Infection in *Drosophila* by PGRP-LC Isoforms and PGRP-LE

Claudine Neyen,* Mickaël Poidevin,† Alain Roussel,‡ and Bruno Lemaitre*

The *Drosophila* antimicrobial response is one of the best characterized systems of pattern recognition receptor-mediated defense in metazoans. *Drosophila* senses Gram-negative bacteria via two peptidoglycan recognition proteins (PGRPs), membrane-bound PGRP-LC and secreted/cytosolic PGRP-LE, which relay diaminopimelic acid (DAP)-type peptidoglycan sensing to the Imd signaling pathway. In the case of PGRP-LC, differential splicing of PGRP domain-encoding exons to a common intracellular domain-encoding exon generates three receptor isoforms, which differ in their peptidoglycan binding specificities. In this study, we used Phi31-mediated recombineering to generate fly lines expressing specific isoforms of PGRP-LC and assessed the tissue-specific roles of PGRP-LC isoforms and PGRP-LE in the antibacterial response. Our in vivo studies demonstrate the key role of PGRP-LCx in sensing DAP-type peptidoglycan-containing Gram-negative bacteria or Gram-positive bacilli during systemic infection. We also highlight the contribution of PGRP-LCa/x heterodimers to the systemic immune response to Gram-negative bacteria through sensing of tracheal cytotoxin (TCT), whereas PGRP-LCy may have a minor role in antagonizing the immune response. Our results reveal that both PGRP-LC and PGRP-LE contribute to the intestinal immune response, with a predominant role of cytosolic PGRP-LE in the midgut, the central section of endodermal origin where PGRP-LE is enriched. Our in vivo model also definitively establishes TCT as the long-distance elicitor of systemic immune responses to intestinal bacteria observed in a loss-of-tolerance model. In conclusion, our study delineates how a combination of extracellular sensing by PGRP-LC isoforms and intracellular sensing through PGRP-LE provides sophisticated mechanisms to detect and differentiate between infections by different DAP-type bacteria in *Drosophila*. *The Journal of Immunology*, 2012, 189: 1886–1897.

In animals, the innate immune system detects bacterial infection through the use of germline-encoded pattern recognition receptors (PRRs) that sense pathogen-associated molecular patterns (PAMPs), such as LPS, peptidoglycan, or flagellin (1). After the identification of PRRs and their respective ligands, a challenge in the field is to understand how each of these various PRRs contributes to an effective and adapted immune response. The study of innate immune recognition is complicated by the existence of multiple PRRs with various expression patterns, variation in PAMP exposure, and modifications through the action of host and bacterial enzymes during the course of infection (2). In addition, PAMP signals intersect with a less well understood but equally complex network of endogenous danger signals, which allow the immune system to discriminate between pathogenic and non-pathogenic microorganisms (3). A better understanding of the mode of action of PRRs ideally requires an in vivo approach in whole organisms using natural routes of infection. The *Drosophila* antimicrobial response is one of the best characterized systems of PRR-mediated defense in metazoans and provides a good model to understand both the logic of pattern recognition and how PRRs shape the ensuing immune response. In this study, we use Phi31-mediated recombineering to generate fly lines expressing specific isoforms of peptidoglycan recognition protein (PGRP)-LC, a *Drosophila* PRR involved in sensing Gram-negative bacteria.

Pattern recognition upstream of the two *Drosophila* innate immune response branches, the Toll and Imd pathways, relies to a large extent on peptidoglycan sensing by PGRPs (4). Peptidoglycan, a cell wall component found in almost all bacteria, is a polymer of alternating *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc), cross-linked by short peptide bridges whose amino acid composition and organization differs among bacteria. As evidenced by Gram staining, peptidoglycan (PGN) forms an abundant external layer in Gram-positive bacteria but is less abundant in Gram-negative bacteria where it is hidden under an external layer of LPS. The structure of PGN from *Bacillus* and Gram-negative bacteria differs from that of most Gram-positive PGN in the third amino acid position of the peptide bridge. Gram-negative and *Bacillus*-type PGNs are cross-linked by a peptide containing a meso-diaminopimelic residue, whereas in other Gram-positive bacterial PGNs a lysine is found in this position (5). In addition, diaminopimelic acid (DAP)-type PGN from Gram-negative bacteria but not DAP-type bacilli contains anhydro-MurNAc residues at the end of each PGN strand, which are distinctive footprints of bacterial PGN synthetic enzymes. Monomers of GlcNAc-1,6-anhydro-MurNAc-L-Ala-γ-D-Glu-meso-DAP-D-Ala, also called tracheal cytotoxin (TCT), represent ~5% of

*Global Health Institute, School of Life Sciences, Swiss Federal Institute of Technology Lausanne, 1015 Lausanne, Switzerland; †Centre de Génétique Moléculaire, Centre National de la Recherche Scientifique, 91198 Gif-sur-Yvette, France; and ‡Immunologie Structurale, Architecture et Fonction des Macromolécules Biologiques, Unité Mixte de Recherche 6098, Centre National de la Recherche Scientifique, 13288 Marseille, France

Received for publication April 10, 2012. Accepted for publication June 5, 2012.

This work was supported by the National Research Fund, Luxembourg (Aides à la Formation-Recherche Postdoctoral Grant 08/037 to C.N.) and by the Bettencourt-Scheller Foundation, a European Research Council Advanced Grant, and the Swiss National Fund (3100A0-12079/1) (to B.L.).

Address correspondence and reprint requests to Prof. Bruno Lemaitre and Dr. Claudine Neyen, Global Health Institute, School of Life Sciences, EPFL, Station 19, 1015 Lausanne, Switzerland. E-mail addresses: bruno.lemaitre@epfl.ch and claudine.neyen@epfl.ch

The online version of this article contains supplemental material.

Abbreviations used in this article: DAP, diaminopimelic acid; *Dpt*, *Diptericin*; GlcNAc, *N*-acetylglucosamine; MurNAc, *N*-acetylmuramic acid; PAMP, pathogen-associated molecular pattern; PGN, peptidoglycan; PGRP, peptidoglycan recognition protein; PRR, pattern recognition receptor; RNAi, RNA interference; RT-qPCR, real-time quantitative PCR; TCT, tracheal cytotoxin.

www.jimmunol.org/cgi/doi/10.4049/jimmunol.1201022

GlcNAc–MurNAc residues and were previously shown to be the minimal PGN motif to elicit Imd responses in flies (6, 7). As anhydro-muropeptides are released during bacterial cell wall synthesis, TCT has been put forward as a specific indicator of potentially dangerous Gram-negative bacterial proliferation (7–9).

Use of highly purified products has demonstrated that in contrast to vertebrates, sensing of Gram-negative bacteria in *Drosophila* is not based on recognition of LPS (6, 10). Rather, the ability of *Drosophila* to discriminate between Gram-positive and Gram-negative bacteria relies on the recognition of specific forms of PGN by PGRPs. The *Drosophila* genome carries a total number of 13 PGRP genes, which give rise to 19 known different receptors (11, 12). The family comprises both enzymatically active, generally secreted, amidase PGRPs that cleave PGN into non-immunogenic fragments and catalytically inactive receptors, generally membrane-bound, which mediate ligand-dependent downstream signaling (13). All family members contain at least one PGRP domain, which is structurally related to bacterial T7 lysozymes and recognizes different types of PGN (11). Whereas PGRP-SA upstream of Toll recognizes mainly lysine-containing PGN from Gram-positive bacteria, Imd-activating PGRP-LC and PGRP-LE exclusively sense DAP-type PGN from Gram-negative bacteria and Gram-positive bacilli (10). In the case of PGRP-LC, differential splicing of PGRP domain-encoding exons to a common intracellular domain-encoding exon generates three receptor isoforms, which differ in their PGN binding specificities but share identical signaling capacities (6, 14). PGRP-LF, a highly similar but signaling-deficient receptor encoded by the locus adjacent to *PGRP-LC*, contains two functional PGRP domains but lacks the intracellular signaling domain and acts as a negative regulator of Imd activation (15, 16). Crystal structures of ligand-binding domains of PGRP-LC isoforms in the presence of monomeric PGN have defined the molecular basis for ligand binding. Only PGRP-LCx contains the characteristic L-shaped PGN binding groove described for mammalian PGRP-Iα (17) and *Drosophila* PGRP-LB (18) and can accommodate polymeric and monomeric PGN. Protruding residues in the ligand binding pocket of PGRP-LCa prevent direct binding of TCT (19, 20), but PGRP-LCa dimerizes with PGRP-LCx–TCT complexes via its PGN binding groove (20, 21). Notably, PGRP domain affinity studies have determined equivalent binding constants for PGRP-LCa and PGRP-LF to PGRP-LCx–TCT complexes (22). Because activation of the Imd pathway relies on ligand-induced receptor homo- or hetero-multimerization, this implies that the stoichiometry of signaling-efficient PGRP-LC isoforms to the signaling-deficient PGRP-LF determines the strength of pathway activation.

Studies in cell culture using RNA interference (RNAi) specific for each PGRP-LC isoform have shown that PGRP-LCx is required for recognition of polymeric PGN, whereas both PGRP-LCa and PGRP-LCx are mandatory for detection of monomeric PGN (6, 14). It has been proposed that signaling is achieved by association of at least two PGRP-LCx molecules in close proximity through binding of polymeric PGN (23). Such an interaction cannot occur with monomeric PGN, and in this case PGRP-LCa is expected to act as an adapter (21). This model is supported by the crystallization of TCT in complex with both PGRP-LCa and PGRP-LCx (20).

Although loss-of-function mutants established the fundamental role of PGRP-LC in survival to Gram-negative infection (24–26), the residual antimicrobial peptide response in flies lacking PGRP-LC compared with Imd-deficient flies suggested a second receptor upstream of Imd. PGRP-LE encodes a PGRP with affinity to DAP-type PGN and is expressed both extracellularly and intracellularly (27, 28). A secreted fragment of PGRP-LE corresponding to the PGRP domain alone functions extracellularly to enhance PGRP-LC–mediated PGN recognition on the cell surface, a role evocative of that of mammalian CD14 in binding of LPS to TLR4 (29). The full-length form of PGRP-LE is cytoplasmic and acts as an intracellular receptor for monomeric PGN, effectively bypassing the requirement for PGRP-LC (29). Thus, both PGRP-LC and PGRP-LE account for sensing of Gram-negative bacteria upstream of the Imd pathway. Finally, detection of DAP-type PGN in *Drosophila* is modulated by amidase PGRPs, which enzymatically degrade PGN and reduce the amount of available immunostimulatory compounds (30, 31). Among these, PGRP-LB has been best characterized as a negative regulator of the Imd pathway (9).

