



# Type VII Secretion Substrates of Pathogenic Mycobacteria Are Processed by a Surface Protease

Maroeska J. Burggraaf,<sup>a</sup> Alexander Speer,<sup>b</sup> Aniek S. Meijers,<sup>b</sup> Roy Ummels,<sup>b</sup> Astrid M. van der Sar,<sup>b\*</sup>  
Konstantin V. Korotkov,<sup>c</sup> Wilbert Bitter,<sup>b,d</sup> Coenraad Kuijl<sup>b</sup>

<sup>a</sup>Amsterdam UMC, Vrije Universiteit Amsterdam, Medical Microbiology and Infection Control, Cancer Center Amsterdam, Amsterdam, Netherlands

<sup>b</sup>Amsterdam UMC, Vrije Universiteit Amsterdam, Medical Microbiology and Infection Control, Amsterdam Institute of Infection & Immunity, Amsterdam, Netherlands

<sup>c</sup>Molecular and Cellular Biochemistry, College of Medicine, University of Kentucky, Lexington, Kentucky, USA

<sup>d</sup>Molecular Microbiology, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

**ABSTRACT** Tuberculosis, one of the world's most severe infectious diseases, is caused by *Mycobacterium tuberculosis*. A major weapon of this pathogen is a unique cell wall that protects the pathogen from eradication by the immune system. Mycobacteria have specialized secretion systems, e.g., type VII secretion or ESX systems, to transport substrates across this cell wall. The largest group of proteins that are secreted by these ESX systems are the PE proteins. Previously, it was shown that the N-terminal PE domain of about 100 amino acids is required for secretion. Here, we describe the identification of an aspartic protease, designated PecA, that removes (part of) this PE domain at the cell surface. Nearly all of the observed PE\_PGRS proteins are processed by PecA. Interestingly, the protease itself is also a secreted PE protein and subject to self-cleavage. Furthermore, a defect in surface processing has no effect on the activity of the PE lipase protein LipY but does seem to affect the functioning of other virulence factors, as a *pecA* mutant strain of *Mycobacterium marinum* shows moderate attenuation in zebrafish larvae. In conclusion, our results reveal the presence of a functional aspartic acid protease in *M. marinum* that cleaves LipY, itself as well as other members of the PE\_PGRS family. Finally, mutants lacking PecA show growth attenuation *in vivo*, suggesting that PecA plays a role during infection.

**IMPORTANCE** Aspartic proteases are common in eukaryotes and retroviruses but are relatively rare among bacteria (N. D. Rawlings and A. Bateman, BMC Genomics 10:437, 2009, <https://doi.org/10.1186/1471-2164-10-437>). In contrast to eukaryotic aspartic proteases, bacterial aspartic proteases are generally located in the cytoplasm. We have identified a surface-associated mycobacterial aspartic protease, PecA, which cleaves itself and many other type VII secretion substrates of the PE\_PGRS family. PecA is present in most pathogenic mycobacterial species, including *M. tuberculosis*. In addition, pathogenicity of *M. marinum* is reduced in the  $\Delta pecA$  mutant, indicating that PecA contributes to virulence.

**KEYWORDS** aspartic protease, *Mycobacterium*, PE proteins, PE\_PGRS, type VII secretion

Tuberculosis is one of the world's deadliest infectious diseases with 1.3 million deaths each year (1). Its causative agent is *Mycobacterium tuberculosis*, a facultative intracellular pathogen with a highly unique di- and tri-acylated cell envelope. This unique cell envelope is highly impermeable and requires specialized secretion systems for protein transport to the cell surface (2, 3). These protein secretion systems are generally known as type VII secretion systems, and pathogenic mycobacteria have up to 5 of these systems, e.g., ESX-1 to ESX-5 (4). ESX-1 and ESX-5 are the best-studied systems. ESX-1

**Citation** Burggraaf MJ, Speer A, Meijers AS, Ummels R, van der Sar AM, Korotkov KV, Bitter W, Kuijl C. 2019. Type VII secretion substrates of pathogenic mycobacteria are processed by a surface protease. mBio 10:e01951-19. <https://doi.org/10.1128/mBio.01951-19>.

**Editor** Christina L. Stallings, Washington University School of Medicine in St. Louis

**Copyright** © 2019 Burggraaf et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Coenraad Kuijl, [c.kuijl@vumc.nl](mailto:c.kuijl@vumc.nl).

\* Present address: Astrid M. van der Sar, NWO, the Hague, Netherlands.

**Received** 24 July 2019

**Accepted** 30 September 2019

**Published** 29 October 2019

is known for its role in virulence, since the deletion of part of the *esx-1* locus, known as the RD-1 region, is the main cause of the attenuation of the *Mycobacterium bovis* BCG vaccine strain (5, 6). ESX-5 is required for the secretion of more than 50 proteins (7, 8) and has been described to be important for immune modulation (9–11) and nutrient uptake (12). Notably, the ESX-5 system is present only in slow-growing mycobacteria (13).

Together with the diversification of ESX systems, two large groups of proteins have evolved, known as the PE and PPE proteins. PE/PPE proteins are unique to mycobacteria and some closely related species and are especially abundant in pathogenic slow-growing species like *M. tuberculosis* and *Mycobacterium marinum* (14). Some of the PE- and PPE-encoding genes are present in the *esx* loci, while others are located elsewhere in the genome (15). These proteins are prominent ESX substrates, and the majority of these proteins are secreted through the ESX-5 system (7, 16). The currently available structural data show that PE and PPE proteins form heterodimers (17–19), which is hypothesized to be a general feature of these proteins.

PE/PPE proteins are characterized by a conserved N-terminal domain of 100 and 180 amino acids (aa), respectively, which includes the Pro-Glu or Pro-Pro-Glu motifs close to the N terminus after which they were named (20). Although both these protein families lack a classical signal peptide, the PE proteins do have a conserved ESX secretion motif, consisting of a helix-turn-helix domain directly followed by YxxxD/E (17).

Some PE proteins contain only a PE domain (PE<sub>only</sub>), while others have C-terminal extensions of variable size. The largest subfamily of PE proteins in *M. tuberculosis* contains a glycine-rich C-terminal domain encoded by so-called polymorphic GC-rich repetitive sequences (PGRS) and are therefore known as PE\_PGRS proteins. These proteins have been linked to immune modulation (9–11). Although the high prevalence of PE\_PGRS proteins in pathogenic mycobacteria and their secretion by the ESX-5 system hint at a role in virulence, the regulation and processing of this group of proteins remain unclear.

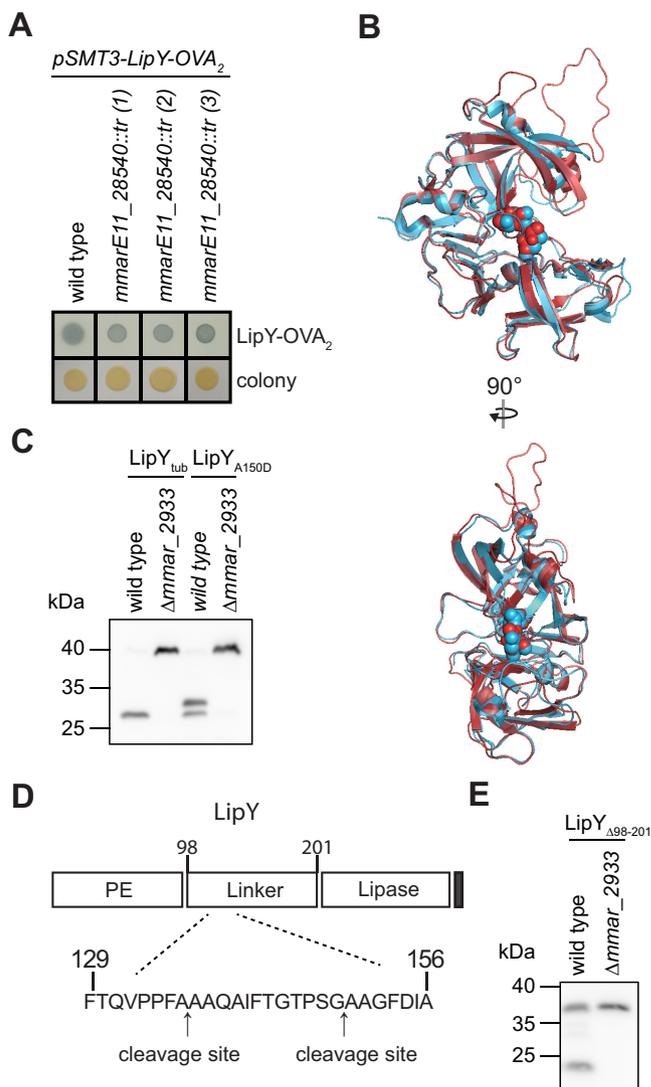
One of the characterized secreted PE proteins is the lipase LipY from *M. tuberculosis* (LipY<sub>tub</sub>). LipY<sub>tub</sub> functions as a surface-exposed triacylglycerol lipase. In *M. tuberculosis*, LipY is mainly expressed under hypoxic conditions and could be involved in lipid body formation (21, 22). LipY consists of a PE domain followed by a linker domain and the lipase domain (Fig. 1D). Previously, we showed that LipY<sub>tub</sub> is secreted by ESX-5 and subsequently processed in the linker domain (23).

