

PE5-PPE4-EspG₃ trimer structure from mycobacterial ESX-3 secretion system gives insight into cognate substrate recognition by ESX systems

Zachary A. Williamson¹, Catherine T. Chaton¹, William A. Ciocca^{1#}, Natalia Korotkova¹, and Konstantin V. Korotkov^{1*}

From the ¹Department of Molecular & Cellular Biochemistry and the Center for Structural Biology, University of Kentucky, Lexington, KY 40536

Running Title: *PE5-PPE4-EspG₃ structure from ESX-3*

[#]Present address: Department of Biological Sciences, Eastern Kentucky University, Richmond, KY 40475

*To whom correspondence should be addressed: Konstantin V. Korotkov: Department of Molecular & Cellular Biochemistry, University of Kentucky, Lexington, KY 40536; kkorotkov@uky.edu; Tel. (859) 323-5493; Fax (859) 257-2283.

Keywords: mycobacteria, tuberculosis, X-ray crystallography, protein complex, protein-protein interaction, type VII secretion system

ABSTRACT

Mycobacterium tuberculosis (*Mtb*) has evolved numerous type VII secretion (ESX) systems to secrete multiple factors important for both growth and virulence across their cell envelope. Three such systems; ESX-1, ESX-3, and ESX-5; have been shown to each secrete a unique set of substrates. A large class of these substrates secreted by these three systems are the PE and PPE families of proteins. Proper secretion of the PE-PPE proteins requires the presence of EspG, with each system encoding its own unique copy. There is no cross-talk between any of the ESX systems and how each EspG is recognizing its subset of PE-PPE proteins is currently unknown. The only current structural characterization of PE-PPE-EspG trimers is from the ESX-5 system. Here we present the crystal structure of the PE5_{mt}-PPE4_{mt}-EspG_{3mm} trimer, from the ESX-3 system. Our trimer reveals that EspG_{3mm} interacts exclusively with PPE4_{mt} in a similar manner to EspG₅, shielding the hydrophobic tip of PPE4_{mt} from solvent. The C-terminal helical domain of EspG_{3mm} is dynamic, alternating between an ‘open’ and ‘closed’ form, and this movement is likely functionally relevant in the unloading of PE-PPE heterodimers at the secretion machinery. In contrast to the previously solved ESX-5 trimers, the PE-PPE heterodimer of our ESX-3 trimer is interacting with its chaperone at a drastically different angle, and presents different faces of the PPE protein to the chaperone. We

conclude that the PPE-EspG interface from each ESX system has a unique shape complementarity that allows each EspG to discriminate amongst non-cognate PE-PPE pairs.

Tuberculosis is currently the deadliest infectious disease in the world, killing 1.5 million people in 2019 (1). The lack of an effective vaccine against the most prevalent pulmonary form of Tuberculosis, as well as the emergence of numerous multi-drug resistant strains of the causative agent, *Mycobacterium tuberculosis* (*Mtb*), highlights the growing need for more effective treatment options. Therefore, a more comprehensive understanding of the *Mtb* virulence machinery is needed to aid the development of new therapeutics.

Mtb, like all mycobacteria, contains a thick hydrophobic cell envelope that aids in protecting the mycobacterium from its environment. To overcome the limited permeability created by this envelope, mycobacteria have evolved a specialized secretion system to export proteins across their cell envelopes, the type VII secretion system, also known as the ESX system (2). Five different ESX systems encoded in the *Mtb* genome, and three are known to secrete proteins; ESX-1, ESX-3, and ESX-5 (3). Recently, the structures of the core complex of both ESX-3 (4,5) and ESX-5 (6) have been solved. The ESX-5 core complex has six-fold symmetry and sits on the inner membrane (6), while the ESX-3 core complex was solved as a dimer that

PE5-PPE4-EspG₃ structure from ESX-3

could be modeled onto the six-fold symmetry of the ESX-5 core complex (4,5). These systems are not functionally redundant, as their substrates are not re-routed to other ESX systems (7). The ESX systems secrete a variety of different substrates, each containing a general type VII secretion motif of YxxxD/E (8). A significant class of substrates being the PE and PPE proteins, named for conserved residues (Pro-Glu for PE and Pro-Pro-Glu for PPE) within their N-terminal domains (9,10). The N-terminal domains are about 110 (PE) or 180 (PPE) amino acids in length and interact together to form a PE-PPE heterodimer. A cytosolic chaperone, EspG, is required for proper folding and/or stability of the PE-PPE proteins and, ultimately their proper secretion (11,12). Each ESX system secretes a unique subset of PE-PPE heterodimers, and therefore each encodes an EspG that binds to only its corresponding heterodimers (11,12). The first structural insight into the EspG and PE-PPE interaction was revealed by analysis of the structure of the PE25-PPE41-EspG₅ complex, a trimer from ESX-5 (12,13). EspG₅ interacts solely with PPE41 at the tip distal to the PE25 interaction and aids in preventing PE-PPE heterodimer aggregation in part by shielding a conserved hydrophobic tip on the PPE proteins, known as the hh motif (12). The additional structure of the ESX-5-related, PE8-PPE15-EspG₅ trimer, revealed similar interactions of the substrate PE-PPE dimer with the EspG₅ chaperone (14). Despite high conservation among PPE proteins in the identified EspG₅ binding region from PPE41, three residues vary depending on whether the PPE protein is secreted by ESX-1, ESX-3, or ESX-5 (12). Alteration of any or all of these positions in the ESX-5-dependent PPE41 did not disrupt PPE41-EspG₅ binding (12). Based on this observation it has been suggested that structural elements outside of the EspG-binding region differentiate the ESX-5-specific PPE proteins from their ESX-1 and ESX-3 homologs to bind EspG₅ (12).