Despite a wealth of studies, several questions remain to be addressed, including the respective contribution of each PGRP-LC isoform and PGRP-LE in response to bacteria as well as the differential requirement of these PRRs in specific tissues, especially in barrier epithelia such as the gut that are constantly exposed to bacterial stimuli. Overexpression studies using full-length and ectodomain-truncated receptors lead to ligand-independent activation of the immune response, probably due to increased receptor proximity in the membrane (23). It is therefore crucial to use an in vivo model with wild-type receptor levels to interpret correctly the mechanism of ligand-specific Imd activation downstream of various PGRP-LC isoforms. We therefore used a genomic complementation approach to supplement *PGRP-LC*–deficient mutants with isoform-specific *PGRP-LC* loci elsewhere in the genome. This approach allowed us to generate wild-type levels of defined PGRP-LC isoforms in vivo and to assess the tissue-specific roles of each isoform, alone or in various combinations. Our results confirm previously described roles of PGRP-LCx and PGRP-LCa/x dimers in polymeric and monomeric PGN sensing, respectively, and uncover a new role for PGRP-LE in the activation of the Imd pathway in the gut. In addition, our in vivo model definitively establishes TCT as the long-distance elicitor of systemic immune responses to intestinal bacteria observed in a case of rupture of tolerance induced by knockdown of amidase *PGRP-LB* (9).

Materials and Methods

Fly stocks and generation of complementation lines

Canton^{S} (*Can*^{S}) and *y,w* flies were used as wild-type controls. The *PGRP-LC*^{E12}, *PGRP-LC*$^{ird7(1)}$, *PGRP-LB*$^{\Delta}$, *PGRP-LE*^{I12}, and *imd*^{1} lines have been described previously (24, 25, 28, 30, 32). *PGRP-LC* rescue lines were generated using gap-repair and recombineering, and final rescue constructs carried by P[acman] vectors were inserted into the PhiC31 landing site 51C on chromosome 2 (BDSC strain 24482) (33, 34). Vectors with *PGRP-LC* only contain the following sequence elements (based on FlyBase release r5.42): *[LC]* 3L:9330982-9342747; *[LCa]* 3L:9330982-9338224;9340740-9342747 (from sequence start until end of exon 7, then from beginning of exon 14 until sequence end); *[LCx]* 3L:9330982-9338167;9338709-9339359;931277-9342747 (from sequence start until end of exon 6, then from beginning of exon 8 until −7 bp from end of 3′UTR of exon 9, then from beginning of 3′UTR of exon 14 until sequence end); *[LCy]* 3L:9330982-9338167;9339691-9340742;9341277-9342747 (from sequence start until end of exon 6, then from beginning of exon 11 until end of exon 13, then from beginning of 3′UTR of exon 14 until sequence end). Vectors with *PGRP-LC* and *PGRP-LF* carry the same sequences as described above but extend until 3L:9345348 instead of 3L:9342747.

Fly genotyping

The following primers were used to genotype flies: *PGRP-LCa* forward 5′-TGGACAACATTGGTGGTGG-3′, reverse 5′-GACCAATGAGTCCAGTTGGC-3′; *PGRP-LCx* forward 5′-GGTGAATGTCGTCCAATCG-3′, reverse 5′-ATTTCGTGTGACCAGTGCG-3′; *PGRP-LCy* forward 5′-TTCCTGTTCTTATTGGACTCGAG-3′, reverse 5′-TTGAGTCATATGCAGATCGGG-3′; primers to verify the *PGRP-LC*^{E12} deletion: forward 5′-CACACGCTGCCATATCAGAC-3′, reverse 5′-TATCGGTTTTCCTGGGTGAG-3′ (will amplify a 212-bp fragment from *PGRP-LC*^{E12} DNA but not from wild type); primers to verify the *PGRP-LC*$^{ird7(1)}$ insertion: forward 5′-TCAAG-

TTAACCGGGACCAAC-3′, reverse 5′-TTTGGCCCATTAAAGCAAAC-3′ (will amplify 1331 bp from wild-type and vector DNA, 2189 bp from *PGRP-LC*$^{ird7(1)}$ DNA, and 2045 bp from vector DNA carrying the eGFP insert).

Bacterial strains and elicitors

All bacteria were grown as shaking cultures in Luria–Bertani medium overnight: *Erwinia carotovora carotovora 15* and *Pseudomonas entomophila* at 29˚C, *Bacillus subtilis* at 37˚C. Bacterial strains were used at the following ODs: Gram-negative DAP-type *Erwinia carotovora carotovora 15* at OD_{600} 200, *P. entomophila* at OD_{600} 200, Gram-positive DAP-type *B. subtilis* at OD_{600} 20. Polymeric PGN purified from *Escherichia coli* and synthetic monomeric PGN (TCT) were gifts from Dominique Mengin-Lecreulx (University of Orsay, Orsay, France).

Infection and survival experiments

Throughout this study infections were performed on 2 to 5-day-old adults, with males for survival and females for mRNA quantification. For systemic infection with bacteria, flies were pricked in the thorax with a metal needle dipped in bacterial pellets and shifted to 29˚C for optimal bacterial proliferation. Flies in survival experiments were kept on medium without fresh yeast and survivors counted daily. Purified elicitors were microinjected using a Nanoject apparatus (Drummond) and pulled glass needles. PGN was injected at a dose of 13.8 nl, 0.05 mg/ml, and TCT at a dose of 13.8 nl, 0.42 mM. For oral infection, flies were starved 2 h at 29˚C then flipped onto fly medium covered with filter disks soaked in a 1:1 mix of bacterial pellets and 5% sucrose or TCT at 0.021 mM in 2.5% sucrose and left to feed for 12–15 h, as previously described (35).

Quantitative RT-PCR

For quantification of *Diptericin* mRNA, whole flies or dissected intestines (cardia to Malpighian tubule branching point) were collected at indicated time points. Total fly RNA was isolated from 5–10 adult flies or 15–20 dissected adult intestines by TRIzol reagent and dissolved in 30 μl (whole flies) or 10 μl (intestines) of RNase-free water. Five hundred nanograms total RNA was then reverse-transcribed in 10 μl reaction volume using Superscript II enzyme (Invitrogen) and random hexamer primers. Quantitative PCR was performed on a LightCycler 480 (Roche) in 96-well plates using the LightCycler 480 SYBR Green I master mix. Primers were as follows: *Diptericin* forward 5′-GCTGCGCAATCGCTTCTACT-3′, reverse 5′-TGGTGGAGTGGGCTTCATG-3′; *RpL32* forward 5′-GACGCTTCAAGGGACAGTATCTG-3′, reverse 5′-AAACGCGGTTCTGCATGAG-3′; *PGRP-LCx* forward 5′-TTTGTCCTTTTCTGCCCAAC-3′, reverse 5′-ATCCAAGTCCGTTTGGTTCA-3′; *PGRP-LCy* forward same as *PGRP-LCx*, reverse 5′-TCAGGATCATTCGTTTGGTTC-3′; *PGRP-LCa* forward 5′- TCTGGACAACATTGGTGGTG-3′, reverse 5′-TCACACTTCTGCGACGATTC-3′.

Immunohistochemistry and confocal microscopy

Fat bodies from *[GFP-LC];PGRP-LC*$^{ird7(1)}$ adult females were dissected and fixed in 4% paraformaldehyde with 0.1% Tween to permeabilize tissues. Fat bodies were stained with rabbit anti-GFP (Interchim) and Alexa 594-coupled mouse anti-rabbit (Molecular Probes) diluted to 1:500 and 1:2000, respectively, in blocking solution (PBS, 0.1% Tween, 2% BSA), and nuclei were stained with DAPI. Images were acquired using a Zeiss LSM700 upright confocal microscope and a ×40/1.30 NA oil immersion objective, with zoom set to ×1.6 and z steps of 0.48 μm. For publication purposes, brightness and contrast were increased identically on control and sample images.

Statistical analysis

All analyses were performed using GraphPad Prism software version 5.02. Survival curves were compared using log-rank tests, and where multiple comparisons were necessary, Bonferroni corrections for p values were applied. All other data were subjected to appropriate ANOVA with Dunnett, Tukey, or Bonferroni post hoc tests.

Results

Generation of flies expressing single PGRP-LC isoforms

The *PGRP-LC* locus including regulatory regions spans ~10 kb between the *PGRP-LA* and *PGRP-LF* loci on chromosome 3L (Fig. 1). Alternative splicing of variable extracellular domain-encoding exons generates three membrane-bound receptor isoforms (PGRP-LCa, PGRP-LCx, and PGRP-LCy) differing in their PGN recognition domain. To build complementation vectors that would mimic wild-type *PGRP-LC* expression when reinserted into the genome, we cloned a genomic fragment including the 3′UTR of *PGRP-LA* and all exons and introns of *PGRP-LC* with or without *PGRP-LF* into P[acman] vectors using gap repair and recombineering techniques (33, 34). These full complementation vectors contain all necessary elements to give rise to all known splice isoforms. Single isoform vectors were then generated by selective deletion of isoform-specific exons as shown in Fig. 1. In addition, a set of full locus vectors (*PGRP-LC* with or without *PGRP-LF*) was also engineered to contain an eGFP coding sequence before the coding sequence of PGRP-LC. All vectors were inserted by PhiC31 integrase-mediated transgenesis into the PhiC31 landing site 51C on chromosome 2 (33). Vector-carrying lines were then combined to two different *PGRP-LC* null mutants,

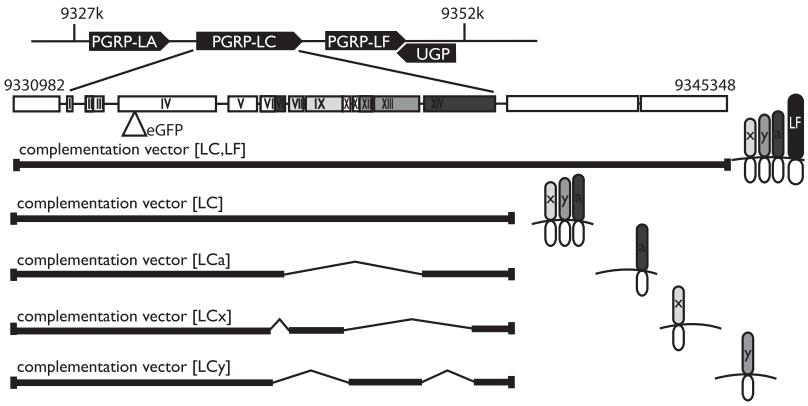

FIGURE 1. Genomic organization of the *PGRP-LC* locus and vector constructs. Genomic organization of the extended *PGRP-LC* locus is based on FlyBase release r5.42. Exons VII to XIV encode extracellular PGRP domains. *PGRP-LC* is preceded by *PGRP-LA* and followed by *PGRP-LF* and *UGP* on the reverse strand. Full-length complementation vectors *[LC,LF]* carry gene region 3L:9330982-9345348 including the 3′UTR of *PGRP-LA*, all exons and introns of *PGRP-LC* and of *PGRP-LF*, as well as the two last exons and 3′UTR of *UGP*. Short complementation vectors *[LC]* carrying gene region 3L:9330982-9342747 are truncated at −39 bp before the start codon of *PGPR-LF*. Single isoform complementation vectors (*[LC]* vectors depicted only) contain the above minus nonspecific exons. Isoform PGRP domains are encoded as follows: PGRP-LCx by exons VIII and IX, PGRP-LCy by exons XII and XIII, and PGRP-LCa by exons VII and XIV. Additional N-terminal fusion constructs of the full locus with *eGFP* were generated by inserting the coding sequence of eGFP (without its start and stop codons) after the start codon in exon 4 in both *[LC,LF]* and *[LC]* vectors.