In this study, we describe the identification and characterization of an aspartic protease, which we named PecA, that processes LipY<sub>tub</sub> and other PE\_PGRS proteins. Interestingly, deletion of *pecA* causes attenuation of *M. marinum* in zebrafish larvae.

## RESULTS

**Identification of MMARE11\_28540.** To study the secretion and processing of PE proteins, we used *M. marinum* and heterologous expression of the model protein LipY-OVA<sub>2</sub>, consisting of the PE domain of LipY linked to an ovalbumin fragment. Previously, we have shown that this fusion construct is secreted, which can be detected by double filter assay (24). To identify proteins involved in PE protein secretion, we generated and screened a transposon mutant library of *M. marinum* expressing LipY-OVA<sub>2</sub>. We expected to find mutants that were blocked in secretion because they contained a mutation in the gene coding for the LipY partner protein, but we did not identify these. Instead, we identified three transposon mutants that reproducibly displayed more intense and more discrete edges on the double filter than the wild-type strain (Fig. 1A).

Analysis of the transposon insertions revealed that all 3 mutants had independent transposon insertions in the same gene, i.e., *MMARE11\_28540*. The encoded protein belongs to the PE\_PGRS family and is 69.6% identical to PE\_PGRS35 (Rv1983) of *M. tuberculosis*. Somewhat weaker similarity (43% identity) is observed with the C-terminal domain of PE\_PGRS16. Interestingly, the structure of the C-terminal domain of PE\_PGRS16 has been solved and belongs to the family of aspartic proteases (25). Although



**FIG 1** MMARE11\_28540 is required for the processing of LipY. (A) Secretion of LipY-OVA<sub>2</sub> by *M. marinum* wild-type E11, *mmare11\_28540::tr*, and complemented strains was detected by double filter assay. (B) Alignment between model of MMARE11\_28540 (red, amino acids 269 to 551 out of 554) and crystal structure of PE\_PGRS16 (blue, 4EHC). Aspartic acids within the putative active site DTG/DSG motifs are depicted as spheres. (C) Immunoblot analysis of whole-cell lysate from *M. marinum* M wild-type and  $\Delta$ *mmar\_2933* strains expressing C-terminally HA-labeled LipY<sub>tub</sub> or LipY<sub>A150D</sub>. LipY was detected with anti-HA antibody. (D) The cleavage sites of MMAR\_2933 in the linker domain of LipY<sub>tub</sub> as determined by N-terminal Edman sequencing; the N-terminal cleavage site (aa 136A to 137A) and the previously identified C-terminal cleavage (aa 149G to 150A) (23) in LipY<sub>tub</sub> are indicated. (E) Immunoblot analysis of whole-cell lysate from *M. marinum* M wild-type and  $\Delta$ *mmar\_2933* strains expressing C-terminally HA-labeled LipY <sub>$\Delta$ 98-201</sub>. Proteins were detected with anti-HA antibody.

the crystal structure of PE\_PGRS16 was solved, no proteolytic activity was observed and putative substrates remained elusive. Based on homology modeling, MMARE11\_28540 has a similar structure, including the location of the DTG/DSG motifs that form the active site of aspartic proteases (Fig. 1B; see also Fig. S1A and B in the supplemental material).

**LipY processing.** To assess the function of MMARE11\_28540, we first analyzed the effect of this gene on full-length LipY<sub>tub</sub>. Previously, we have shown that LipY<sub>tub</sub> is cleaved in the linker domain after amino acid 149 (23). For follow-up experiments, we generated a clean deletion mutant of the *mmare11\_28540* orthologue in *M. marinum* M (*mmar\_2933*). We expressed LipY<sub>tub</sub> with a C-terminal hemagglutinin (HA) tag in *M. marinum* wild type and the  $\Delta$ *mmar\_2933* strain. Western blot analysis showed that

deletion of *mmar\_2933* completely abolished the processing of LipY<sub>tub</sub> (Fig. 1C), indicating that processing of LipY<sub>tub</sub> is performed by MMAR\_2933. From our previous experiments, we knew that there is an additional cleavage site in LipY<sub>tub</sub>, because mutation of the original cleavage site (by the A150D mutation) resulted in a slightly larger but still cleaved product of around 33 kDa (Fig. 1C) (23). Interestingly, cleavage at this additional site was also not observed in the *mmar\_2933* mutant, which indicates that processing at this alternative site is also mediated by this putative protease.

To determine the localization of this additional cleavage site, we isolated processed LipY<sub>A150D</sub> from the Genapol supernatant (GS) (Fig. S2) and performed Edman sequencing. The resulting sequence showed that LipY<sub>A150D</sub> is cleaved between two alanine residues at positions 136 and 137, resulting in the N-terminal sequence AAQAI (Fig. 1D). This cleavage site corresponds to the observed molecular weight of the processed fragment (~33 kDa).

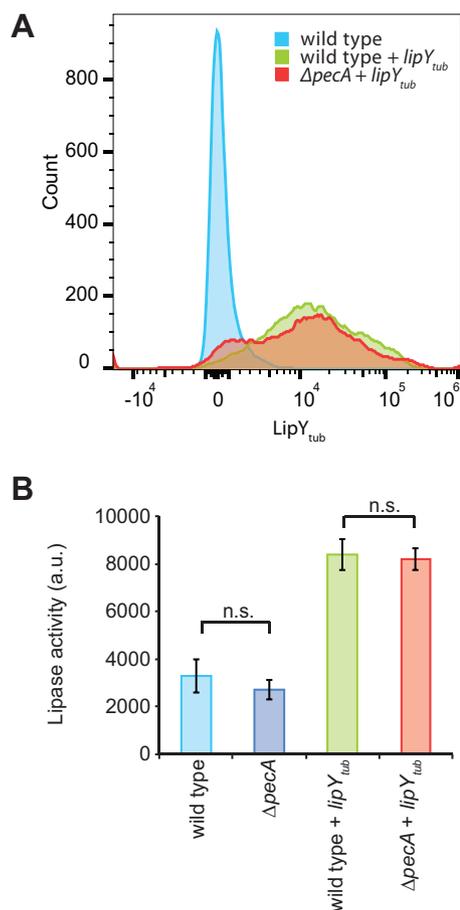
Cleavage of both LipY<sub>tub</sub> and LipY<sub>A150D</sub> takes place in the linker domain. However, LipY-OVA<sub>2</sub>, whose secretion was affected in *mmarE11\_28540::tr*, does not contain this linker domain. This hinted at the possibility that MMAR\_2933 also cleaves at a third location, inside the PE domain of LipY-OVA<sub>2</sub>. To explore cleavage inside the PE domain, we expressed LipY without the linker domain (LipY<sub>Δ98-201</sub>) in wild-type and *Δmmar\_2933* strains. In the wild-type strain, we observed both the full-length product of LipY<sub>Δ98-201</sub> at 36 kDa and a processed fragment with an apparent weight of 25 kDa (Fig. 1E). Similarly to the full-length protein, this shorter LipY<sub>tub</sub> product was not processed in the *Δmmar\_2933* mutant strain. This indicates that MMAR\_2933 can cleave LipY both inside the PE domain, before amino acid 98, and after amino acids 136 and 149. Strikingly, the apparent molecular weight of the N-terminal part that is cleaved off (11 kDa) corresponds to the size of the PE domain. These experiments demonstrate that MMAR\_2933 is a protease that cleaves PE proteins, and therefore, we would like to rename this protein **PE cleavage protein A (PecA)**.

**LipY surface localization and activity.** Since the majority of LipY in the Genapol supernatant fraction is processed by PecA, we questioned whether LipY processing is required for proper surface localization. This could also elucidate at which cellular location cleavage is taking place. Surface localization of LipY<sub>tub</sub> was analyzed by flow cytometry. This experiment showed that cell surface exposures of LipY<sub>tub</sub> were similar for the wild-type and the *ΔpecA* mutant strains (Fig. 2A), demonstrating that cleavage is not required for secretion and is likely to occur at the cell surface.