This study was initiated to understand the how each EspG from the different ESX systems specifically recognizes its unique subset of cognate PE-PPE heterodimers. Here we present the structure of PE5-PPE4-EspG₃ from ESX-3. This structure reveals a novel binding mode of PE-PPE proteins with the EspG chaperone and suggests the molecular mechanism by which the PE-PPE dimers are specifically targeted by cognate chaperones.

Results

EspG₃ forms a complex with PE5-PPE4 and binding is conserved across species

In order to understand the mechanism for the specificity of PE-PPE recognition by cognate chaperones a high-resolution structure of a trimer produced by the ESX systems, other than ESX-5, was needed. Our efforts have focused on optimizing the ESX-3 PE-PPE-EspG trimer for X-ray structural studies. Constructs of full-length PE5 (*Rv0285*), the conserved N-terminal PPE domain of PPE4 (*Rv0286*, residues 1-181), in a complex with the cognate full-length EspG₃ (*Rv0289*) from *Mycobacterium tuberculosis* (Fig. 1a) never formed high-resolution diffraction quality crystals, despite our best efforts. The difficulty could be due to some heterogeneity in the processing of EspG_{3mt} within the *Escherichia coli* cell, as seen by the double band in Fig. 1b and Sup. Fig. 1a. Numerous variations of PE5-PPE4-EspG₃ constructs were screened utilizing multiple mycobacterial species, different fusion approaches, and even mixing PE5-PPE4 dimers with EspG₃ chaperones from different species (Sup. Table 1). This latter approach was inspired by the work done on the *Plasmodium* aldolase-thrombospondin-related anonymous protein complex (15), and in the end, produced the best crystals for further diffraction experiments. To ensure the mixed trimers behaved the same in solution as the wild-type (WT) trimer, size-exclusion chromatography with multi-angle light scattering (SEC-MALS) experiment was performed on both the WT PE5_{mt}-PPE4_{mt}-EspG_{3mt} trimer as well as the mixed PE5_{mt}-PPE4_{mt}-EspG_{3mm} trimer which contained the *Mycobacterium marinum* EspG₃ gene (*MMAR_0548*) (Fig. 1b-c). Both trimers form a 1:1:1 complex with experimental molecular weights of 56.2 kDa (Fig. 1b) for the full *M. tuberculosis* trimer and 54.6 kDa (Fig. 1c) for the mixed trimer with the *M. marinum* EspG₃. Co-purification assays were run with both *M. tuberculosis* PE4-PPE5, along with EspG₃'s from *Mycobacterium smegmatis* (*MSMEG_0622*), *Mycobacterium hassiacum* (*MHAS_04631*), and *Mycobacterium kansasii* (*MKAN_17015*). Due to the His₆-tag only being present on PE5_{mt}, EspG₃ co-purification required interaction with the PE5_{mt}-PPE4_{mt} heterodimer. Across all species that were

PE5-PPE4-EspG₃ structure from ESX-3

tested, EspG₃ co-purified with the PE5_{mt}-PPE4_{mt} heterodimer (Supplementary Fig. 1a-e). The binding of different EspG₃s to the same PE-PPE heterodimer suggests a common protein-protein recognition mechanism within the ESX-3 family.

Overall structure of PE5_{mt}-PPE4_{mt}-EspG_{3mm}

The PE5_{mt}-PPE4_{mt}-EspG_{3mm} trimer was able to form diffraction quality crystals, and two different crystal forms were observed that diffracted to 3.3 Å (I422) and 3.0 Å (P2₁2₁2₁) (Table 1). Final refinement and data statistics are in Table 1. Overall there is little structural variation between the individual proteins across the copies present in the two crystal forms (Table 2).