namely *PGRP-LC*^{E12}, lacking the whole *PGRP-LC* locus (25), and *PGRP-LC*$^{ird7(1)}$, which cannot translate the full PGRP-LC protein due to an insertion in the first coding exon (24). Assuming integration does not significantly alter the expression of the transgene, this should result in full or isoform-specific rescue of the mutation. We tested this hypothesis by infecting these flies through septic injury with the *Drosophila* Gram-negative pathogen *Erwinia carotovora carotovora 15* and monitoring survival (Fig. 2A). *PGRP-LC* null mutants generally succumb to infection within 24 to 48 h, whereas wild-type flies survive the infection (24–26). Both *[LC]* and *[LC,LF]* vectors were able to rescue *PGRP-LC*$^{ird7(1)}$ and *PGRP-LC*^{E12} to wild-type levels, although the rescue in *PGRP-LC*^{E12} was slightly less efficient. The additional genomic copy of *PGRP-LF* from *[LC,LF]* vectors did not impact significantly on survival. However, the intensity of the immune response is influenced by the stoichiometry of genomic copies of *PGRP-LC* and *PGRP-LF*. Indeed, increasing the genomic ratios of *PGRP-LF* to *PGRP-LC* correlated with reduced expression of the antimicrobial peptide *Diptericin* (*Dpt*), one of the downstream targets of the Imd pathway (Supplemental Fig. 1). This is in agreement with a role of *PGRP-LF* as negative regulator of the Imd pathway (15).

We also verified whether the GFP tag on the N-terminal, cytoplasmic tail of the receptor would interfere with function. As shown in Fig. 2B, the presence of GFP does not significantly alter survival levels. Using the GFP tag, we analyzed PGRP-LC tissue expression by immunohistochemistry. Remarkably, PGRP-LC was enriched in the membrane of adult fat body cells facing the hemocele, consistent with a role in sensing ligands carried in the hemolymph (Fig. 2C, *upper panels*). This apico-lateral localization of a receptor also indicates that adult fat body cells are polarized. The localization of PGRP-LC in the fat body did not change noticeably upon infection with *Erwinia carotovora carotovora 15* (Fig. 2C, *lower panels*). Although *PGRP-LC* is expressed in other organs such as the gut or the trachea [Ref. 13 and FlyAtlas (36), Supplemental Fig. 2, and real-time quantitative PCR (RT-qPCR), data not shown), GFP-tagged PGRP-LC in these tissues was below the detection limits of immunochemistry.

Finally, expression dynamics of single isoforms produced from the chromosome 2 *[LC]* transgene were comparable to those from the wild-type *PGRP-LC* locus on chromosome 3 (Fig. 2D), suggesting that all major regulatory elements are present in the rescue vectors.

PGRP-LCx is necessary and sufficient in vivo to mediate Imd pathway activation during systemic infection

To test the in vivo role of each isoform, single isoform complementation lines were established on the *PGRP-LC*^{E12} background (*[LCa/a];PGRP-LC*^{E12} or *[LCx/x];PGRP-LC*^{E12}). For the complementation with *PGRP-LCy*, only the extended locus insertion with *PGRP-LF* was available, resulting in the *[LCy,LF];PGRP-LC*^{E12} complementation line. These lines were then crossed between each other to obtain flies expressing a combination of two distinct isoforms (*[LCa/x];PGRP-LC*^{E12} or *[LCx/y,LF];PGRP-LC*^{E12} or *[LCa/y,LF];PGRP-LC*^{E12}). Correct isoform genotype in these lines was verified by PCR (data not shown). As shown in Fig. 3A, the presence of *PGRP-LCx* alone or in combination with *PGRP-LCa* or *PGRP-LCy* is necessary and sufficient on its own for survival to septic injury with *Erwinia carotovora carotovora 15*, whereas neither *PGRP-LCa* nor *PGRP-LCy*, alone or in combination as *PGRP-LCa/y*, is able to mount a protective response. To assess the ability of each isoform to activate the Imd pathway, *Dpt* mRNA levels were quantified at 8 h after bacterial infection with live or heat-killed *Erwinia carotovora carotovora 15* or live *B. subtilis* (Fig. 3B). Combinations with *[LCy,LF];*

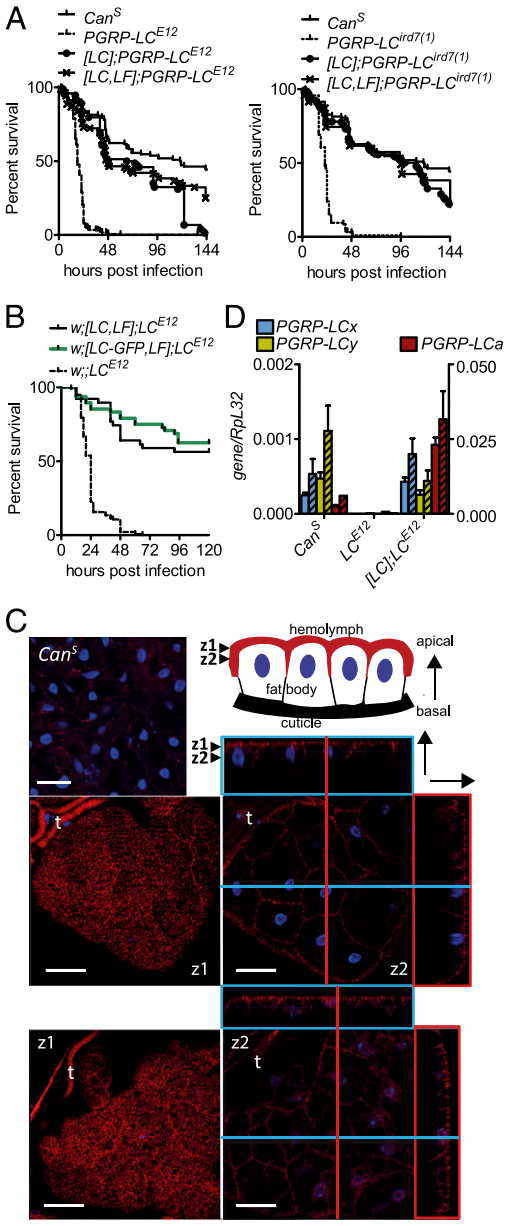

FIGURE 2. In vivo complementation of *PGRP-LC* mutants. (**A**) Both *[LC,LF]* and *[LC]* complementation vectors reestablish a wild-type phenotype in loss-of-function mutants of *PGRP-LC* [alleles *E12*, *ird7(1)*]. Males were infected with *Erwinia carotovora carotovora 15* by septic injury and monitored for survival at 29°C. Results are pooled data from at least three independent experiments each with 20–30 males per genotype. Log-rank test comparing complemented lines to *PGRP-LC*^{E12} and *PGRP-LC*$^{ird7(1)}$, respectively: $p < 0.0001$. (**B**) The presence of an N-terminal GFP tag does not interfere significantly with receptor function. Flies were treated as in (A). Log-rank test comparing lines with and without GFP tag: $p = 0.5001$. Raw survival data for (A) and (B) can be found in Supplemental Table I. (**C**) GFP–PGRP-LC localizes to the apico-lateral membrane of adult fat body cells. Fat bodies from *[GFP-LC];PGRP-LC*$^{ird7(1)}$ adult females (uninfected, *upper panels*; 6 h after pricking with *Erwinia carotovora carotovora 15*, *lower panels*) were dissected, fixed, and stained for GFP. PGRP-LC is shown in red, nuclei are in blue. Shown are top views and orthogonal views from a z-stack taken through a fat body lobe attached to the cuticle, as shown on the diagram. Arrows indicate the position of top views along the *z*-axis of apical (z1) and central (z2) slices. *Inset* shows background staining on *Can*^{S} control. Scale bar in all images, 20 μm. t, trachea. (**D**) Receptor levels in complemented lines mimic wild-type expression dynamics. Adult females were left unchallenged (open bars) or systemically infected with *Erwinia carotovora carotovora 15* for 8 h (hatched bars), and *PGRP-LC* mRNA levels were quantified by RT-qPCR.

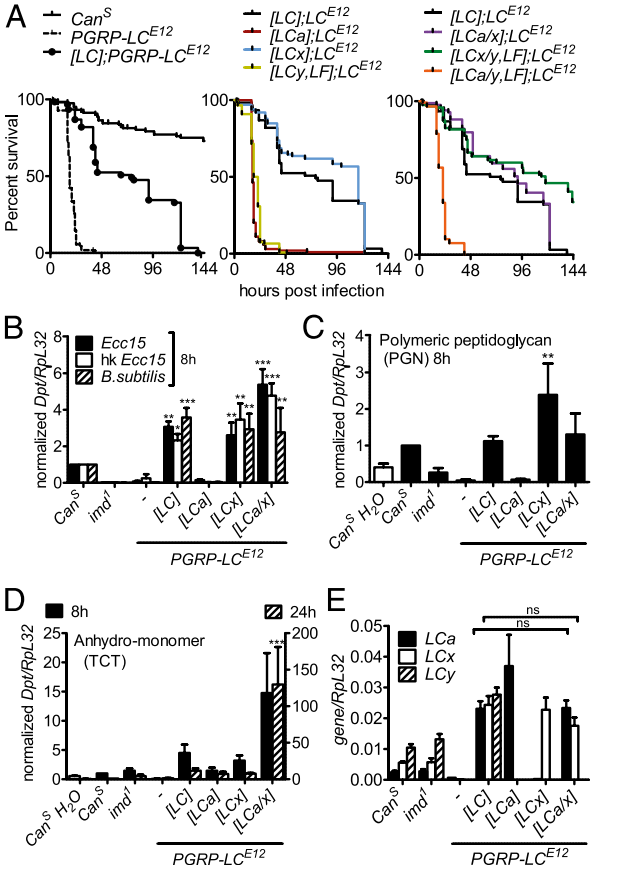

FIGURE 3. PGRP-LCx is necessary and sufficient to respond to Gram-negative bacteria. (**A**) Survival analysis to systemic challenge with Gram-negative bacteria. Males were infected with *Erwinia carotovora carotovora 15* by septic injury and monitored for survival at 29°C. Twenty to thirty males were infected per experiment, and experiments were repeated at least three times. Log-rank test, Bonferroni corrected threshold for multiple comparisons: *PGRP-LC*^{E12} complemented with *[LC]*, *[LCx]*, *[LCa/x]*, or *[LCx/y,LF]* significantly different ($p < 0.0001$) from *PGRP-LC*^{E12}. *PGRP-LC*^{E12} complemented with *[LCa]*, *[LCy,LF]*, *[LCa/y,LF]* not significantly different ($p > 0.05$) from *PGRP-LC*^{E12}. Raw survival data can be found in Supplemental Table I. (**B**) *Dpt* expression in response to systemic challenge with DAP-type Gram-negative and Gram-positive bacteria. Females were infected with live *Erwinia carotovora carotovora 15* (*Ecc15*) or heat-killed *Erwinia carotovora carotovora 15* (hk *Ecc15*) (OD_{600} 200) or live *B. subtilis* (OD_{600} 25) by septic injury and collected at indicated time points for mRNA quantification by RT-qPCR. Data represent the mean + SEM from 10 females per genotype from at least three experiments. For each type of infection, each genotype was compared with the respective *PGRP-LC*^{E12} infected control using one-way ANOVA with Dunnett post hoc test. (**C** and **D**) *Dpt* expression in response to injected polymeric PGN (C) and monomeric PGN [TCT (D)]. Females were microinjected with *Escherichia coli* PGN (13.8 nl, 0.05 mg/ml) or synthetic anhydro-monomeric PGN (TCT, 13.8 nl, 0.42 mM) and collected at indicated time points for mRNA quantification by RT-qPCR (8 h, left *y*-axis; 24 h, right *y*-axis). Results are shown as in (B), and comparisons are as in (B) to injected *PGRP-LC*^{E12} control. (**E**) mRNA levels of all *PGRP-LC* isoforms in unchallenged flies used in (B–D). $*0.01 < p < 0.05$, $**0.001 < p < 0.01$, $***p < 0.001$. ns, Not significant, $p > 0.05$.