Previous research has shown that the presence of the PE domain reduces the lipase activity of recombinant LipY<sub>tub</sub> produced in *Escherichia coli* (26). To test whether the presence of the PE domain reduces lipase activity in its native environment, we analyzed the lipase activity of LipY<sub>tub</sub> in *M. marinum* wild-type and *ΔpecA* strains. Surprisingly, no change in lipase activity was observed between the wild type and the *ΔpecA* strain overexpressing LipY<sub>tub</sub> (Fig. 2B). Furthermore, lipase activity increased 2.7-fold upon LipY<sub>tub</sub> overexpression. Lipase activity was almost fully inhibited upon treatment with the lipase inhibitor paraoxon (POX) (Fig. S3), demonstrating the specificity of the lipase assay. Overall, our evidence suggests that processing of LipY<sub>tub</sub> by PecA does not affect LipY<sub>tub</sub> lipase activity in *M. marinum*.

**PecA cleavage.** PecA is a PE\_PGRS protein and has a similar domain structure as LipY<sub>tub</sub>: a PE domain followed by a linker domain and a C-terminal domain with enzymatic activity. To investigate whether PecA has the capacity to cleave itself, PecA with a C-terminal Myc tag was introduced in the *M. marinum* *ΔpecA* strain. Western blot analysis showed that PecA had an observed molecular weight of approximately 42 kDa, instead of the expected 53 kDa, suggesting the removal of the 11-kDa PE domain (Fig. 3A).

The active site of PecA was predicted to consist of two aspartic acid residues, the DTG and DSG motifs found at amino acids 293 to 295 and 470 to 472, respectively. These residues are positioned in the center of the putative cleft of the predicted protein structure (Fig. 1B). To confirm whether aspartic protease activity was required for

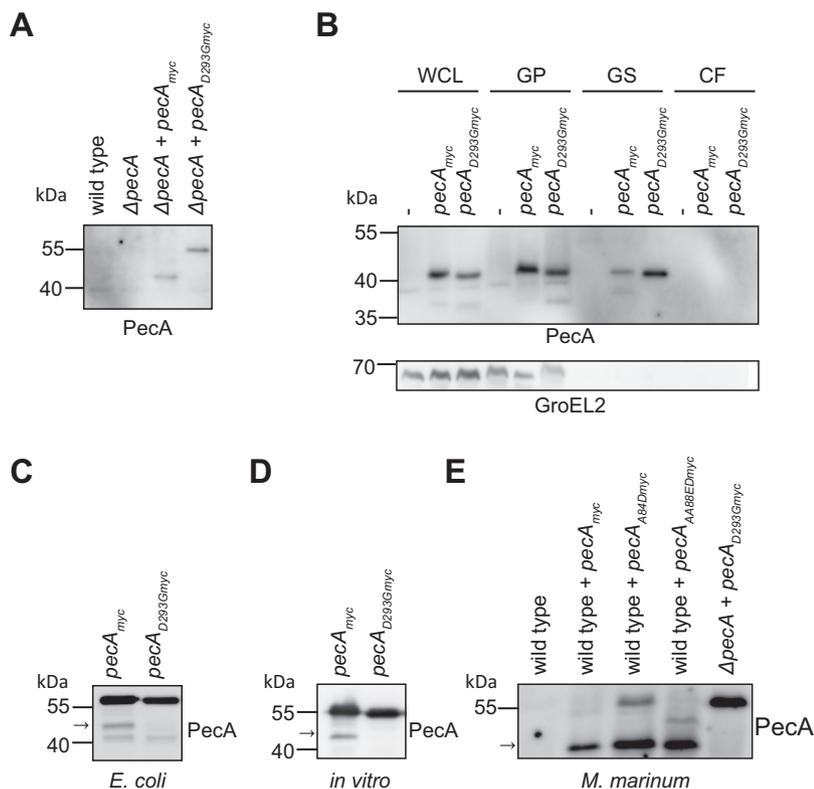


**FIG 2** PecA does not affect LipY surface localization and lipase activity. (A) Surface localization of LipY was measured on whole cells by flow cytometry for *M. marinum* M wild-type and  $\Delta pecA$  strains expressing LipY<sub>tub</sub>. LipY<sub>tub</sub> was detected via its C-terminal HA tag by anti-HA antibody (10,000 bacteria were counted per condition). (B) Lipase activity was measured by the fluorescently based DGGR assay. Standard deviation values are shown ( $n = 3$ ). a.u., arbitrary units; n.s., not significant.

proteolytic processing of PecA, we aimed to create a catalytically inactive site mutant by replacing the conserved aspartic acid residue at position 293 with glycine (PecA<sub>D293G</sub>). This mutated PecA was not cleaved when expressed in the  $\Delta pecA$  strain, demonstrating that the aspartic protease activity of PecA is indeed required for PecA self-cleavage (Fig. 3A).

Interestingly, when the same construct was expressed in wild-type *M. marinum*, PecA<sub>D293G</sub> was cleaved, revealing that PecA can be cleaved in *trans* (Fig. 3B). Furthermore, cleaved PecA is present in the Genapol supernatant (GS), but not in the culture supernatant, indicating that PecA is located at the mycobacterial surface (Fig. 3B, GS fraction).

**PecA is directly involved in PE processing.** To confirm that PecA is cleaved by itself and not by another protease, PecA was introduced in *E. coli*. When expressed in *E. coli*, PecA generated similarly sized cleavage products as observed in *M. marinum*. Again, the active site mutant PecA<sub>D293G</sub> did not produce the specific processed form seen for proteolytic active PecA (Fig. 3C). An additional fragment of lower molecular weight was detected for both PecA and PecA<sub>D293G</sub>, indicating that this fragment was produced by an *E. coli* protease. To obtain additional proof for self-cleavage of PecA, we performed *in vitro* transcription/translation experiments. Although PecA activity was low under these conditions, we could clearly observe that, also in this cell-free environment, PecA partially cleaved itself, while PecA<sub>D293G</sub> was not cleaved (Fig. 3D). Taken together, these results demonstrate that PecA is an aspartic protease that can cleave



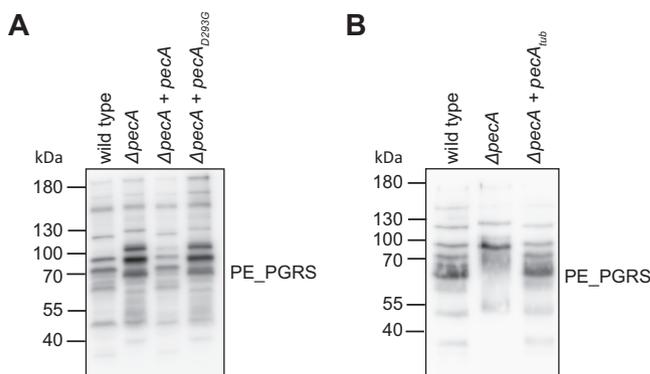
**FIG 3** PecA is subjected to self-cleavage. (A) Expression of PecA<sub>myc</sub> and PecA<sub>D293Gmyc</sub> in *M. marinum*  $\Delta pecA$  was detected by immunoblotting of whole-cell lysate. (B) Subcellular localization and secretion analysis of *M. marinum* wild type (indicated with "-") or wild type expressing PecA<sub>myc</sub> or PecA<sub>D293Gmyc</sub>. WCL, whole-cell lysate; GP, Genapol-treated pellet (membrane associated and intracellular); GS, Genapol supernatant (membrane associated); CF, culture filtrate. GroEL2 was probed as a cytosolic control. (C) Immunoblot analysis of PecA<sub>myc</sub> and PecA<sub>D293Gmyc</sub> expressed in *E. coli* (whole-cell lysate). (D) Immunoblot analysis of *in vitro* translation of PecA<sub>myc</sub> and PecA<sub>D293Gmyc</sub>. (E) Immunoblot analysis of *M. marinum* wild type; wild type expressing PecA<sub>myc</sub>, PecA<sub>A84Dmyc</sub> or PecA<sub>AA88EDmyc</sub>; and  $\Delta pecA$  mutant expressing PecA<sub>D293Gmyc</sub> (whole-cell lysate). The arrow indicates the fully cleaved form of PecA. Proteins were detected by their C-terminal Myc tag with an anti-Myc antibody.

not only LipY but also itself. PecA cleavage is dependent on the conserved aspartic protease residues and can occur in the absence of other mycobacterial proteins.

When we compared the two identified cleavage sites in the linker domain of LipY<sub>tub</sub>, the presence of two consecutive small amino acid (glycine or alanine) residues stood out (Fig. 1D). The variable residues within the YxxxD/E secretion motif of both LipY<sub>tub</sub> and PecA contain three consecutive alanine residues (YAAAE). In addition, proximal to the secretion motif is a conserved glycine alanine motif at position 84. The presence of consecutive small amino acids within the PE domain combined with the approximately 11-kDa size of the cleaved N-terminal domain led us to hypothesize that the PE domain of PecA and LipY<sub>tub</sub> is cleaved near or inside the secretion signal motif (17).