For all structural analysis and comparisons, the first copy of the PE5_{mt}-PPE4_{mt}-EspG_{3mm} trimer from the higher resolution P2₁2₁2₁ crystal form was used because it diffracted at a higher resolution and has the lowest *B*-factors from the non-crystallographic copies in the P2₁2₁2₁ form. EspG_{3mm} interacts solely with the tip of PPE4_{mt} (Fig. 2), similar to EspG₅ in the previously solved ESX-5 trimers (12-14). However, the orientation of PE5_{mt}-PPE4_{mt} relative to EspG_{3mm} is dramatically different than was observed for either ESX-5 trimer, and the differences between them will be described in later sections. The YXXD/E motif for ESX secretion of PE5_{mt} is accessible for interactions with the rest of the ESX machinery, as it is located distal to the EspG_{3mm} interaction (11). In both crystal forms, this secretion motif is disordered, similar to the motif in PE8_{mt} from the PE8_{mt}-PPE15_{mt}-EspG_{5mt} trimer (14). The individual components of the PE5_{mt}-PPE4_{mt}-EspG_{3mm} trimer align well to the individual components of the previously reported ESX-5 trimers, both PE25_{mt}-PPE41_{mt}-EspG_{5mt} (4KXR and 4W4L) and PE8_{mt}-PPE15_{mt}-EspG_{5mt} (5XFS), with only moderate variations (Table 3).

In a previous study on EspG structures (16), a small-angle X-ray scattering (SAXS) experiment was done on the PE5-PPE4-EspG₃ trimer from *M. smegmatis*. Comparisons between this SAXS analysis and our crystal structure were performed to see if the solution-based characterization of the trimer matched the X-ray-based characterization. We ran CRYSOLOG (17) on our crystal structure in comparison to the experimental scattering data from the *M. smegmatis* trimer. The overall χ^2 is 2.53, which is acceptable

given that the trimers are from different species with only 54.0-73.8% sequence identity across the different components (Sup. Fig. 2). The main differences are in the extreme high- and low-resolution areas, likely arising from differences in the primary structure between the two samples and from aggregation in the SAXS sample, respectively. Therefore, we are confident that the crystal structure is an appropriate model of the ESX-3 trimer as it exists in solution.

Interface between PPE4_{mt} and EspG_{3mm}

The interface between EspG_{3mm} and PPE4_{mt} contains numerous hydrophobic interactions, multiple hydrogen bonds, and two salt bridges centered around E140^{PPE4_{mt}} (Fig. 2b-f). Overall the interface buries 3,121 Å² of solvent-accessible surface area, as calculated by the PISA server (18) and is comprised of 30 total residues from PPE4_{mt} and 49 residues from EspG_{3mm} (Fig. 3). The tip of PPE4_{mt} containing the ends of α 4 and α 5 and the loop between them is inserted into a groove on EspG₃ composed of its central β sheet and C-terminal helical bundle. This bundle shields the hydrophobic tip of PPE4_{mt}, including the hh motif of FF128, from solvent access. The tip of PPE4_{mt} is interacting with EspG₃ in such a way that the complex is unlikely to disengage at the ESX secretion machinery without structural rearrangement of the chaperone.

Mutations cause disruptions in the PPE4-EspG₃ interface

In order to probe the interface of the crystal structure and test the importance of interacting residues, we made several mutations on both PPE4 and EspG₃ sides of the interface and opted to use the cognate PE5_{mt}-PPE4_{mt}-EspG_{3mt} trimer to test our mutations. The PISA output (18) of the interface was analyzed along with sequence alignments of the current known ESX-3 PPE proteins (Sup. Fig. 3) and alignments of the EspG₃ used in this study (Sup. Fig. 4) to select which residues in the interface would be mutated.

PPE4_{mt} is well conserved along the interface among ESX-3-specific PPE proteins (Sup. Fig. 3), and we targeted strictly conserved residues in the interface. We selected N127 and N132 because they contain buried hydrogen bonds, F128 and F129 because they are the hh motif and contribute a large amount of solvation energy to the

PE5-PPE4-EspG₃ structure from ESX-3

interface according to PISA (18), and E140 because it is part of the salt bridges in the interface. We ran co-purification pull-down assays with mutated PPE4_{mt} and EspG_{3mt} (Table 4). As described earlier, EspG_{3mt} is only co-purified with the PE5-PPE4 heterodimer if it forms a complex. The introduction of charges into the buried hydrogen bonds with N127D and N132E was unable to break the PPE4_{mt}-EspG_{3mt} interaction, and neither was the charge reversal of E140R, as all three mutations co-purify with EspG_{3mt} (Sup. Fig. 5a). This suggests that disruption of any of these single positions is not sufficient to abolish PPE4_{mt}-EspG_{3mt} interaction. Conversely, the introduction of charged residues into the hh motif with F128R or F129E did disrupt the interface and prevented EspG_{3mt} from being co-purified (Sup. Fig. 5a), as it interrupts with the hydrophobic environment deep within the EspG_{3mt} binding pocket.