PGRP-LC^{E12} were omitted from this experiment because of the aforementioned non-negligible effects of *PGRP-LC* to *PGRP-LF* ratios on *Dpt* mRNA levels. Nevertheless, data discussed later in this article using *[LCy,LF];PGRP-LC*^{E12}flies (see Fig. 5) show that PGRP-LCy alone is unable to confer immune responsiveness. Again, *PGRP-LCx* appears necessary and sufficient to activate the Imd response after sensing DAP-type bacteria (Gram-negative bacteria or Gram-positive bacilli), whereas *PGRP-LCa* remains unable to respond. Notably, the combination of *PGRP-LCx* and *PGRP-LCa* induces a stronger Imd pathway activation than any other combination of isoforms upon challenge with Gram-negative bacteria but not *B. subtilis*. This is consistent with in vitro data indicating a role of the PGRP-LCa/x heterodimer in sensing the PGN monomer TCT (6). Given that TCT is released by Gram-negative bacteria but not DAP-type Gram-positive bacilli such as *B. subtilis*, this also suggests a likely explanation as to why the contribution of *PGRP-LCa* is observed exclusively in response to Gram-negative bacteria.

To confirm this hypothesis, we microinjected flies with either PGN purified from *Escherichia coli* (composed mostly of polymeric PGN) or with a TCT solution. As apparent from Fig. 3C, the presence of *Escherichia coli* PGN in the hemolymph is mainly sensed via PGRP-LCx, whereas TCT induces a very potent and prolonged response only via the PGRP-LCa/x dimer (Fig. 3D). Intriguingly, flies carrying the full *PGRP-LC* locus do not reach a similar magnitude of response, despite identical expression levels of *PGRP-LC* isoforms (Fig. 3E), suggesting that some form of negative regulation is lacking in *PGRP-LCa/x* reconstituted flies. Possibly *PGRP-LCy*, which is absent from *PGRP-LCa/x* reconstituted flies, might dampen the response to TCT in wild-type flies.

To gain further insight into the possible regulatory role of PGRP-LCy, we used 3D modeling to determine the likelihood of PGRP-LCy interacting with PGN. When compared with PGRP-LCx, PGRP-LCy presents several sequence changes affecting the PGN binding cleft (Fig. 4). The resulting structure is unlikely to bind PGN, as it exhibits several features described for PGRP-LCa and PGRP-LFz, two PGRPs that cannot bind to PGN (19, 22). This analysis suggests that PGRP-LCy is structurally incapable of binding PGN and exerts its regulatory function through a mechanism different from that of PGRP-LF.

Collectively, our in vivo studies demonstrate the key role of PGRP-LCx in sensing DAP-type PGN-containing bacteria. We also highlight the contribution of PGRP-LCa/x heterodimers to the systemic immune response to Gram-negative bacteria through sensing of TCT, whereas PGRP-LCy may have a minor role in antagonizing the immune response.

PGPR-LE is dispensable for survival to Gram-negative infections and for systemic sensing of monomeric PGN

The residual level of *Dpt* upon TCT stimulation in a supposedly nonresponsive line (*PGRP-LCa/a*) but not in *PGRP-LC*^{E12} (Fig. 3D) indicates an additional sensing mechanism upstream of PGRP-LC. Previous studies have demonstrated a contribution of PGRP-LE in the sensing of Gram-negative bacteria, acting in cooperation with PGRP-LC (28, 29). In particular, secreted or intracellular forms of PGRP-LE were shown to contribute to TCT sensing upstream of or in parallel to PGRP-LC. If PGRP-LE played a substantial role in the systemic protection against Gram-negative infection, then complementation with *PGRP-LC* isoforms on a *PGRP-LE*^{112}*;;PGRP-LC*^{E12} double mutant background should be less efficient than in single *PGRP-LC*^{E12} mutants. However, when comparing the survival capacity of isoform complementation lines between single and double mutant backgrounds, no significant difference was apparent (Fig. 5A and Supplemental Fig. 3), although there seemed to be a trend toward slightly diminished survival in complemented *PGRP-LE*^{112}*;; PGRP-LC*^{E12} mutants. However, results should be evaluated with caution, as *PGRP-LC*^{E12} complemented lines are on a *w* background, whereas *PGRP-LE*^{112}*;;PGRP-LC*^{E12} complemented lines are *y,w*. Similarly, we observed no difference in *Dpt* induction

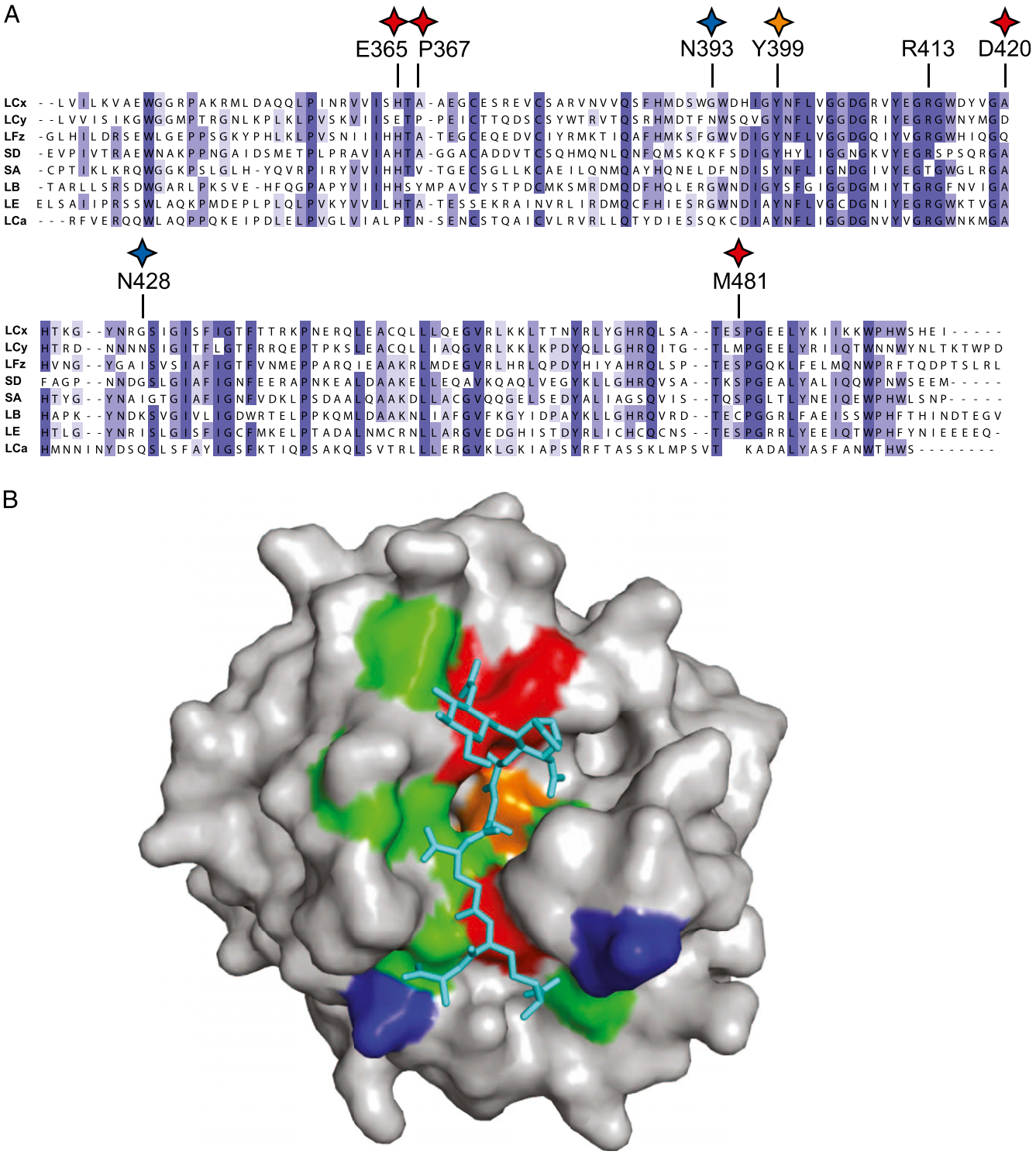

FIGURE 4. Three-dimensional model of the PGRP domain of PGRP-LCy suggests lack of PGN binding ability. (**A**) Alignment of *Drosophila* PGRP domains of known structure. Amino acids critical for binding to TCT in PGRP-LCx are highlighted above the sequence in the color code described later in this caption. (**B**) The 3D model of the PGRP domain of LCy was calculated with the Modeler program (46) using seven homologous structures found by the server HH-pred (47).The position of TCT (stick mode, cyan) at the molecular surface of PGRP-LCy (gray) derives from superimposing the structure of PGRP-LCx in complex with TCT (PDB 2F2L) with the model obtained for LCy. Residues of the PGN binding site that are conserved or interact similarly between LCy and LCx are colored in green. The key residue Tyr399 (LCx numbering is used throughout) is conserved in LCy but oriented differently due to the Ser433 to Thr change in LCy (orange). Residues that differ between LCy and LCx and that create strong clashes with the TCT are colored in red (His365Glu, Ala367Pro, Ala420Asp, and Ser481Met). In the structure of LCx in complex with TCT (20), the imidazole group of His365 contacts the carbonyl group of 2-acetamide and the equatorial 3-OH of GlcNAc. His365 is strictly conserved in all the PGRPs of known structure except in LCa. In LCy, His365 is replaced by Glu, which cannot play a similar role due to its negative charge. The presence of Met instead of Ser481 creates strong clashes. This position is always occupied by a serine in PGRPs devoid of amidase activity (except LCa, which has a phenylalanine at this position) and by a cysteine that coordinates the zinc ion in PGRPs with amidase activity. Asp420 in LCy bonds Arg413 and blocks access to this arginine, which engages in an electrostatic interaction with TCT that is essential for an efficient interaction between LCx and TCT (20). Asp420 plays a role similar to that described for Gln143 in LFz (22). Finally, two Asn residues (N393 and N428; blue) create putative glycosylation sites in the vicinity of the binding crevice that may greatly impair access of PGN.