To test this hypothesis, we generated *pecA* constructs harboring point mutations near and in the region coding for the PE secretion signal (PecA<sub>A84D</sub> and PecA<sub>AA88ED</sub>) and studied its processing in *M. marinum*. The PecA<sub>A84D</sub> and PecA<sub>AA88ED</sub> mutants were both partially processed as wild-type PecA (indicated with arrow in Fig. 3E) but differed in size regarding to the secondary products. The secondary product of PecA<sub>A84D</sub> appeared at full length whereas the secondary product of PecA<sub>AA88ED</sub> was alternatively processed, shown by an additional band of ~47 kDa (Fig. 3E and Fig. S4). These data demonstrate that PecA cleaves itself within the PE domain.

**PecA processes PE\_PGRS proteins.** With the aid of ESX-5 mutant strains, we previously showed that PecA and LipY are both ESX-5 substrates (7, 23) and belong to the PE\_PGRS protein family. Based on the conservation of the PE domains among



**FIG 4** PecA is required for the processing of PE\_PGRS proteins. Immunoblotting of whole-cell lysate from *M. marinum* wild-type M strain,  $\Delta pecA$  strain, and  $\Delta pecA$  strain complemented with *M. marinum*  $pecA_{myc}$  and  $pecA_{D293G}$  (A) or *M. tuberculosis*  $pecA_{tub}$  (B) for endogenous PE\_PGRS proteins. PE\_PGRS proteins were detected with monoclonal antibodies recognizing the PGRS domain.

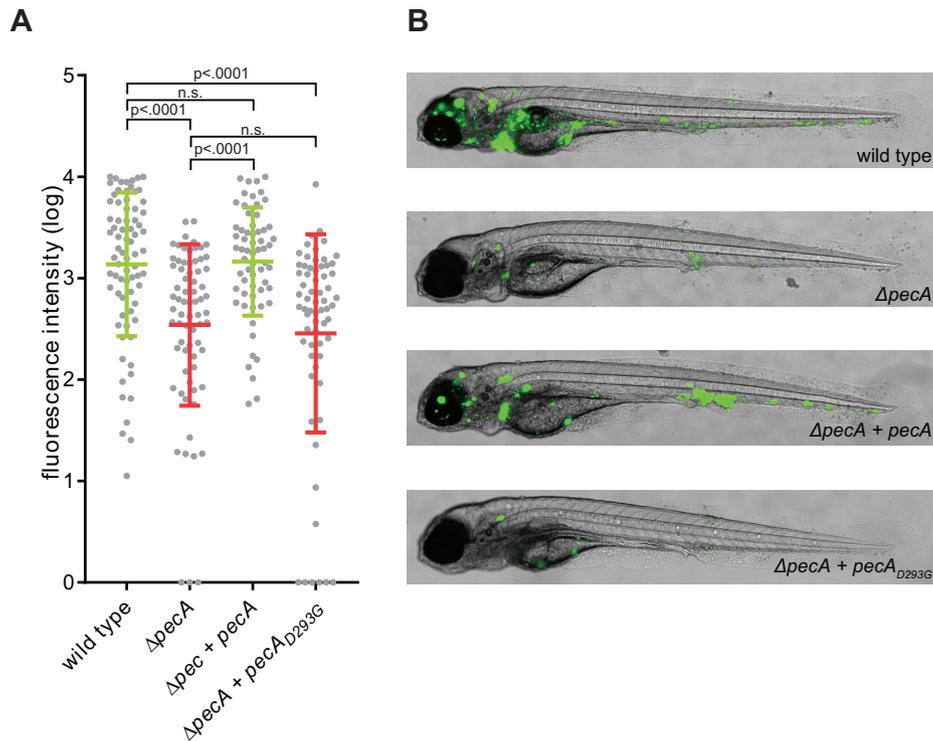
PE\_PGRS proteins and the proteolytic cleavage inside the PE domain of both LipY<sub>tub</sub> and PecA, we hypothesized that PecA could also cleave other PE\_PGRS proteins. We tested this hypothesis by immunoblotting lysates of wild-type and  $\Delta pecA$  strains using a monoclonal antibody directed against the PGRS domain, recognizing various PE\_PGRS proteins (7). Although secretion of PE\_PGRS proteins seemed to be normal in all strains tested, the PE\_PGRS protein pattern of the *M. marinum*  $\Delta pecA$  mutant appeared different from the wild type (Fig. 4A). Specifically, the  $\Delta pecA$  strain showed several proteins with an increased molecular weight. The different pattern shows an increase of approximately  $\sim 11$  kDa for each PE\_PGRS protein, corresponding with the size of a PE domain.

This indicates that several or perhaps most PE\_PGRS proteins that are secreted under these conditions are dependent on PecA for their processing. Importantly, the PE\_PGRS pattern on an immunoblot was restored to wild type when PecA was expressed in the  $\Delta pecA$  strain. Notably, the expression of the active site mutant  $pecA_{D293G}$  did not restore the PGRS phenotype, confirming that the aspartic protease activity is required for proteolytic processing of PE\_PGRS proteins.

***M. tuberculosis* PecA complements the *M. marinum*  $\Delta pecA$  mutant.** As mentioned previously, the genome of *M. tuberculosis* contains a gene (*Rv1983*, *PE\_PGRS35*) with high sequence identity (69.6% identical amino acids) and conserved synteny. The identity is especially conserved in the protease domain, implying a similar function. To investigate whether PE\_PGRS35 processes PE\_PGRS proteins in a similar fashion, we expressed this gene in *M. marinum*  $\Delta pecA$  and visualized the PE\_PGRS proteins by immunoblotting. We observed full complementation, based on the identical PE\_PGRS patterns between the complemented strain and the wild-type *M. marinum* strain (Fig. 4B). We conclude that PE\_PGRS35 from *M. tuberculosis* has proteolytic activity toward PE\_PGRS proteins from *M. marinum* and potentially *M. tuberculosis*, and therefore, we designate it PecA<sub>tub</sub>.

**PecA is a virulence factor *in vivo*.** PE\_PGRS proteins have been implicated to play a role in immune modulation, macrophage invasion, and B-cell responses (9, 27). However, the high similarity between members of this family implies functional redundancy, which hinders their functional resolution. Since PecA processes many of these proteins, including itself, we assessed whether its proteolytic activity contributes to virulence *in vivo*.

We first established that the absence of PecA has no negative effect on replication *in vitro* (Fig. S5A). Next, we monitored intracellular replication in RAW 264.7 macrophages and *Acanthamoeba castellanii* with mEosFP-expressing wild-type *M. marinum* and  $\Delta pecA$  strains. Although PecA is responsible for the processing of many PE\_PGRS proteins, no or only very small differences in intracellular replication were observed



**FIG 5** Deletion of PecA causes attenuation of virulence in an *in vivo* zebrafish infection model. (A) A total of 273 zebrafish larvae were infected 1 day postfertilization (dpf) with wild-type (75 larvae),  $\Delta pecA$  (71 larvae),  $\Delta pecA$  plus  $pecA_{myc}$  (62 larvae), or  $\Delta pecA$  plus  $\Delta pecA_{D293Gmyc}$  (65 larvae) *M. marinum* M strains (see Table S2 for input CFU). Graph shows relative infection from 4 independent experiments as calculated by normalized integrated fluorescent intensity per larva at 4 days postinfection, with mean intensity and standard deviation for the entire population. Similar effects were obtained for CFU counts at 4 dpi with wild-type,  $\Delta pecA$ , and  $\Delta pecA$  plus  $pecA_{myc}$  strains (Fig. S6). (B) Representative images of infected larvae are shown for each strain.

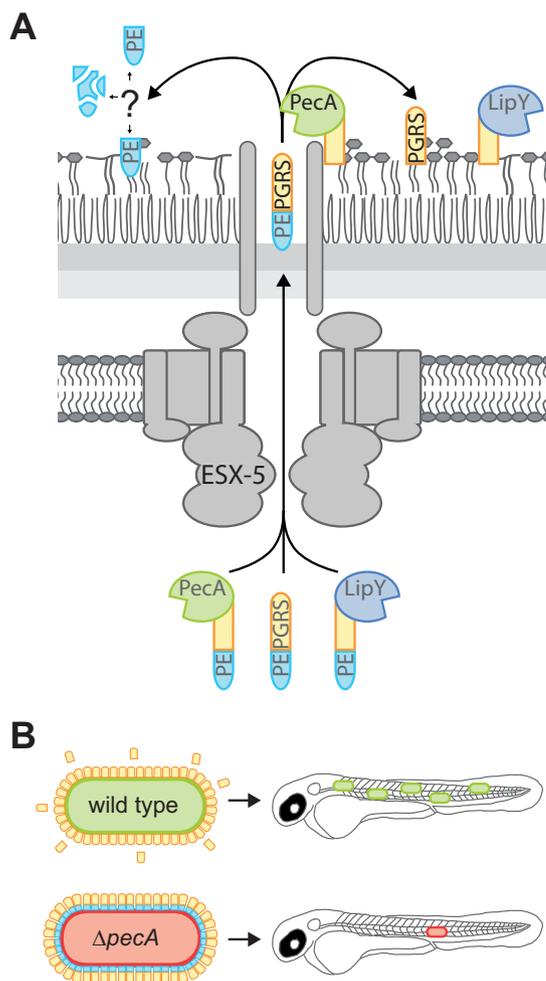
between the wild-type and the  $\Delta pecA$  strains in the phagocytic host cells studied (Fig. S5B to D).