The interface of EspG_{3mt} is also well conserved amongst the various EspG₃s tested in this study (Sup. Fig. 4), and again, we targeted strictly conserved residues. We selected R208 and E212 because they contain buried hydrogen bonds, R87 and R102 because they form the salt bridge within the interface, and S231 because it sits at the top of the groove of EspG₃ and could sterically block entrance into the pocket. Neither single mutation of the salt bridge, R87E or R102E, was able to prevent co-purification of EspG_{3mt} (Sup. Fig. 5b). Also, the introduction of a charged residue with R208E was unable to prevent the interaction (Sup. Fig. 5b). In contrast, E212R was sufficient to prevent co-purification, as well as S231Y (Sup. Fig. 5b), as both prevent the hydrophobic tip of PPE4_{mt} from interacting with the binding pocket of EspG_{3mt} either by charge repulsion or steric hindrance. Thus, our mutations on both PPE4_{mt} and EspG_{3mt} highlight the importance of the hydrophobic environment deep within the PPE4_{mt}-EspG_{3mt} interface.

Structure of EspG₃ in and out of trimer complex

Our structure is the first of EspG₃ solved in complex with a cognate PE-PPE dimer, and thus we wanted to compare it to the previously solved unbound EspG₃ structures. In total, there are six available EspG₃ structures, four of EspG_{3ms} (PDB codes: 4L4W, 4RCL, 5SXL, and 4W4J (13,16)), one EspG_{3mt} (4W4I (13)), and one EspG_{3mm} (5DLB (16)). These six structures can be classified into two

different forms, an ‘open’ and a ‘closed’ form. The differentiation between these two forms is the orientation of the C-terminal helical bundle relative to the core β -sheet. The EspG_{3mm} structure (5DLB) is representative of the ‘open’ form, and one of the EspG_{3ms} structures (4RCL) is representative of the ‘closed’ form. Analysis of EspG_{3mm} as it exists in the PE5_{mt}-PPE4_{mt}-EspG_{3mm} trimer was done relative to these two representative structures. The overall alignment of the representative structures to the bound EspG_{3mm} was good with r.m.s.d. of 2.1 Å and 1.9 Å for the ‘open’ and ‘closed’ forms, respectively (Fig. 4a). Inspection of these alignments show the majority of differences to be within the arrangement of the C-terminal helical bundles, with the bound form of EspG_{3mm} being in close to the orientation found in the ‘closed’ form (Fig. 4b-c). The bound EspG_{3mm} cannot be any closer to the ‘closed’ form orientation because the C-terminal helical bundle is making contacts with PPE4_{mt}. We hypothesized that this C-terminal helical bundle is dynamic and closes on cognate PPE proteins upon interaction. A comparison between the bound EspG_{3mm} structure and the ‘open’ EspG_{3mm} was performed with the DynDom server to test this hypothesis (19). DynDom identified a moving domain within the structures that was located in the C-terminal helical bundle (Fig. 4d). DynDom’s analysis also performed a whole structure alignment that agreed with the previous Dali alignment in Fig 4a-b. DynDom performed alignments between the fixed domains (residues 11-168 and 189-279) and the moving domains (residues 168-188), which resulted in much better alignments with r.m.s.d. of 1.76 Å and 0.86 Å, respectively. Therefore, the moving domain, the C-terminal helical bundle, is essentially structurally identical between PPE4_{mt}-bound EspG_{3mm} and the ‘open’ EspG_{3mm} and its rotation of 30.2° and translation of 0.8 Å is moderately perturbing the fixed domain. Since the moving domain is making extensive contact with PPE4_{mt} and PPE4_{mt} would sterically clash with the current orientation of the C-terminal helical bundle, the movement from the ‘closed’ to the ‘open’ orientation could be significant in releasing the secreted PE-PPE dimers from the chaperone at the secretion machinery.

Comparison of ESX-3 and ESX-5 PE-PPE-EspG trimers

PE5-PPE4-EspG₃ structure from ESX-3

A vastly different binding mode is observed when comparing the ESX-3-specific PE_{5mt}-PPE_{4mt}-EspG_{3mm} trimer to the previously published ESX-5-specific trimers. As mentioned earlier, there is good agreement when comparing individual components of the ESX-3-specific trimer to the available ESX-5-specific trimers (Table 3). The difference between the two sets of trimers became apparent when they were aligned via EspG (Fig. 5a-b). Our results focused on comparisons with the PE_{25mt}-PPE_{41mt}-EspG_{5mt} (4KXR) trimer, but the same differences were present with the PE_{8mt}-PPE_{15mt}-EspG_{5mt} (5XFS) trimer. The interaction angle of the different PE-PPE heterodimer with EspG is drastically different between the two trimers, with 30° angle difference (Fig. 5b). Another difference lies within the hh motif loops of PPE_{25mt} (α 4- α 5 loop) and PPE_{4mt} (α 5- α 6 loop) (Fig. 5c). In PPE_{25mt}, this loop is seven residues long and undertakes a compact conformation that is not altered during EspG_{5mt} binding (12). In contrast, in PPE_{4mt}, this loop is nine residues long and has an extended conformation, and this difference was rapidly apparent when PPE_{25mt} and PPE_{4mt} were aligned (Fig. 5c).