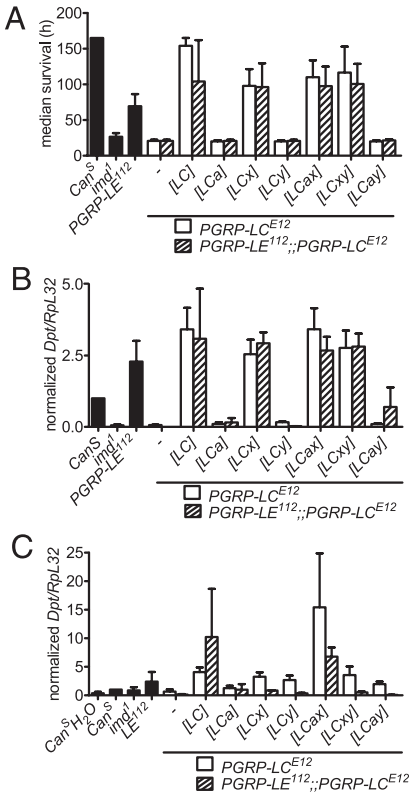

FIGURE 5. PGPR-LE is dispensable for survival to Gram-negative infections and for systemic sensing of monomeric PGN. (**A**) Survival analysis to systemic challenge with Gram-negative bacteria in *w;;PGRP-LC*^{E12} or *y,w,PGRP-LE*^{112}*;;PGRP-LC*^{E12} backgrounds complemented with *[LC,LF]* vectors (LF is omitted from legend for clarity). Males were infected with *Erwinia carotovora carotovora 15* by septic injury and monitored for survival at 29˚C. Results are shown as mean + SEM of median survival from at least three independent experiments each with 20–30 males per genotype. Log-rank test, Bonferroni corrected threshold for multiple comparisons: no significant difference between single or combined isoforms in *PGRP-LC*^{E12} versus *PGRP-LE*^{112}*;;PGRP-LC*^{E12} backgrounds. Raw survival data can be found in Supplemental Table I. (**B**) *Dpt* expression in response to systemic infection with *Erwinia carotovora carotovora 15*. Females were infected with *Erwinia carotovora carotovora 15* by septic injury and collected at 8 h for mRNA quantification by RT-qPCR. (**C**) *Dpt* expression in response to injection of TCT. Females were injected with TCT (13.8 nl, 0.42 mM) and collected at 24 h for mRNA quantification by RT-qPCR. Results in (B) and (C) represent the mean + SEM from 10 females per genotype from three independent experiments. Two-way ANOVA with Bonferroni post hoc tests comparing single or combined isoforms shows no significant difference between *PGRP-LC*^{E12} and *PGRP-LE*^{112}*;;PGRP-LC*^{E12} backgrounds.

after challenge with live *Erwinia carotovora carotovora 15* when comparing single or combined isoforms on *PGRP-LE*^{112}*;;PGRP-LC*^{E12} and *PGRP-LC*^{E12} backgrounds (Fig. 5B), confirming previous results that *PGRP-LE* is dispensable in vivo for Imd activation upon systemic infection with Gram-negative bacteria (28). When flies were injected with TCT, *PGRP-LC*^{E12} mutants invariably retained a very low capacity to induce Imd signaling, more marked when complemented with unresponsive *PGRP-LC* isoforms, indicating a role of PGRP-LE to sense TCT (Fig. 5C). In the double *PGRP-LE*^{112}*;;PGRP-LC*^{E12} mutants, TCT responsiveness was lost entirely, in agreement with a previous study (29). Notably, use of complemented *PGRP-LE*^{112}*;;PGRP-LC*^{E12} mutants demonstrates that detection of TCT now relied exclusively on the combined presence of *PGRP-LCa/x*.

Taken together, these results stress the self-sufficiency of PGRP-LC in combating Gram-negative infection. They also demonstrate that TCT sensing by the fat body can be mediated either by the PGRP-LCa/x heterodimer or by PGRP-LE, but that the contribution of PGRP-LCa/x to TCT sensing prevails over that of PGRP-LE, likely due to the fact that soluble PGRP-LE partially signals via membrane-bound PGRP-LC (29). The observation that *PGRP-LE*^{112}*;[LCx];PGRP-LC*^{E12} flies, which are unable to recognize TCT due to the lack of both *PGRP-LE* and *PGRP-LCa*, still retain some resistance to infection with *Erwinia carotovora carotovora 15* indicates that *Erwinia carotovora carotovora 15* is detected mostly through polymeric PGN.

Both PGRP-LC and PGRP-LE contribute to local Imd activation in the gut

It has been suggested previously that PGRP-LE engages different mechanisms of signal transduction depending on its localization: intracellular, full-length PGRP-LE can signal independently in response to intracellular ligands, whereas extracellular or secreted PGRP-LE is signaling-deficient and acts as a soluble coreceptor of PGRP-LC, reminiscent of CD14 and TLR4 (29). Moreover, different tissues may have different tolerance thresholds to infection: whereas the hemocele would need to remain strictly sterile, organs such as the digestive tract or trachea, which are in direct contact with the external environment, might tolerate higher loads of bacterial elicitors and keep a tighter control on immune activation. This prompted us to investigate the role of PGRP-LC isoforms and PGPP-LE in the intestinal immune response. *PGRP-LC* and *PGRP-LE* are differentially expressed in the fat body and gut compartments (Supplemental Fig. 2). *PGRP-LE* is strongly enriched in the midgut, which is of endodermal origin (depicted as dashed lines in Fig. 6A), whereas *PGRP-LC* is slightly enriched in both the hindgut and the foregut (comprising the crop and the first half of the cardia), which are of ectodermal origin (depicted as solid lines in Fig. 6A) [Fig. 6A, 6B; FlyAtlas (36)]. We have previously shown that PGRP-LC contributes to the gut response through sensing of PGN, whereas the role of PGRP-LE was not assessed (9, 37). In particular, Zaidman-Rémy et al. (9) found that ingestion of Gram-negative PGN or *Erwinia carotovora carotovora 15* induced PGRP-LC–dependent *Dpt* expression in the cardia. This intestinal Imd activation is kept in check by amidase PGRPs, especially PGRP-LB, which degrades PGN (9, 30). We therefore wanted to dissect the contribution of PGRP-LC isoforms and PGRP-LE to the gut immune response. To this end, flies were orally infected with *Erwinia carotovora carotovora 15*, which is nonlethal in this type of infection, and the *Dpt* response was quantified in dissected intestinal tissues comprising the cardia and midgut without Malpighian tubules. As observed for the systemic immune response in adults, the local Imd pathway activation in the gut seems to rely on PGRP-LCx (Fig. 6C), albeit without reaching significance. Along with the residual levels of activation in *PGRP-LC*^{E12} flies, this suggests that another PRR, most likely PGRP-LE, contributes significantly to the gut immune response. We therefore reevaluated the local immune response to *Erwinia carotovora carotovora 15* exclusively in the endodermal section of the gut (midgut without cardia) where the relative enrichment of *PGRP-LE* over *PGRP-LC* suggests a sizeable contribution of PGRP-LE. *PGRP-LE*^{112} mutants are on a *y,w* background, which generally reacts stronger to immune stimuli. Therefore, *PGRP-LC*^{E12} and *PGRP-LE*^{112} single mutants cannot be compared directly. However, when normalized to their respective genetic background controls, both *PGRP-LC*^{E12} and *PGRP-LE*^{112} single mutants showed a significant, around 50% contribution to Imd pathway activation in the midgut (Fig. 6D), and only the combined

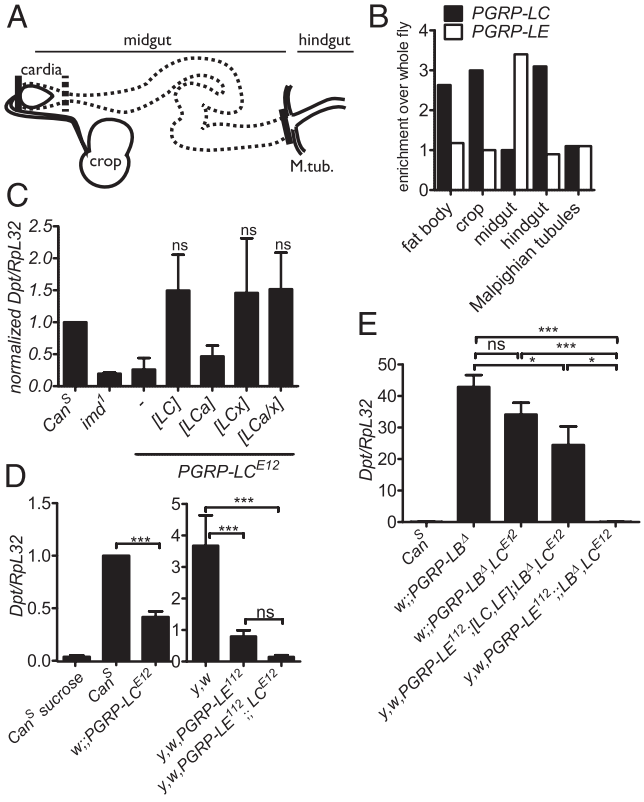

FIGURE 6. Local responses to oral infection rely on combined action of PGRP-LC and PGRP-LE. (**A**) Adult *Drosophila* digestive system, comprising crop, cardia, midgut, Malpighian tubules (M. tub., truncated), and hindgut (truncated). Solid line, ectodermal origin; dashed line, endodermal origin. Full bars indicate dissection points for experiments using midgut with cardia (mixed endodermal and ectodermal tissue); hatched bar indicates dissection point for experiments using exclusively endodermal midgut tissue. (**B**) Expression levels of different PGRP genes based on FlyAtlas. (**C**) Intestinal *Dpt* expression in response to oral challenge with a nonlethal Gram-negative pathogen. Females were fed a mix of sucrose/*Erwinia carotovora carotovora 15* for 12–15 h, after which intestines were collected for mRNA quantification by RT-qPCR. Results represent the mean + SEM from 15–20 females per genotype from three independent experiments. All genotypes were compared with *PGRP-LC*^{E12} using one-way ANOVA with Dunnett post hoc test; no significant difference was observed. (**D**) Both PGRP-LC and PGRP-LE contribute to intestinal *Dpt* expression in response to *Erwinia carotovora carotovora 15*. Females were treated as in (B) but intestines were dissected to include only tissues of endodermal origin. Results represent the mean + SEM from 20–25 females per genotype from three independent experiments expressed as *Dpt/RpL32* ratios. Note different *y*-axis scales: the immune response in *y,w* flies is three times stronger than in Can^S. One-way ANOVA with Tukey post hoc test was performed separately for each background. (**E**) The intestinal response to TCT relies on PGRP-LE and PGRP-LC. Females of indicated genotypes were fed a mix of sucrose/TCT (0.02 mM) for 12–15 h, after which intestines were collected for mRNA quantification by RT-qPCR. Results are shown as described in (C). All genotypes were compared pairwise using one-way ANOVA with Tukey post hoc test. *$0.01 < p < 0.05$, **$0.001 < p < 0.01$, ***$p < 0.001$. ns, Not significant, $p > 0.05$.

loss of both receptors completely abolished the midgut response to infection. This clearly indicates that PGRP-LC and PGRP-LE contribute equally to the activation of the Imd pathway by Gram-negative bacteria in the midgut.