To assess the role of PecA in an animal model, zebrafish larvae were infected with *M. marinum* wild-type and  $\Delta pecA$  strains. In contrast to the experiments with phagocytic cells, the  $\Delta pecA$  strain showed reduced outgrowth (5-fold change) after 4 days compared to wild-type bacteria, when measured by fluorescence microscopy (Fig. 5A and B) and CFU plating (Fig. S6). This phenotype was fully complemented by reintroducing  $pecA$  in the deletion mutant when measured by fluorescence microscopy and partially when measured by CFU counting. Importantly, the active site mutant  $PecA_{D293G}$  failed to complement the  $\Delta pecA$  phenotype, demonstrating that proteolytic activity of PecA is directly contributing to virulence *in vivo*.

## DISCUSSION

For most bacterial pathogens, secreted and surface-associated proteins play a crucial role in host-pathogen interactions. The largest families of secreted proteins in pathogenic mycobacteria are the PE and PPE protein families, whose role in host-pathogen interactions remains enigmatic (9). A large subset of the PE proteins belongs to the PE\_PGRS family and is secreted by the ESX-5 system (7, 8). PE\_PGRS proteins have a conserved N-terminal PE domain (100 amino acids [aa], ~11 kDa) containing the YxxD/E secretion motif and a larger C-terminal domain that is characterized by GG AGGX repeats. In this study, we have provided new mechanistic insights into the processing of *M. marinum*-secreted PE\_PGRS proteins by PecA, a PE\_PGRS protein with aspartic protease activity (Fig. 6).

Previously, Barathy and Suguna (25) resolved the structure of the C-terminal domain



**FIG 6** Model: PecA processes PE-PGRS proteins and is important for virulence. (A) PecA removes the PE domain of PE-PGRS proteins, including LipY<sub>tub</sub> and itself, at the surface of *Mycobacterium marinum*. The processed PE\_PGRS proteins might be of importance during infection, although their exact role, as well as the fate of the PE domains, remains unknown. (B) Proteolytic activity of PecA is suggested to play a role in virulence of *M. marinum*, as was observed during infection in zebrafish larvae.

of PE\_PGRS16 (Rv0977) and showed that it has an aspartic protease fold. In the same study, they described two other PE proteins of *M. tuberculosis* that could have aspartic protease activity: PE26 (Rv2519) and PE\_PGRS35 (Rv1983, PecA<sub>tub</sub>). All three proteins contain the DTG/DSG sequences that are characteristic for the catalytic domain of aspartic proteases. Whereas this previous study was unable to demonstrate proteolytic activity for PE\_PGRS16, in our study we conclusively show that PecA is an active protease that is essential to process PE\_PGRS proteins at the cell surface.

LipY<sub>tub</sub>, one of the PecA substrates, contains at least three different processing sites, two in the linker domain and one near the end of the PE domain. This extensive processing could be an exception, since we observe an apparent molecular weight difference consistent with the removal of the PE domain (or part thereof) for PecA and other PE\_PGRS proteins. Mutations near or inside the secretion signal motif YxxxD/E of PecA resulted in reduced and even altered processing of the PE domain, demonstrating that the PE domain is cleaved by PecA. We were, however, unable to create a mutant that fully blocked processing by PecA. This is not unexpected, since cleavage sites of aspartic proteases, such as HIV-1 protease, are known to be difficult to predict. The HIV-1 protease cleavage sites not only rely on sequence but also are determined by local structure and physicochemical features of the cleavage site (28).

As mentioned above, the processing of other PE\_PGRS proteins also seems to

indicate the removal of a PE domain-sized fragment. However, some caution is warranted since we did not confirm the identity of the PE\_PGRS proteins visualized by Western blotting. Whether the PE domain remains intact after cleavage and associated with the bacterium to fulfill a function beyond enabling secretion of the C-terminal domain remains a question to be answered.

We found evidence that processing of PE proteins takes place at the cell surface. First, PecA itself is an ESX-5 substrate, since the culture filtrate of an *esx-5* mutant of *M. marinum* lacks PecA (MMAR\_2933), as Abdallah et al. demonstrated by mass spectrometry (MS) analysis (7). Second, processing of LipY does not affect its surface localization and lipase activity. This last result is in sharp contrast with results obtained with recombinant LipY, which showed increased lipase activity upon removal of the PE domain (22, 26). To reconcile these observed differences, we speculate that the PE domain is interacting with (glyco)lipids in the mycobacterial membrane, preventing inhibition of lipase activity. However, prior to secretion by the ESX-5 system, the PE domain could be inhibiting LipY to prevent undesired intracellular lipase activity. Similarly, PecA exhibited strongly reduced protease activity, measured by self-cleavage, when expressed in *E. coli* or *in vitro*. This suggests that its natural surrounding at the cell surface is important for optimal activity. Improper folding could be an alternative explanation for the observed reduction in enzymatic activities of both LipY and PecA when produced in *E. coli*.

Although no obvious difference in intracellular growth was observed in phagocytic cells, deletion of *pecA* in *M. marinum* resulted in moderate attenuation in the zebrafish larva infection model. This phenotype was restored by introducing the wild-type gene coding for PecA but not by the inactive PecA<sub>D293G</sub>, indicating that substrate processing affects their biological function. It would be interesting to identify which PE or PE\_PGRS proteins are required to be cleaved to function as a virulence factor. However, this is challenging, since there are as many as 110 PE\_PGRS proteins in *M. marinum* (63 in *M. tuberculosis*), and the exact role of each PE\_PGRS proteins during infection is unknown. Previously, we have demonstrated that PE\_PGRS proteins reduce virulence, since strains lacking PE\_PGRS (and PPE-MPTR) secretion, exhibited increased virulence (16). These contradictory results could be explained by the possibility that PecA might have other substrates in addition to PE\_PGRS proteins that are involved in virulence. Follow-up research has to reveal whether PecA substrate specificity goes beyond the PE\_PGRS proteins or even beyond ESX-5 substrates. By looking into the substrate specificity and the role of these different substrates, we will further elucidate the role of PE\_PGRS proteins and PecA during infection.

PecA from *M. marinum* shares 69.6% identity with PecA from *M. tuberculosis*. And indeed, we showed that a *pecA* mutant in *M. marinum* was complemented with *pecA<sub>tub</sub>*. It is tentative to speculate that PecA<sub>tub</sub> plays a similar function in virulence as PecA from *M. marinum*. However, possibly PecA in *M. tuberculosis* has (partial) redundancy with the two homologs, i.e., PE26 and PE\_PGRS16.

Next to the attenuation *in vivo*, the importance of PecA in pathogenic mycobacteria can be illustrated by its high degree of conservation in the mycobacterial genome. Most PE\_PGRS proteins show high variability between clinical *M. tuberculosis* strains and accumulate mutations faster than expected (29). In contrast, PecA is one of the most conserved PE\_PGRS proteins. The mutation rate of PecA is even lower than the average mutation rate of the entire genome, again suggesting an important and conserved role for this protease (30).

Taken together, we show that PecA is a unique aspartic protease, in respect to its bacterial origin and extracellular localization. PecA is required for the processing of many PE\_PGRS proteins, the largest family of secreted proteins in pathogenic mycobacteria. Future investigations will have to reveal whether PecA-dependent processing is limited to PE\_PGRS proteins and will address the role of PecA and its two homologs in *M. tuberculosis*.

## MATERIALS AND METHODS

**Bacterial strains and growth conditions.** *Mycobacterium marinum* E11 (31) and M (32) were grown on 7H10 agar plates supplemented with Middlebrook oleic acid-albumin-dextrose-catalase (OADC) or in 7H9 broth (Difco-BD Biosciences) supplemented with Middlebrook ADC and 0.05% Tween 80 (Sigma) at 30°C with shaking at 90 rpm. *M. tuberculosis* CDC1551 was grown in 7H9 supplemented with Middlebrook ADC and 0.05% Tween 80 (Sigma) at 37°C.