This loop conformation also made each PPE protein incompatible with the other's binding mode. When looking at the PPE alignment in the context of the ESX-3 trimer, the α 4- α 5 loop of PPE_{25mt} does not align over the central groove of EspG_{3mm} and instead sterically clashes the central β sheet of the chaperone (Fig. 5d). The tip of PPE_{41mt} would have to undergo a drastically new tip confirmation in order to bind in the opening of EspG_{3mm}. In the context of the ESX-5 trimer, the α 5- α 6 loop does not align with the central groove of the chaperone, and instead, PPE_{4mt}'s hh motif sterically clashes with the C-terminal helical bundle of EspG_{5mt} (Fig. 5e). Also, none of the salt bridges between PPE_{41mt} and EspG_{5mt} are conserved in PPE_{4mt}. Specifically, D134^{PPE41mt}-K235^{EspG5mt}, D140^{PPE41mt}-R109^{EspG5mt}, and D144^{PPE41mt}-R27^{EspG5mt}, that are all replaced with hydrophobic residues in PPE_{4mt}: either T137^{PPE4mt} or L138^{PPE4mt}, V144^{PPE4mt}, and L147^{PPE4mt}, respectively.

Discussion

In this work, we present the first structure of the PE_{5mt}-PPE_{4mt}-EspG_{3mm} trimer, which is from the ESX-3 system. Our structure is a mixed trimer,

and we presented evidence that EspG₃ from numerous mycobacterial species can bind the PE_{5mt}-PPE_{4mt} heterodimer. Conservation of the EspG₃s used in this study ranged from 57-83% identity, yet an enrichment in conservation is observed within PPE4-interacting residues (Sup. Fig. 4). The ability of EspG₃ from numerous mycobacterial species to bind PE_{5mt}-PPE_{4mt} suggests that the recognition mechanism is conserved within ESX systems across species. Overall the PE_{5mt}-PPE_{4mt} interaction is similar to the previously reported PE-PPE-EspG trimers (12-14) in that PPE_{4mt}'s tip is solely interacting with EspG_{3mm} and the general secretion motif of YxxxD/E, on PE_{5mt}, is at the distal end of the PE_{5mt}-PPE_{4mt} heterodimer. In all copies of PE_{5mt}, this motif is unstructured as it is in the PE₈-PPE₁₅-EspG₅ trimer (14), and similarly, W63^{PPE4mt} is pointed away from this secretion motif. This arrangement is distinct from the PE₂₅-PPE₄₁-EspG₅ trimers (12,13) and EspB, an ESX-1 substrate that has a similar structural fold to the PE-PPE heterodimers (20,21). PE_{8mt} contains an expanded C-terminal domain, and since the secretion motif is located in the linker between the C-terminal domain and the PE domain the orientation of the secretion motif was unclear (14). PE_{5mt} does not have an expanded C-terminal domain and is just the conserved PE domain, yet its secretion motif is still unstructured in our trimer. Therefore, it is still unclear the exact significance of the structural variations in the ESX secretion motif, and further work is still needed.

Our structure is the first of EspG₃ bound to a cognate PE-PPE heterodimer. In comparisons of the various published EspG₃ structures, we identified two different forms that relate to the orientation of the C-terminal helical bundle, an 'open' and a 'closed' form. EspG_{3mm}, when bound to the PE_{5mt}-PPE_{4mt} heterodimer, is in a conformation slightly different than the 'closed' form due to interactions with the tip of PPE_{4mt}. We also found that the C-terminal helical bundle is a dynamic domain and shifts between the 'open' and 'closed' forms via a hinge movement (Fig. 4d). The functional significance of this domain movement could be two-fold. First, the plasticity of the C-terminal helical bundle could allow EspG₃ to accommodate any variation in the ESX-3-specific PPE tips. While the tip of ESX-3-specific PPE proteins is mostly conserved (Sup. Fig. 3), there is

PE5-PPE4-EspG₃ structure from ESX-3

some variations at the end of $\alpha 5$ that could alter the tertiary structure and thus slightly alter the interactions with the EspG₃ chaperone and the PPE protein. Secondly, the movement of the C-terminal helical bundle could be critical to the release of the PE-PPE heterodimers at the ESX-3 secretion machinery. It is unlikely that PPE4_{mt} could be removed from its interactions with EspG_{3mm} without either movement of the C-terminal helical bundle or steric clashes with the C-terminal helical bundle. Movement of this helical bundle and release of PPE4_{mt} would likely require energy input, and a candidate to provide that energy is EccA. EccA is an ATPase (2) and interacts with both EspG and PPE proteins in yeast two-hybrid experiments (13,22). Recent structures of the ESX machinery from both ESX-3 (4) and ESX-5 (6) suggests overall six-fold symmetry of the core ESX machinery within the inner membrane, and EccA could not only be acting to provide the energy required to uncouple the PE-PPE heterodimers from their EspG chaperone but also to provide a platform for interaction with the core secretion machinery as EccA is likely hexameric when functional.