Sensing of TCT in the gut

We next wanted to address whether TCT is directly detected in the gut and to what extent each receptor contributes to this detection. PGRP-LE has been implicated in the sensing of TCT in absorptive tissues like the Malpighian tubules (29), but its role in the intestinal epithelium remains untested. In a preliminary set of experiments, we were unable to detect any strong activation of the Imd pathway upon oral ingestion of TCT. As mentioned above, PGRP-LB rapidly hydrolyzes free intestinal PGN (9, 30), raising the hypothesis that ingested TCT was immediately degraded prior to detection by gut PRRs. To detect a local epithelial response to ingested TCT, we analyzed the contribution of PGRP-LC or PGRP-LE in *PGRP-LB*$^{\Delta}$ mutant flies. Notably, the response to TCT in intestinal tissues (midgut with cardia) was only slightly affected by the loss of *PGRP-LC* alone, more so by the loss of *PGRP-LE* alone, and was completely abolished in mutants lacking both *PGRP-LC* and *PGRP-LE* (Fig. 6E). Note that complemented *PGRP-LE*^{112}*;[LC,LF];PGRP-LB*$^{\Delta}$*,PGRP-LC*^{E12} flies were used as a surrogate for *PGRP-LE*^{112}*;;PGRP-LB*$^{\Delta}$ single mutant flies because of the poor growth of *PGRP-LE*^{112}*;;PGRP-LB*$^{\Delta}$ stocks. From this experiment, we conclude that intestinal cells have the capacity to detect TCT prevalently through PGRP-LE, with a smaller contribution of PGRP-LC.

Long-range activation of immune responses to Gram-negative bacteria relies on PGRP-LCa/x dimers

A subset of Gram-negative bacteria, including *P. entomophila*, is lethal to *Drosophila* even when given by oral route and results in both local and systemic immune responses (38, 39). The fat body response to gut bacteria is thought to be caused by active or passive translocation of bacterial products through the gut barrier, loss of epithelial integrity, or escape of bacteria from the gut. PGRP-LB is proposed to play a major role in preventing this ectopic response by digesting intestinal PGN, a hypothesis supported by loss of tolerance to nonvirulent gut bacteria in *PGRP-LB* knockdown flies (9). We have proposed but not formally demonstrated that small PGN fragments, especially TCT, act as diffusible signaling molecules from the gut to elicit a systemic immune response in *PGRP-LB* knockdown or mutant flies (8, 9). Therefore, we tested which PGRP-LC isoforms were required for long-range activation of Imd both in the response to lethal *P. entomophila* as well as in *PGRP-LB*$^{\Delta}$ mutant flies fed nonlethal *Erwinia carotovora carotovora 15*.

Oral infection with *Erwinia carotovora carotovora 15* in *PGRP-LC* isoform-complemented *PGRP-LB*$^{\Delta}$ mutant flies generated a systemic response that was mainly mediated through PGRP-LCa/x dimers (Fig. 7A). Taking into consideration that TCT is the only ligand sensed by the PGRP-LCa/x heterodimer, our result indicates that this response is mediated by TCT sensing. To verify that TCT is indeed the long-distance elicitor generating this phenotype, we directly fed TCT to these flies and observed the same pattern of PGRP-LC isoform dependence (Fig. 7B). In this case, the high background levels of TCT responsiveness even in the complete absence of *PGRP-LC* were reminiscent of our results with systemic injection of TCT (compare with Fig. 3D). Therefore, we also tested the relative contribution of *PGRP-LE* and *PGRP-LC* in the systemic response of *PGRP-LB*$^{\Delta}$ mutant flies to oral infection with *Erwinia carotovora carotovora 15*. Fig. 7C shows that both PGRP-LE and PGRP-LC significantly contribute to systemic activation of the Imd pathway by TCT feeding of *PGRP-LB*$^{\Delta}$–deficient flies.

Likewise, when wild-type flies were orally infected with lethal *P. entomophila*, the ensuing systemic immune response relied mainly on the PGRP-LCa/x combination (Fig. 7D), indicating that this response is mediated by TCT. The observation that wild-type flies or *PGRP-LC*^{E12} mutants carrying the full *PGRP-LC* locus do not reach the same amplitude of response compared with *PGRP-LCa/x* reconstituted flies (Fig. 7B, 7D) is puzzling. A possible expla-

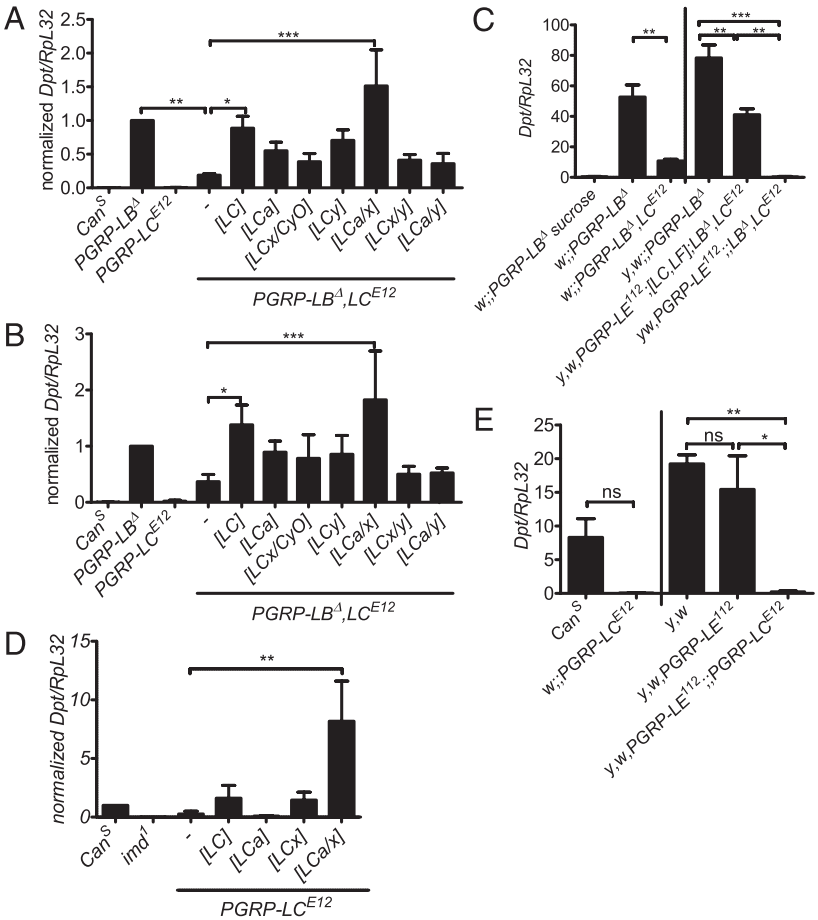

FIGURE 7. Long-range activation of immune response to Gram-negative bacteria relies on TCT recognition by PGRP-LCa/x dimers and PGRP-LE. (**A** and **B**) Loss of systemic tolerance to oral infection with a nonlethal Gram-negative pathogen in *PGRP-LB*$^{\Delta}$,*PGRP-LC*^{E12} backgrounds complemented with [LC,LF] vectors (LF omitted in legend for clarity). Females were fed *Erwinia carotovora carotovora 15* (A) or synthetic TCT (B) for 12 h, then whole flies were collected for mRNA quantification by RT-qPCR. Results are shown as mean + SEM from 10 females per genotype from at least three independent experiments. All genotypes were compared with *PGRP-LC*^{E12} using one-way ANOVA with Dunnett post hoc test. All genotypes were compared with *PGRP-LC*^{E12} using two-way ANOVA with Bonferroni post hoc test. (**C**) PGRP-LE contributes to the systemic response to oral infection in *PGRP-LB*$^{\Delta}$-deficient flies. Females were fed a mix of sucrose/ *Erwinia carotovora carotovora 15* for 12–15 h, after which whole flies were collected for mRNA quantification by RT-qPCR. Data represents the mean + SEM from 10 females per genotype from at least three independent experiments. One-way ANOVA with Tukey post hoc test was performed separately for each background. (**D**) Systemic *Dpt* expression in response to oral challenge with a lethal Gram-negative pathogen. Females were fed sucrose/*P. entomophila* for 12 h, after which whole flies were collected for mRNA quantification by RT-qPCR. Results are shown as in (A). (**E**) Experiments in (D) were repeated with mutants for either PGRP-LE or PGRP-LC and their respective background controls. Results represent the mean + SEM from 20–25 females per genotype from at least three independent experiments, expressed as *Dpt/RpL32* ratios. Note the stronger immune response in *y,w* flies. Results were analyzed using one-way ANOVA with Tukey post hoc test. *$0.01 < p < 0.05$, **$0.001 < p < 0.01$, ***$p < 0.001$. ns, Not significant, $p > 0.05$.

nation brought forward before is that PGRP-LCy, which is absent from *PGRP-LCa/x* reconstituted flies, might dampen the response to TCT in wild-type flies. As apparent from Fig. 7E, loss of *PGRP-LE* had no significant effect on the systemic response to *P. entomophila.* However, the levels of systemic *Dpt* induction in the *P. entomophila* model are five times lower than those observed in the TCT-fed *PGRP-LB*$^{\Delta}$ mutant model (compare non-normalized *Dpt/RpL32* ratios in Fig. 7C and 7E), a situation that might mask the possible contribution of PGRP-LE.

Taken together, these experiments prove that long-range activation of the systemic response upon oral infection is mediated by TCT and that the PGRP-LCa/x heterodimer plays a major role in sensing TCT. However, we noted a discernible role of PGRP-LE in eliciting systemic responses to orally administered TCT in *PGRP-LB*$^{\Delta}$ mutants.