*Escherichia coli* DH5 $\alpha$  and *E. coli* BL21 were grown in LB medium (Difco-BD Biosciences) at 37°C with shaking at 200 rpm.

For all cultures, hygromycin was used at a concentration of 50  $\mu$ g/ml, kanamycin was used at 25  $\mu$ g/ml, and streptomycin was used at 30  $\mu$ g/ml.

**Construction of *pecA* deletion mutant.** *M. marinum*  $\Delta$ *pecA* was constructed using the mycobacteriophage method developed by the Jacobs laboratory (33) and described in reference 12. Briefly, *MMAR\_2933* was deleted by synthesizing fragments of flanking regions of *MMAR\_2933* by PCR (primer set *MMAR\_2933* LF and *MMAR\_2933* LR for the 5' region and primer set *MMAR\_2933* RF and *MMAR\_2933* RR for the 3' region; primers are listed in Table S1 in the supplemental material). Resulting amplicons were digested with AlwNI and DraIII, cloned into the Van911-digested p0004s plasmid containing a hygromycin resistance cassette and the *sacB* gene to enable selection on sucrose sensitivity, introduced to the PacI site of phasmid pHA159, and electroporated into *Mycobacterium smegmatis* mc<sup>2</sup>155 according to the method of Bardarov et al. (34). The phage was isolated in high titers and incubated with *M. marinum* strain M. Colonies were selected on hygromycin plates and verified for sucrose sensitivity. The deletion was confirmed with PCR and sequencing.

**Plasmid construction.** The secretion-optimized synthetic LipY<sub>tub</sub> construct, named pSMT3-*lipY-OVA*<sub>2</sub>, contains amino acids 1 to 98 and 201 to 205 from LipY<sub>tub</sub> fused to a fragment of chick ovalbumin (amino acids 259 to 357). This construct was identified by screening for enhanced secretion after random mutagenesis of LipY<sub>tub</sub>-1–205 fused to chick ovalbumin (amino acids 259 to 357) in the pSMT3 plasmid. This work is described in the work of Burggraaf et al. (24).

To create pMV361-*pecA*, pMV361-*eccB<sub>5</sub>-mycP<sub>5</sub>-HA* (described by reference 35) was restricted with EcoRI and HindIII. *MMARE11\_28540* (*MMAR\_2933*, *pecA*) was amplified by PCR using primers PecA-pMV-Fw and PecA-pMV-Rv (primers are listed in Table S1) and subsequently inserted by In-Fusion cloning (TaKaRa Bio). After restriction with XhoI and HindIII to introduce an annealed oligonucleotide containing the Myc tag, the mutations D293G, A84D, and AA88ED were introduced using site-directed mutagenesis (Agilent Technologies). pST5552-*pecA<sub>tub</sub>* was created by restriction cloning. *PecA<sub>tub</sub>* was PCR amplified from genomic H37Rv DNA with the primers Rv1983-Fw and Rv1983-Rv. The PCR product was digested with EcoRI and XhoI and ligated into the pST5552 plasmid (a pMV261-based vector with a theophylline-inducible riboswitch) digested with EcoRI and Sall. The plasmids were verified with sequencing.

To enable *in vitro* transcription, pET28A-*iCAP* was restricted with XhoI and NcoI. *pecA* and *pecA<sub>D293G</sub>* were amplified by PCR from pMV-*pecA* and pMV-*pecA<sub>D293G</sub>*, respectively, using primers PecA-pET-Fw and PecA-pET-Rv and introduced in pET28A (NcoI/XhoI) by infusion cloning.

Plasmids were introduced to *M. marinum* by electroporation and to *E. coli* DH5 $\alpha$  by transformation according to standard heat shock protocol.

**Transposon mutagenesis and double filter screening.** A transposon library was generated for strain *M. marinum* E11 transformed with pSMT3-*lipY-OVA*<sub>2</sub> (published in reference 24) using the mycobacterium-specific phage phiMycMarT7 containing mariner-like transposon Himar1 (36).

After transposon mutagenesis, *M. marinum* pSMT3-*lipY-OVA*<sub>2</sub> was grown on a nitrocellulose membrane (0.45  $\mu$ m; Merck Millipore) placed on a 7H10 plate supplemented with Middlebrook OADC (Difco-BD Biosciences). Plates were incubated at 30°C. When colonies formed, the membrane was transferred onto a new membrane already placed on a fresh plate and incubated overnight at 30°C. The second membrane was incubated with antibodies directed against the influenza virus hemagglutinin epitope (HA.11; Covance) and the secondary antibody horseradish peroxidase-conjugated goat anti-mouse IgG (American Qualex Antibodies). Signal was detected with enhanced chemiluminescence (ECL) Western blotting detection reagent (Amersham Bioscience). Selected transposon mutants and complemented strains were grown to an optical density at 600 nm (OD<sub>600</sub>) of 0.8 to 1.2 in 7H9 supplemented with Middlebrook ADC (Difco Biosciences), 10- $\mu$ l drops were placed on the nitrocellulose filter on a 7H10 plate, and the procedure for analyzing secretion was repeated.

**In vitro transcription and translation.** pET28A-*pecA* and pET28A-*pecA<sub>D293G</sub>* were expressed with the PURExpress *in vitro* protein synthesis kit (New England Biolabs [NEB]) according to the manufacturer's protocol. Briefly, 100 ng plasmid DNA was added to a mixture of solutions A and B, supplemented with RNase inhibitor (Thermo Scientific). The reaction mixture was incubated for 2 h at 37°C and subsequently put on ice for 15 min. Solubilization/denaturation (SD) buffer (containing 100 mM dithiothreitol and 2% SDS) was added before loading samples on the gel for SDS-PAGE.

**Secretion analysis.** One day prior to making samples, cultures were washed with 7H9-0.2% dextrose-0.2% glycerol-0.05% Tween 80 to remove ADC. Cultures were grown in this medium at 30°C, with shaking at 90 rpm to an OD<sub>600</sub> of 0.8 to 1.2. Cultures were harvested, and the supernatant was filtered through an 0.2- $\mu$ m Millipore filter and precipitated with 10% trichloroacetic acid (Sigma) and 36  $\mu$ g/ml tRNA (Roche). Pellets were washed with phosphate-buffered saline (PBS). Two OD units was taken and dissolved in SD buffer to a concentration of 0.02 OD/ $\mu$ l, making the pellet sample. Two OD units was incubated with 0.5% Genapol X-080 (in PBS) for 30 min at room temperature (RT). After spinning down, the supernatant was taken and diluted with SD buffer to a final concentration of 0.016 OD/ $\mu$ l, forming the Genapol supernatant fraction. The remaining pellet was washed with PBS and dissolved in SD buffer

to a concentration of 0.02 OD/ $\mu$ l, forming the Genapol pellet fraction. Pellet samples were sonicated. All samples were incubated 5 min at 98°C.

**SDS-PAGE and immunoblotting.** Samples were loaded on a 12.5% SDS-PAGE gel. Subsequently, proteins were transferred to nitrocellulose membranes for Western blotting. Blots were blocked with 5% skimmed milk in PBS and incubated with antibodies recognizing the influenza virus hemagglutinin tag (HA.11; Covance), GroEL2 (CCs44; J. Belisle, NIH, Bethesda, MD, USA), PGRS domain (7), or the Myc tag (dMyc; Abcam). Secondary antibodies used were horseradish peroxidase conjugated to goat anti-mouse IgG or to goat anti-rabbit (Rockland). Secondary antibodies were detected with chemiluminescence (ECL).

**Flow cytometry analysis.** Bacteria were grown to an OD<sub>600</sub> of 0.8 to 1.2 in 7H9 with ADC and without Tween 80. Bacteria were pelleted, washed with PBS with 0.5% bovine serum albumin (BSA; Sigma), and incubated for 1 h with antibodies recognizing the influenza virus hemagglutinin tag (HA.11; Covance). After washing with PBS with 0.5% BSA, bacteria were incubated with secondary goat anti-mouse IgG conjugated to Alexa 647 antibodies for 1 h. After washing with PBS with 0.5% BSA, bacteria were analyzed by flow cytometry (Attune NxT; ThermoFisher). As a control, bacteria were incubated only with the secondary antibodies.

**Bioinformatic analysis.** Data obtained with the lipase assay were analyzed with an unpaired *t* test. The fluorescent intensity data obtained from the zebrafish experiments were normalized and log<sub>10</sub> transformed to obtain a Gaussian distribution. Normalization was performed by subtracting the geometric mean of uninfected larvae and setting the maximum value for each experiment arbitrarily to 10,000. Larvae with a negative fluorescence were set to 1 to allow for log<sub>10</sub> transformation. A 1-way analysis of variance (ANOVA) was performed on the transformed data and on the CFU data to test the null hypothesis: there is no difference between the groups and equality between means. When statistical significance (*P* < 0.05) is reached, the *P* value is specified.