Previous studies showed that each EspG only recognize PE-PPE heterodimers from their cognate systems (11,12). Despite the structures of two different PE-PPE-EspG trimers from ESX-5 (12-14), it was still unclear how EspG₅ was differentiating from cognate and non-cognate PE-PPE heterodimers. Our structure represents the first PE-PPE-EspG trimer from ESX-3 and allows for direct comparisons between the ESX-3 and ESX-5 trimers. Our structure reveals that PE5_{mt}-PPE4_{mt} interacts with EspG_{3mm} at a different angle of interaction than what was shown for either ESX-5 trimer. This difference in interaction angle presents a different face of PPE4_{mt} to EspG_{3mm}. We hypothesize that this is a conserved feature of the ESX-3 PPE-EspG₃ interaction, as both characterized ESX-5 PE-PPE heterodimers (12-14) display the same face to EspG₅ despite 33% sequence identity between PPE41 and PPE15. Therefore, we hypothesize that each ESX system has a unique shape complementarity between its subset of PPE proteins and their cognate EspG chaperone, and these unique shapes are likely not compatible for interaction with non-cognate chaperones. Our structure is also the first of an ESX-3-specific PE-PPE heterodimer. PE5_{mt}-

PPE4_{mt} shares the same global conformation as the previously solved PE-PPE heterodimers, yet it differs drastically in PPE4_{mt} in the loop between $\alpha 5$ - $\alpha 6$, which contains the hh motif. This longer, more extended loop interacts deeper in the cleft of EspG_{3mm} and is subsequently much more shielded from solvent. It is possible that the longer, extended loop conformation is a feature of ESX-3 PPE proteins and could play an essential role in EspG₃ recognition.

In conclusion, we presented the first structure of a PE-PPE-EspG trimer from the ESX-3 system. This structure allowed us to compare the interactions of EspG₃ and a cognate PPE protein to the previously described EspG₅-PPE interactions. We hypothesize that shape complementarity is a key feature of distinguishing cognate and non-cognate PPE proteins from the EspG chaperones.

Experimental Procedures

Bacterial strains and growth conditions

The *Escherichia coli* Rosetta2(DE3) strains grown in Luria-Bertani (LB) medium or on LB agar at 37 °C. When needed, antibiotics were included at the following concentrations: chloramphenicol at 10 μ g/ml, streptomycin at 50 μ g/ml, and kanamycin at 50 μ g/ml.

Expression and purification of PE5-PPE4-EspG₃ heterotrimers

Optimized DNA sequences based on the amino acids of full-length PE5 and PPE4 residues 1-180 from *M. tuberculosis* were obtained from Invitrogen and put into a pRSF-NT vector (23) using NcoI and HindIII restriction sites, which contains an N-terminal His₆ tag on PE5 that is cleavable by TEV protease. EspG_{3mt}. EspG_{3mm} expression plasmid was constructed as described previously (12). Mutations in PPE4_{mt}, EspG_{3mt}, and EspG_{3mm} were introduced with Gibson assembly mutagenesis (SGI-DNA).

Co-expression of all heterotrimers was performed as described previously (12). Briefly, *E. coli* strains containing the appropriate PE5_{mt}-PPE4_{mt} and EspG₃ plasmids were induced with 0.5 mM IPTG when they reached an OD at 600 nm of 0.5-0.8 and then continued to shake at 16 °C for 20 h. Cells were harvested by centrifugation. Cells were then resuspended in lysis buffer (300 mM NaCl, 20 mM Tris pH 8.0, and 10 mM imidazole)

PE5-PPE4-EspG₃ structure from ESX-3

and 1:100 Halt protease inhibitor cocktail (Thermo Fisher Scientific, Waltham, MA). Cells were lysed using an EmulsiFlex-C5 homogenizer (Avestin, Ottawa, ON, Canada). The soluble lysate was purified over a Ni-NTA column (G-Biosciences, St. Louis, MO). Eluted protein was dialyzed against lysis buffer without imidazole and incubated with 1:20 mg of TEV protease at 4 °C for 20 h before being re-applied to the Ni-NTA column. Flow-through and washes were pooled and concentrated for size-exclusion chromatography over a Superdex 200 Increase 10/300 GL column (GE Healthcare Life Sciences, Marlborough, MA) that was equilibrated in buffer A (100 mM NaCl and 20 mM HEPES pH 7.5).