Discussion

Our initial aim was to define the role of each PGRP-LC isoform in vivo. Using Phi31-mediated recombineering, we successfully inserted loci of full and isoform-specific *PGRP-LC* constructs into the fly genome and proved they were capable of complementing *PGRP-LC* null mutations. Our in vivo approach confirms and extends previous in vitro and RNAi experiments in proving that PGRP-LCx is indeed necessary and sufficient to respond to challenge with live or dead Gram-negative bacteria and to Gram-positive, DAP-type bacilli. Moreover, PGRP-LCx alone induces the in vivo immune response to polymeric PGN, whereas combined presence of PGRP-LCx and PGRP-LCa is necessary to sense the anhydro-monomer TCT. The differential requirement of PGRP-LC isoforms in response to Gram-positive DAP-type (PGRP-LCx alone) and Gram-negative bacteria (both PGRP-LCx and PGRP-LCa/x) indicates that flies are able to discriminate between the two types of DAP-type PGN-containing bacteria and to mount appropriate responses. Notably, injection of TCT in contrast to polymeric PGN leads to an increase in amplitude and duration of Imd pathway activation (Ref. 9 and this study). Thus, TCT detection by PGRP-LCa/x allows flies to mount a strong response to Gram-negative bacteria despite the fact that DAP-type

PGN is not exposed (masked by the LPS layer) and is less abundant compared with DAP-type PGN-containing Gram-positive bacteria (40).

Consistent with previous reports that showed no effect of *PGRP-LCy* RNAi on PGN sensing in cells (6, 14), PGRP-LCy on its own did not show any induction of the Imd pathway. However, bacterial infection or injection of immunostimulatory compounds repeatedly produced a stronger response in flies carrying *PGRP-LCa/x* isoforms than in flies carrying the whole *PGRP-LC* locus. Although we cannot exclude subtle differences in isoform expression from the intact, full locus compared with the engineered isoform loci, this suggests that the full locus carries an additional regulatory element lacking in heterozygous *PGRP-LCa/x* flies. It is tempting to speculate that PGRP-LCy, present in the full locus but absent from *PGRP-LCa/x* flies, might help to regulate response levels. PGRP-LCy is structurally unlikely to bind PGN but, unlike PGRP-LF, retains a signaling-competent cytoplasmic tail. If any regulatory activity was associated with the PGRP-LCy isoform, it would therefore have to act extracellularly, possibly by competing with other isoforms for cell surface localization and thereby diluting receptor availability. Thus, the only function we can attribute to PGRP-LCy from this study is a regulatory role in adjusting the amplitude of Imd pathway activation.

The importance of wild-type receptor levels in any study of isoform function is crucial because overexpression of receptors is sufficient on its own to stimulate the Imd pathway (24, 25). Our *PGRP-LC* complemented system mimics wild-type receptor expression dynamics, and we did not detect any elevated background levels of Imd activation in complemented *PGRP-LC* mutant flies (data not shown). However, alterations in the genomic ratio of *PGRP-LC* to *PGRP-LF*, achieved by combining *[LC]* or *[LC,LF]* vector-carrying lines with wild-type or different *PGRP-LC*–deficient backgrounds, showed a significant correlation between *Dpt* levels and *PGRP-LC/LF* ratios in infected flies, consistent with an inhibitory role of PGRP-LF (15, 16). This indicates that the stoichiometry of activating and regulating receptors matters, as foreshadowed by affinity studies between signaling-competent PGRP-LCx–TCT–PGRP-LCa and signaling-deficient PGRP-LCx–TCT–PGRP-LF complexes (22).

Several overexpression studies in S2 cells already localized PGRP-LC to the plasma membrane (6, 41). We extend this finding to wild-type receptor expression levels in an immunocompetent tissue and provide evidence that PGRP-LC localizes to the apical and lateral plasma membrane in fat body cells, revealing a previously undescribed polarity in this immune-responsive tissue.

Similar to Takehana et al. (28), who described no significant difference between *Dpt* expression in *PGRP-LC* versus *PGRP-LE;; PGRP-LC* mutants after stimulation with *B. subtilis* and *Escherichia coli*, we see no additional decrease in survival rates to *Erwinia carotovora carotovora 15* when comparing single *PGRP-LC* and double *PGRP-LE;;LC* mutants, and no significant underlying reduction in *Dpt* levels. This underlines the major role of PGRP-LC to survey a defined compartment—the insect hemolymph—and to preferentially activate immune responses in the fat body.

In general agreement with Kaneko et al. (29), we confirmed a role of PGRP-LE in the systemic immune response to TCT, albeit depending on the route of administration. On one hand, we note a predominant role of PGRP-LCa/x over PGRP-LE in sensing injected TCT in the hemolymph. In this context, the contribution of PGRP-LE was discernible in the presence of any PGRP-LC isoform but was less marked in the absence of the full locus, consistent with the concept that hemolymph PGRP-LE cannot signal directly but depends on membrane-bound PGRP-LC to relay information (29). However, even though secreted PGRP-LE might contribute to Imd activation by delivering hemolymph TCT/PGN to membrane-bound PGRP-LC, the effect of complete *PGRP-LE* loss on systemic immune activation after injection of TCT into the hemocele was not significant. This suggests that the cytosolic, autonomous PGRP-LE form does not contribute significantly to the activation of the Imd pathway by injected TCT and establishes PGRP-LC as the predominant receptor eliciting systemic responses in the hemocele.

On the other hand, when Imd activation in the fat body was triggered by oral ingestion of TCT in the *PGRP-LB* mutant background, we observed a non-negligible contribution of PGRP-LE to this systemic response in the absence of PGRP-LC. This indicates that when TCT reached the hemocele by active or passive transport from the intestine, the role of cytosolic PGRP-LE became more prominent. Although we have no explanation for this discrepancy, one might speculate that even though cytosolic PGRP-LE does not significantly contribute to TCT sensing when injected into the hemolymph, possibly because the fat body lacks transporters present in absorptive organs, this intracellular mode of recognition gains in importance when TCT transits through cells.

Taken together, the subordinate role of secreted PGRP-LE compared with PGRP-LC might suggest that the main contribution of PGRP-LE is as an intracellular sensor, which will only spring into action when systemic levels of TCT have reached a critical threshold and permeated the cytosol.

Determining the mechanisms by which barrier epithelia sense bacteria and differentiate between acceptable and nonacceptable intruders is a major issue in the field of innate immunity. Previous studies proposed PRR compartmentalization as an essential mechanism to discriminate between pathogenic versus beneficial bacterial colonization (42). Although we observed a clear role of PGRP-LC sensing in the gut, consistent with previous studies, we cannot conclude whether this reflects direct sampling of the gut lumen by PGRP-LC. Unfortunately, the expression of the *PGRP-LC-GFP* fusion construct was not strong enough to determine whether PGRP-LC is expressed at the apical or the basal side of enterocytes. Of note, recognition PGRPs involved in the sensing of Gram-negative bacteria show differential expression patterns along the gut, with enrichment of *PGRP-LE* in the endodermally derived midgut and a modest enrichment of *PGRP-LC* in ectodermally derived foregut and hindgut. Moreover, PGRPs in these sections are more or less accessible to gut contents. A relatively impermeable cuticle protects ectodermal epithelia in the foregut and hindgut, whereas the peritrophic matrix covering the *PGRP-LE*-rich section of the midgut is permeable to allow passage of digested nutrients (43). It is therefore more likely for bacterial compounds to reach midgut epithelia, and a reduction in surface receptors capable of mounting potentially detrimental immune responses to commensals in this compartment would make sense. Cytosolic receptors expressed in this compartment would be able specifically to detect absorbed or diffusible bacterial compounds such as TCT, which may be a hallmark of proliferation and/or harmful bacteria. Consistent with this, we saw a major contribution of PGRP-LE (most probably of the cytosolic form as PGRP-LE signaling did not depend on PGRP-LC) and less of PGRP-LC when we assessed the midgut-specific response to Gram-negative bacteria. More strikingly, the midgut response to ingested TCT relied mostly on PGRP-LE, supporting a role of this receptor in danger detection in the gut. Thus, our study uncovers a key role of PGRP-LE in the *Drosophila* midgut and suggests that intracellular sensing of TCT is used in *Drosophila* as a mechanism to recognize infectious bacteria.

We previously put forward a model whereby long-range activation of the systemic immune response in *Drosophila* is mediated

by the translocation of small PGN fragments from the gut lumen or other barrier epithelia to the hemolymph. This view was supported by the observation that ingestion of monomeric PGN can stimulate a strong systemic immune response in *PGRP-LB* knockdown flies with reduced amidase activity and that deposition of PGN or TCT on the genitalia is sufficient to induce a systemic immune response (8, 9). Moreover, because TCT consistently elicited stronger responses than PGN, these models proposed an involvement of active or passive transport of the elicitor to the hemocele. On the basis of our results, we cannot yet conclude the mechanism of TCT delivery to the hemocele. However, the unique and well-characterized interaction of TCT–PGRP-LCa–PGRP-LCx (20) and the primordial role of PGRP-LCa/x heterodimers in mediating TCT-specific systemic activation of the Imd pathway (this study) demonstrates that TCT is indeed a crucial element in the long-range activation of the immune response.

In conclusion, our study shows that a combination of extracellular sensing by PGRP-LC isoforms and intracellular sensing through PGRP-LE provides sophisticated mechanisms to detect and differentiate between infections by different DAP-type bacteria in *Drosophila*. It is probable that the absence of LPS sensing in *Drosophila* has imposed some constraints on the system and that sensing of TCT through PGRP-LCa/x and PGRP-LE evolved as a surrogate way to distinguish Gram-negative bacteria from Gram-positive DAP-type PGN-containing bacteria. Because TCT is released during bacterial division, intracellular sensing through PGRP-LE provides an adequate mechanism of detection in the gut, reminiscent of the intracellular sensing of Gram-negative muropeptides by intracellular NOD1 in epithelia (44, 45). To date, the existence of a mode of recognition of lysine-type bacteria in the midgut remains unexplored. A simple explanation could be that lysine-type bacteria do not represent a threat for flies as they rarely infect via the oral route and are therefore not detected. Indeed, DAP-type PGN-containing bacteria of either Gram-negative type (*Serratia*, *Pseudomonas*) or bacillus-type (*Bacillus thuringiensis*) are the only characterized naturally occurring insect pathogens to date.

Acknowledgments

We thank Quentin Geissmann for early work on the project and Fanny Schüpfer for technical support, Juan Paredes and Halit Ongen for critical reading of the manuscript, D. Mengin-Lecreulx for PGN samples, and the Bloomington Stock Centre and S. Kurata for fly stocks.

Disclosures

The authors have no financial conflicts of interest.