**Structure prediction.** The homology model of the protease domain of PecA was obtained using RaptorX (37) and the crystal structure of the protease domain of PE\_PGRS16 (PDB identifier [ID] 4EHC) as a template (25). The model was refined using ModRefiner (38). The model was adjusted to remove clashes using Coot (39) and energy minimized using Phenix (40).

**N-terminal sequencing.** Bacterial cultures were washed with 7H9-0.2% dextrose-0.2% glycerol and grown to an OD<sub>600</sub> of 0.8 to 1.0. Bacterial pellets were incubated with 0.5% Genapol X-080 for 30 min at RT. The supernatant was collected and precipitated with 10% trichloroacetic acid (TCA) and 29  $\mu$ g/ml tRNA. Samples were incubated 1 h on ice, before being washed with acetone. The supernatant was discarded, and pellets were dried at RT. Pellets were dissolved in SD buffer and incubated for 5 min at 98°C.

Samples were run on a 12.5% SDS-PAGE gel supplemented with 0.4 mM gluconic acid (Merck). The gel was blotted on a polyvinylidene difluoride (PVDF) membrane (Bio-Rad). The blot was stained with 0.1% Coomassie brilliant blue R250, 50% methanol, 1% acetic acid, and destained with 50% methanol. The band corresponding to the cleaved product of LipY<sub>A150D</sub> was excised, and the 5 N-terminal amino acids were analyzed by Edman N-terminal sequencing (Alphalyse, Odense, Denmark).

**Lipase activity assay.** Bacteria were grown to mid-log phase in 7H9 medium (Difco) supplemented with 10% ADC (BD Bioscience) with 0.05% Tween 80 and diluted to an OD<sub>600</sub> of 0.2. One hundred sixty microliters of bacterial suspension was mixed with 25  $\mu$ l development solution DGGR [10 mg/ml BSA, 120 mM NaCl, 50 mM Tris, pH 7.5, 1% Triton X-100, 4  $\mu$ M 1,2-di-*o*-lauryl-*rac*-glycero-3-(glutaric acid 6-methylresorufin ester)]. Fluorescence was measured in a plate reader at an excitation of 530 nm and emission of 600 nm. Paraoxon (POX; 10  $\mu$ M; Sigma) is a lipase inhibitor, used as a control for inhibition of LipY lipase activity.

**Zebrafish larva infection procedure.** Transparent casper zebrafish larvae were removed from their chorion with tweezers and infected at 1 day postfertilization (dpf) via the caudal vein with a bacterial suspension. The injected CFU was determined before and after injection of the larvae to rule out blockage of the needle (see Table S2 in the supplemental material). Injection stocks were prepared by growing bacteria until the logarithmic phase (OD<sub>600</sub> of 0.7 to 1). Bacteria were spun down at low speed for 1 min to remove the largest clumps, washed with 0.3% Tween 80 in phosphate-buffered saline (PBS), and sonicated briefly for declumping. Bacteria were resuspended in PBS with 20% glycerol and 2% polyvinylpyrrolidone and stored at -80°C. Before use, bacteria were resuspended in PBS containing 0.17% (vol/vol) phenol red (Sigma) to aid visualization of the injection process. To determine the exact number of bacteria injected, the injection volume was plated on 7H10 plates containing the proper antibiotic selection. At 4 days postinfection (dpi), larvae were analyzed with an Olympus IX83 fluorescence microscope. Bright-field and fluorescence images were acquired with an Olympus IX83 microscope equipped with an Orca-flash 4.0 LT camera. Infection levels were quantified with CellProfiler 3.15. Zebrafish were manually delineated followed by integrating the fluorescent intensity of the entire fish.

**CFU counting.** Viable bacterial counts present in whole larvae were determined by plating larva homogenates, decontaminated with BBL MycoPrep (BD), on Middlebrook 7H10 agar.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mBio.01951-19>.

**FIG S1**, PDF file, 1.9 MB.

**FIG S2**, PDF file, 0.3 MB.

**FIG S3**, PDF file, 0.1 MB.

**FIG S4**, PDF file, 0.4 MB.

**FIG S5**, PDF file, 0.5 MB.

**FIG S6**, PDF file, 0.4 MB.

**TABLE S1**, PDF file, 0.1 MB.

**TABLE S2**, PDF file, 0.1 MB.

## ACKNOWLEDGMENTS

We thank Peter van Ulsen for his help with sample preparation for N-terminal sequencing. We thank Marion Sparrius for phage preparation for the transposon screen.

This work was supported by funding from the National Institute of Allergy and Infectious Diseases (R01AI119022 to K.V.K.), the Cancer Center Amsterdam (to M.J.B. and A.S.M.) and by the Netherlands Organization for Scientific Research (NWO) through a VENI grant (016.Veni.171.090 to A.S.).