Crystallization, data collection, and structure solution

Purified protein was concentrated to 4.195 mg/ml. Initial screening was done using the MCSG Crystallization Suite (Anatrace, Maumee, OH). This initial screening produced the P₂₁₂₁ crystals that were grown in 200 mM NH₄ tartrate and 20% PEG 3350. Optimization around three others hits from the initial crystal screening containing NaCl as the precipitant and various buffers ranging from a pH 5.5 to 8.0 produced the I422 crystals, which were grown in 2.0 M NaCl and 100 mM bis-tris, pH 6.5. Crystals were transferred to cryoprotectant solution, which contained the crystallization solution supplemented with either 20% (P₂₁₂₁) or 25% (I422) glycerol and then flash-cooled in liquid N₂. Data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Data were processed using XDS and XSCALE (24). Molecular replacement using Phaser (25) was used to solve the structure of both crystal forms. First, the PE₂₅_{mt}-PPE₄₁_{mt} dimer (PDB: 4KXR (12)) and EspG_{3mm} (PDB: 5DLB (16)) were used as search models for the I422 dataset. Later, an early model of the I422 structure was used as a search model for the P₂₁₂₁ dataset. DENSITY MODIFICATION AND ORIGINAL MODEL. The starting model for both forms was then iteratively rebuilt and refined using Coot and phenix.refine (26,27). The final structure for both crystal forms was refined in phenix.refine, with the P₂₁₂₁ form using non-crystallographic symmetry restraints (TLS?). All data collection and refinement statistics are listed in Table 1. The final

model was assessed using Coot and the MolProbity server (28) for quality.

Size-exclusion chromatography multi-angle light scattering (SEC-MALS)

Proteins were expressed and purified as described and then passed over an AKTA pure with an inline Superdex 200 Increase 10/300 GL column (GE Healthcare Life Sciences), miniDAWN TREOS, and Optilab T-rEX (Wyatt Technologies, Santa Barbara, CA). The system was equilibrated and run in buffer A. Samples were loaded at a volume of 500 µL at a concentration of 2-4 mg/ml and the system was run at 0.5 mL/min. Analysis of light scattering data was done using Astra (Wyatt Technologies). Molecular weight determination was done by analyzing peaks at one-half their maximum. Graphics were prepared using Prism (Graphpad Software, La Jolla, CA).

Sequence analysis

Sequence analysis was performed using the EMBL-EBL analysis tools, specifically the Clustal Omega program (29). Rendering of sequence analysis was done with the ESPript server (30).

Structural analysis

All structural figures were generated using PyMol (<http://www.pymol.org>) and Chimera/ChimeraX (chimera ref). Electrostatic surface potentials were calculated using the APBS Electrostatics plugin in PyMol (31). Structural alignments were performed with the Dali server (32).

SAXS Data Comparison and Ab Initio Model Reconstruction

PE_{5ms}-PPE_{4ms}-EspG_{3ms} trimer SAXS data (SASDDX2 (16)) was compared to a single copy of the mixed PE_{5mt}-PPE_{4mt}-EspG_{3mm} trimer structure (PDB code: 6UUJ) using CRY SOL (17). Ab initio reconstruction of the envelope was completed using GASBOR (33). Monomeric symmetry was used as a constraint for GASBOR. Twenty ab initio models were generated and averaged using the DAMAVER software package (34). DAMSEL rejected only one model.

Accession codes

PE5-PPE4-EspG₃ structure from ESX-3

Coordinates and structure factors were deposited in the Protein Data Bank with accession codes 6UUJ (P2₁2₁) and 6VHR (I422).

Acknowledgments

The work in this study was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health grants P20GM103486 and P30GM110787, and by the National Institute of Allergy and Infectious Diseases grant R01AI119022 to K.V.K. W.A.C. was supported by the National Science Foundation Research Experiences for Undergraduates (REU) grant 1358627. Use of SER-CAT is supported by its member institutions (see www.ser-cat.org/members.html), and equipment grants (S10_RR25528 and S10_RR028976) from the National Institutes of Health. Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, contract W-31-109-Eng-38. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH.

Conflicts of Interest

The authors declare that they have no conflicts of interest with the contents of this article.