References

1. Janeway, C. A., Jr., and R. Medzhitov. 2002. Innate immune recognition. *Annu. Rev. Immunol.* 20: 197–216.
2. Vance, R. E., R. R. Isberg, and D. A. Portnoy. 2009. Patterns of pathogenesis: discrimination of pathogenic and nonpathogenic microbes by the innate immune system. *Cell Host Microbe* 6: 10–21.
3. Blander, J. M., and L. E. Sander. 2012. Beyond pattern recognition: five immune checkpoints for scaling the microbial threat. *Nat. Rev. Immunol.* 12: 215–225.
4. Lemaitre, B., and J. Hoffmann. 2007. The host defense of *Drosophila melanogaster. Annu. Rev. Immunol.* 25: 697–743.
5. Schleifer, K. H., and O. Kandler. 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol. Rev.* 36: 407–477.
6. Kaneko, T., W. E. Goldman, P. Mellroth, H. Steiner, K. Fukase, S. Kusumoto, W. Harley, A. Fox, D. Golenbock, and N. Silverman. 2004. Monomeric and polymeric gram-negative peptidoglycan but not purified LPS stimulate the Drosophila IMD pathway. *Immunity* 20: 637–649.
7. Stenbak, C. R., J. H. Ryu, F. Leulier, S. Pili-Floury, C. Parquet, M. Hervé, C. Chaput, I. G. Boneca, W. J. Lee, B. Lemaitre, and D. Mengin-Lecreulx. 2004. Peptidoglycan molecular requirements allowing detection by the Drosophila immune deficiency pathway. *J. Immunol.* 173: 7339–7348.
8. Gendrin, M., D. P. Welchman, M. Poidevin, M. Hervé, and B. Lemaitre. 2009. Long-range activation of systemic immunity through peptidoglycan diffusion in Drosophila. *PLoS Pathog.* 5: e1000694.
9. Zaidman-Rémy, A., M. Hervé, M. Poidevin, S. Pili-Floury, M. S. Kim, D. Blanot, B. H. Oh, R. Ueda, D. Mengin-Lecreulx, and B. Lemaitre. 2006. The Drosophila amidase PGRP-LB modulates the immune response to bacterial infection. *Immunity* 24: 463–473.
10. Leulier, F., C. Parquet, S. Pili-Floury, J. H. Ryu, M. Caroff, W. J. Lee, D. Mengin-Lecreulx, and B. Lemaitre. 2003. The Drosophila immune system detects bacteria through specific peptidoglycan recognition. *Nat. Immunol.* 4: 478–484.
11. Dziarski, R., and D. Gupta. 2006. The peptidoglycan recognition proteins (PGRPs). *Genome Biol.* 7: 232.
12. Royet, J., D. Gupta, and R. Dziarski. 2011. Peptidoglycan recognition proteins: modulators of the microbiome and inflammation. *Nat. Rev. Immunol.* 11: 837–851.
13. Werner, T., G. Liu, D. Kang, S. Ekengren, H. Steiner, and D. Hultmark. 2000. A family of peptidoglycan recognition proteins in the fruit fly *Drosophila melanogaster. Proc. Natl. Acad. Sci. USA* 97: 13772–13777.
14. Werner, T., K. Borge-Renberg, P. Mellroth, H. Steiner, and D. Hultmark. 2003. Functional diversity of the Drosophila PGRP-LC gene cluster in the response to lipopolysaccharide and peptidoglycan. *J. Biol. Chem.* 278: 26319–26322.
15. Maillet, F., V. Bischoff, C. Vignal, J. Hoffmann, and J. Royet. 2008. The Drosophila peptidoglycan recognition protein PGRP-LF blocks PGRP-LC and IMD/JNK pathway activation. *Cell Host Microbe* 3: 293–303.
16. Persson, C., S. Oldenvi, and H. Steiner. 2007. Peptidoglycan recognition protein LF: a negative regulator of Drosophila immunity. *Insect Biochem. Mol. Biol.* 37: 1309–1316.
17. Guan, R., A. Roychowdhury, B. Ember, S. Kumar, G. J. Boons, and R. A. Mariuzza. 2004. Structural basis for peptidoglycan binding by peptidoglycan recognition proteins. *Proc. Natl. Acad. Sci. USA* 101: 17168–17173.
18. Kim, M. S., M. Byun, and B. H. Oh. 2003. Crystal structure of peptidoglycan recognition protein LB from *Drosophila melanogaster. Nat. Immunol.* 4: 787–793.
19. Chang, C. I., K. Ihara, Y. Chelliah, D. Mengin-Lecreulx, S. Wakatsuki, and J. Deisenhofer. 2005. Structure of the ectodomain of Drosophila peptidoglycan-recognition protein LCa suggests a molecular mechanism for pattern recognition. *Proc. Natl. Acad. Sci. USA* 102: 10279–10284.
20. Chang, C. I., Y. Chelliah, D. Borek, D. Mengin-Lecreulx, and J. Deisenhofer. 2006. Structure of tracheal cytotoxin in complex with a heterodimeric pattern-recognition receptor. *Science* 311: 1761–1764.
21. Mellroth, P., J. Karlsson, J. Håkansson, N. Schultz, W. E. Goldman, and H. Steiner. 2005. Ligand-induced dimerization of Drosophila peptidoglycan recognition proteins in vitro. *Proc. Natl. Acad. Sci. USA* 102: 6455–6460.
22. Basbous, N., F. Coste, P. Leone, R. Vincentelli, J. Royet, C. Kellenberger, and A. Roussel. 2011. The Drosophila peptidoglycan-recognition protein LF interacts with peptidoglycan-recognition protein LC to downregulate the Imd pathway. *EMBO Rep.* 12: 327–333.
23. Choe, K. M., H. Lee, and K. V. Anderson. 2005. Drosophila peptidoglycan recognition protein LC (PGRP-LC) acts as a signal-transducing innate immune receptor. *Proc. Natl. Acad. Sci. USA* 102: 1122–1126.
24. Choe, K. M., T. Werner, S. Stöven, D. Hultmark, and K. V. Anderson. 2002. Requirement for a peptidoglycan recognition protein (PGRP) in Relish activation and antibacterial immune responses in Drosophila. *Science* 296: 359–362.
25. Gottar, M., V. Gobert, T. Michel, M. Belvin, G. Duyk, J. A. Hoffmann, D. Ferrandon, and J. Royet. 2002. The Drosophila immune response against Gram-negative bacteria is mediated by a peptidoglycan recognition protein. *Nature* 416: 640–644.
26. Rämet, M., P. Manfruelli, A. Pearson, B. Mathey-Prevot, and R. A. Ezekowitz. 2002. Functional genomic analysis of phagocytosis and identification of a Drosophila receptor for *E. coli. Nature* 416: 644–648.
27. Takehana, A., T. Katsuyama, T. Yano, Y. Oshima, H. Takada, T. Aigaki, and S. Kurata. 2002. Overexpression of a pattern-recognition receptor, peptidoglycan-recognition protein-LE, activates imd/relish-mediated antibacterial defense and the prophenoloxidase cascade in Drosophila larvae. *Proc. Natl. Acad. Sci. USA* 99: 13705–13710.
28. Takehana, A., T. Yano, S. Mita, A. Kotani, Y. Oshima, and S. Kurata. 2004. Peptidoglycan recognition protein (PGRP)-LE and PGRP-LC act synergistically in Drosophila immunity. *EMBO J.* 23: 4690–4700.
29. Kaneko, T., T. Yano, K. Aggarwal, J. H. Lim, K. Ueda, Y. Oshima, C. Peach, D. Erturk-Hasdemir, W. E. Goldman, B. H. Oh, et al. 2006. PGRP-LC and PGRP-LE have essential yet distinct functions in the drosophila immune response to monomeric DAP-type peptidoglycan. *Nat. Immunol.* 7: 715–723.
30. Paredes, J. C., D. P. Welchman, M. Poidevin, and B. Lemaitre. 2011. Negative regulation by amidase PGRPs shapes the Drosophila antibacterial response and protects the fly from innocuous infection. *Immunity* 35: 770–779.
31. Mellroth, P., J. Karlsson, and H. Steiner. 2003. A scavenger function for a Drosophila peptidoglycan recognition protein. *J. Biol. Chem.* 278: 7059–7064.
32. Georgel, P., S. Naitza, C. Kappler, D. Ferrandon, D. Zachary, C. Swimmer, C. Kopczynski, G. Duyk, J. M. Reichhart, and J. A. Hoffmann. 2001. Drosophila immune deficiency (IMD) is a death domain protein that activates antibacterial defense and can promote apoptosis. *Dev. Cell* 1: 503–514.
33. Venken, K. J., Y. He, R. A. Hoskins, and H. J. Bellen. 2006. P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. *Science* 314: 1747–1751.
34. Warming, S., N. Costantino, D. L. Court, N. A. Jenkins, and N. G. Copeland. 2005. Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res.* 33: e36.

35. Romeo, Y., and B. Lemaitre. 2008. Drosophila immunity: methods for monitoring the activity of Toll and Imd signaling pathways. *Methods Mol. Biol.* 415: 379–394.
36. Chintapalli, V. R., J. Wang, and J. A. Dow. 2007. Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nat. Genet.* 39: 715–720.
37. Buchon, N., N. A. Broderick, M. Poidevin, S. Pradervand, and B. Lemaitre. 2009. Drosophila intestinal response to bacterial infection: activation of host defense and stem cell proliferation. *Cell Host Microbe* 5: 200–211.
38. Vodovar, N., M. Vinals, P. Liehl, A. Basset, J. Degrouard, P. Spellman, F. Boccard, and B. Lemaitre. 2005. Drosophila host defense after oral infection by an entomopathogenic *Pseudomonas* species. *Proc. Natl. Acad. Sci. USA* 102: 11414–11419.
39. Liehl, P., M. Blight, N. Vodovar, F. Boccard, and B. Lemaitre. 2006. Prevalence of local immune response against oral infection in a Drosophila/*Pseudomonas* infection model. *PLoS Pathog.* 2: e56.
40. Mengin-Lecreulx, D., and B. Lemaitre. 2005. Structure and metabolism of peptidoglycan and molecular requirements allowing its detection by the Drosophila innate immune system. *J. Endotoxin Res.* 11: 105–111.
41. Lhocine, N., P. S. Ribeiro, N. Buchon, A. Wepf, R. Wilson, T. Tenev, B. Lemaitre, M. Gstaiger, P. Meier, and F. Leulier. 2008. PIMS modulates immune tolerance by negatively regulating Drosophila innate immune signaling. *Cell Host Microbe* 4: 147–158.
42. Sansonetti, P. J. 2004. War and peace at mucosal surfaces. *Nat. Rev. Immunol.* 4: 953–964.
43. Hegedus, D., M. Erlandson, C. Gillott, and U. Toprak. 2009. New insights into peritrophic matrix synthesis, architecture, and function. *Annu. Rev. Entomol.* 54: 285–302.
44. Chamaillard, M., M. Hashimoto, Y. Horie, J. Masumoto, S. Qiu, L. Saab, Y. Ogura, A. Kawasaki, K. Fukase, S. Kusumoto, et al. 2003. An essential role for NOD1 in host recognition of bacterial peptidoglycan containing diaminopimelic acid. *Nat. Immunol.* 4: 702–707.
45. Girardin, S. E., I. G. Boneca, L. A. Carneiro, A. Antignac, M. Jéhanno, J. Viala, K. Tedin, M. K. Taha, A. Labigne, U. Zähringer, et al. 2003. Nod1 detects a unique muropeptide from gram-negative bacterial peptidoglycan. *Science* 300: 1584–1587.
46. Sali, A., and T. L. Blundell. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234: 779–815.
47. Remmert, M., A. Biegert, A. Hauser, and J. Söding. 2012. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods* 9: 173–175.