## REFERENCES

- WHO. 2018. Global tuberculosis report 2018. World Health Organization, Geneva, Switzerland.
- Bottai D, Groschel M, Brosch R. 2017. Type VII secretion systems in Gram-positive bacteria, p 235–266. In Bagnoli F, Rappuoli R (ed), Protein and sugar export and assembly in Gram-positive bacteria. Springer Nature, Cham, Switzerland.
- Houben EN, Korotkov KV, Bitter W. 2014. Take five—type VII secretion systems of mycobacteria. *Biochim Biophys Acta* 1843:1707–1716. <https://doi.org/10.1016/j.bbamcr.2013.11.003>.
- Bitter W, Houben EN, Bottai D, Brodin P, Brown EJ, Cox JS, Derbyshire K, Fortune SM, Gao LY, Liu J, Gey van Pittius NC, Pym AS, Rubin EJ, Sherman DR, Cole ST, Brosch R. 2009. Systematic genetic nomenclature for type VII secretion systems. *PLoS Pathog* 5:e1000507. <https://doi.org/10.1371/journal.ppat.1000507>.
- Pym AS, Brodin P, Brosch R, Huerre M, Cole ST. 2002. Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. *Mol Microbiol* 46: 709–717. <https://doi.org/10.1046/j.1365-2958.2002.03237.x>.
- Lewis KN, Liao R, Guinn KM, Hickey MJ, Smith S, Behr MA, Sherman DR. 2003. Deletion of RD1 from *Mycobacterium tuberculosis* mimics Bacille Calmette-Guérin attenuation. *J Infect Dis* 187:117–123. <https://doi.org/10.1086/345862>.
- Abdallah AM, Verboom T, Weerdenburg EM, Gey van Pittius NC, Mahasha PW, Jimenez C, Parra M, Cadieux N, Brennan MJ, Appelmelk BJ, Bitter W. 2009. PPE and PE\_PGRS proteins of *Mycobacterium marinum* are transported via the type VII secretion system ESX-5. *Mol Microbiol* 73:329–340. <https://doi.org/10.1111/j.1365-2958.2009.06783.x>.
- Ates LS, Sayes F, Frigui W, Ummels R, Damen MPM, Bottai D, Behr MA, van Heijst JWJ, Bitter W, Majlessi L, Brosch R. 2018. RD5-mediated lack of PE\_PGRS and PPE-MPTR export in BCG vaccine strains results in strong reduction of antigenic repertoire but little impact on protection. *PLoS Pathog* 14:e1007139. <https://doi.org/10.1371/journal.ppat.1007139>.
- Sampson SL. 2011. Mycobacterial PE/PPE proteins at the host-pathogen interface. *Clin Dev Immunol* 2011:497203. <https://doi.org/10.1155/2011/497203>.
- Saini NK, Baena A, Ng TW, Venkataswamy MM, Kennedy SC, Kunnath-Velayudhan S, Carreno LJ, Xu J, Chan J, Larsen MH, Jacobs WR, Jr, Porcelli SA. 2016. Suppression of autophagy and antigen presentation by *Mycobacterium tuberculosis* PE\_PGRS47. *Nat Microbiol* 1:16133. <https://doi.org/10.1038/nmicrobiol.2016.133>.
- Brennan MJ, Delogu G. 2002. The PE multigene family: a ‘molecular mantra’ for mycobacteria. *Trends Microbiol* 10:246–249. [https://doi.org/10.1016/S0966-842X\(02\)02335-1](https://doi.org/10.1016/S0966-842X(02)02335-1).
- Ates LS, Ummels R, Commandeur S, van de Weerd R, Sparrius M, Weerdenburg E, Alber M, Kalscheuer R, Piersma SR, Abdallah AM, Abd El Ghany M, Abdel-Haleem AM, Pain A, Jiménez CR, Bitter W, Houben ENG. 2015. Essential role of the ESX-5 secretion system in outer membrane permeability of pathogenic mycobacteria. *PLoS Genet* 11:e1005190. <https://doi.org/10.1371/journal.pgen.1005190>.
- Abdallah AM, Verboom T, Hannes F, Safi M, Strong M, Eisenberg D, Musters RJ, Vandenbroucke-Grauls CM, Appelmelk BJ, Luirink J, Bitter W. 2006. A specific secretion system mediates PPE41 transport in pathogenic mycobacteria. *Mol Microbiol* 62:667–679. <https://doi.org/10.1111/j.1365-2958.2006.05409.x>.
- Gey van Pittius NC, Sampson SL, Lee H, Kim Y, van Helden PD, Warren RM. 2006. Evolution and expansion of the *Mycobacterium tuberculosis* PE and PPE multigene families and their association with the duplication of the ESAT-6 (esx) gene cluster regions. *BMC Evol Biol* 6:95. <https://doi.org/10.1186/1471-2148-6-95>.
- Stinear TP, Seemann T, Harrison PF, Jenkin GA, Davies JK, Johnson PDR, Abdallah Z, Arrowsmith C, Chillingworth T, Churcher C, Clarke K, Cronin A, Davis P, Goodhead I, Holroyd N, Jagels K, Lord A, Moule S, Mungall K, Norbertczak H, Quail MA, Rabbinowitsch E, Walker D, White B, Whitehead S, Small PLC, Brosch R, Ramakrishnan L, Fischbach MA, Parkhill J, Cole ST. 2008. Insights from the complete genome sequence of *Mycobacterium marinum* on the evolution of *Mycobacterium tuberculosis*. *Genome Res* 18:729–741. <https://doi.org/10.1101/gr.075069.107>.
- Ates LS, Dippenaar A, Ummels R, Piersma SR, van der Woude AD, van der Kuij K, Le Chevalier F, Mata-Espinosa D, Barrios-Payán J, Marquina-Castillo B, Guapillo C, Jiménez CR, Pain A, Houben ENG, Warren RM, Brosch R, Hernández-Pando R, Bitter W. 2018. Mutations in *ppe38* block PE\_PGRS secretion and increase virulence of *Mycobacterium tuberculosis*. *Nat Microbiol* 3:181–188. <https://doi.org/10.1038/s41564-017-0090-6>.
- Daleke MH, Ummels R, Bawono P, Heringa J, Vandenbroucke-Grauls CM, Luirink J, Bitter W. 2012. General secretion signal for the mycobacterial type VII secretion pathway. *Proc Natl Acad Sci U S A* 109:11342–11347. <https://doi.org/10.1073/pnas.1119453109>.
- Strong M, Sawaya MR, Wang S, Phillips M, Cascio D, Eisenberg D. 2006. Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* 103:8060–8065. <https://doi.org/10.1073/pnas.0602606103>.
- Chen X, Cheng HF, Zhou J, Chan CY, Lau KF, Tsui SK, Au SW. 2017. Structural basis of the PE-PPE protein interaction in *Mycobacterium tuberculosis*. *J Biol Chem* 292:16880–16890. <https://doi.org/10.1074/jbc.M117.802645>.
- Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, III, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor S, Whitehead S, Barrell BG. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 394: 651–653. <https://doi.org/10.1038/24206>.
- Deb C, Daniel J, Sirakova TD, Abomoelak B, Dubey VS, Kolattukudy PE. 2006. A novel lipase belonging to the hormone-sensitive lipase family induced under starvation to utilize stored triacylglycerol in *Mycobacterium tuberculosis*. *J Biol Chem* 281:3866–3875. <https://doi.org/10.1074/jbc.M505556200>.
- Santucci P, Diomandé S, Poncin I, Alibaud L, Viljoen A, Kremer L, de Chastellier C, Cnaan S. 2018. Delineating the physiological roles of the PE and catalytic domains of LipY in lipid consumption in mycobacterium-infected foamy macrophages. *Infect Immun* 86:e00394-18. <https://doi.org/10.1128/IAI.00394-18>.

23. Daleke MH, Cascioferro A, de Punder K, Ummels R, Abdallah AM, van der Wel N, Peters PJ, Luirink J, Manganelli R, Bitter W. 2011. Conserved Pro-Glu (PE) and Pro-Pro-Glu (PPE) protein domains target LipY lipases of pathogenic mycobacteria to the cell surface via the ESX-5 pathway. *J Biol Chem* 286:19024–19034. <https://doi.org/10.1074/jbc.M110.204966>.
24. Burggraaf MJ, Ates LS, Speer A, van der Kuij K, Kuijl C, Bitter W. 2019. Optimization of secretion and surface localization of heterologous OVA protein in mycobacteria by using LipY as a carrier. *Microb Cell Fact* 18:44. <https://doi.org/10.1186/s12934-019-1093-1>.
25. Barathy DV, Suguna K. 2013. Crystal structure of a putative aspartic proteinase domain of the *Mycobacterium tuberculosis* cell surface antigen PE\_PGRS16. *FEBS Open Bio* 3:256–262. <https://doi.org/10.1016/j.fob.2013.05.004>.
26. Garrett CK, Broadwell LJ, Hayne CK, Neher SB. 2015. Modulation of the activity of *Mycobacterium tuberculosis* LipY by its PE domain. *PLoS One* 10:e0135447. <https://doi.org/10.1371/journal.pone.0135447>.
27. Abdallah AM, Savage ND, van Zon M, Wilson L, Vandenbroucke-Grauls CM, van der Wel NN, Ottenhoff TH, Bitter W. 2008. The ESX-5 secretion system of *Mycobacterium marinum* modulates the macrophage response. *J Immunol* 181:7166–7175. <https://doi.org/10.4049/jimmunol.181.10.7166>.
28. Singh O, Su EC. 2016. Prediction of HIV-1 protease cleavage site using a combination of sequence, structural, and physicochemical features. *BMC Bioinformatics* 17:478. <https://doi.org/10.1186/s12859-016-1337-6>.
29. McEvoy CR, Cloete R, Muller B, Schurch AC, van Helden PD, Gagneux S, Warren RM, Gey van Pittius NC. 2012. Comparative analysis of *Mycobacterium tuberculosis* pe and ppe genes reveals high sequence variation and an apparent absence of selective constraints. *PLoS One* 7:e30593. <https://doi.org/10.1371/journal.pone.0030593>.
30. Copin R, Coscolla M, Seiffert SN, Bothamley G, Sutherland J, Mbayo G, Gagneux S, Ernst JD. 2014. Sequence diversity in the pe\_pgrs genes of *Mycobacterium tuberculosis* is independent of human T cell recognition. *mBio* 5:e00960-13. <https://doi.org/10.1128/mBio.00960-13>.
31. Puttinaowarat S, Thompson K, Lilley J, Adams A. 1999. Characterization of *Mycobacterium* spp. isolated from fish by pyrolysis mass spectrometry (PyMS) analysis. *AAHRI News* 8:4–8.
32. Ramakrishnan L, Falkow S. 1994. *Mycobacterium marinum* persists in cultured mammalian cells in a temperature-restricted fashion. *Infect Immun* 62:3222–3229.
33. Braunstein M, Bardarov SS, Jacobs WR, Jr. 2002. Genetic methods for deciphering virulence determinants of *Mycobacterium tuberculosis*. *Methods Enzymol* 359:67–99. [https://doi.org/10.1016/s0076-6879\(02\)58081-2](https://doi.org/10.1016/s0076-6879(02)58081-2).
34. Bardarov SS, Bardarov SS, Jr, Jacobs WR, Jr. 2001. Transposon mutagenesis in mycobacteria using conditionally replicating mycobacteriophages. *Methods Mol Med* 54:43–57. <https://doi.org/10.1385/1-59259-147-7-043>.
35. van Winden VJ, Ummels R, Piersma SR, Jimenez CR, Korotkov KV, Bitter W, Houben EN. 2016. Mycosins are required for the stabilization of the ESX-1 and ESX-5 type VII secretion membrane complexes. *mBio* 7:e01471-16. <https://doi.org/10.1128/mBio.01471-16>.
36. Sasseti CM, Boyd DH, Rubin EJ. 2003. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* 48:77–84. <https://doi.org/10.1046/j.1365-2958.2003.03425.x>.
37. Kallberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, Xu J. 2012. Template-based protein structure modeling using the RaptorX web server. *Nat Protoc* 7:1511–1522. <https://doi.org/10.1038/nprot.2012.085>.
38. Xu D, Zhang Y. 2011. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys J* 101:2525–2534. <https://doi.org/10.1016/j.bpj.2011.10.024>.
39. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501. <https://doi.org/10.1107/S0907444910007493>.
40. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221. <https://doi.org/10.1107/S0907444909052925>.