References

1. Organization, W. H. (2019) Tuberculosis Fact Sheet 2019. World Health Organization, Geneva, Switzerland
2. Bitter, W., Houben, E. N., Bottai, D., Brodin, P., Brown, E. J., Cox, J. S., Derbyshire, K., Fortune, S. M., Gao, L. Y., Liu, J., Gey van Pittius, N. C., Pym, A. S., Rubin, E. J., Sherman, D. R., Cole, S. T., and Brosch, R. (2009) Systematic genetic nomenclature for type VII secretion systems. *PLoS Pathog* **5**, e1000507
3. Groschel, M. I., Sayes, F., Simeone, R., Majlessi, L., and Brosch, R. (2016) ESX secretion systems: mycobacterial evolution to counter host immunity. *Nat Rev Microbiol* **14**, 677-691
4. Famelis, N., Rivera-Calzada, A., Degliesposti, G., Wingender, M.,

- Mietrach, N., Skehel, J. M., Fernandez-Leiro, R., Bottcher, B., Schlosser, A., Llorca, O., and Geibel, S. (2019) Architecture of the mycobacterial type VII secretion system. *Nature* **576**, 321-325
5. Poweleit, N., Czudnochowski, N., Nakagawa, R., Trinidad, D., Murphy, K. C., Sasseti, C. M., and Rosenberg, O. S. (2019) The structure of the endogenous ESX-3 secretion system. *Elife* **8**
6. Beckham, K. S., Ciccarelli, L., Bunduc, C. M., Mertens, H. D., Ummels, R., Lugmayr, W., Mayr, J., Rettel, M., Savitski, M. M., Svergun, D. I., Bitter, W., Wilmanns, M., Marlovits, T. C., Parret, A. H., and Houben, E. N. (2017) Structure of the mycobacterial ESX-5 type VII secretion system membrane complex by single-particle analysis. *Nat Microbiol* **2**, 17047
7. Abdallah, A. M., Gey van Pittius, N. C., Champion, P. A., Cox, J., Luirink, J., Vandenbroucke-Grauls, C. M., Appelmelk, B. J., and Bitter, W. (2007) Type VII secretion--mycobacteria show the way. *Nat Rev Microbiol* **5**, 883-891
8. Daleke, M. H., Ummels, R., Bawono, P., Heringa, J., Vandenbroucke-Grauls, C. M., Luirink, J., and Bitter, W. (2012) General secretion signal for the mycobacterial type VII secretion pathway. *Proc Natl Acad Sci U S A* **109**, 11342-11347
9. Ates, L. S. (2019) New insights into the mycobacterial PE and PPE proteins provide a framework for future research. *Mol Microbiol*
10. Delogu, G., Brennan, M. J., and Manganelli, R. (2017) PE and PPE Genes: A Tale of Conservation and Diversity. *Adv Exp Med Biol* **1019**, 191-207
11. Daleke, M. H., van der Woude, A. D., Parret, A. H., Ummels, R., de Groot, A. M., Watson, D., Piersma, S. R., Jimenez, C. R., Luirink, J., Bitter, W., and Houben, E. N. (2012) Specific chaperones for the type VII protein secretion pathway. *J Biol Chem* **287**, 31939-31947
12. Korotkova, N., Freire, D., Phan, T. H., Ummels, R., Creekmore, C. C., Evans, T. J., Wilmanns, M., Bitter, W., Parret, A.

PE5-PPE4-EspG₃ structure from ESX-3

13. H., Houben, E. N., and Korotkov, K. V. (2014) Structure of the Mycobacterium tuberculosis type VII secretion system chaperone EspG5 in complex with PE25-PPE41 dimer. *Mol Microbiol* **94**, 367-382
14. Ekiert, D. C., and Cox, J. S. (2014) Structure of a PE-PPE-EspG complex from Mycobacterium tuberculosis reveals molecular specificity of ESX protein secretion. *Proc Natl Acad Sci U S A* **111**, 14758-14763
15. Chen, X., Cheng, H. F., Zhou, J., Chan, C. Y., Lau, K. F., Tsui, S. K., and Au, S. W. (2017) Structural basis of the PE-PPE protein interaction in Mycobacterium tuberculosis. *J Biol Chem* **292**, 16880-16890
16. Bosch, J., Buscaglia, C. A., Krumm, B., Ingason, B. P., Lucas, R., Roach, C., Cardozo, T., Nussenzweig, V., and Hol, W. G. (2007) Aldolase provides an unusual binding site for thrombospondin-related anonymous protein in the invasion machinery of the malaria parasite. *Proc Natl Acad Sci U S A* **104**, 7015-7020
17. Tuukkanen, A. T., Freire, D., Chan, S., Arbing, M. A., Reed, R. W., Evans, T. J., Zenkeviciute, G., Kim, J., Kahng, S., Sawaya, M. R., Chaton, C. T., Wilmanns, M., Eisenberg, D., Parret, A. H. A., and Korotkov, K. V. (2019) Structural Variability of EspG Chaperones from Mycobacterial ESX-1, ESX-3, and ESX-5 Type VII Secretion Systems. *J Mol Biol* **431**, 289-307
18. Svergun, D., Barberato, C., and Koch, M. H. J. (1995) CRY SOL— a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography* **28**, 768-773
19. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797
20. Hayward, S., and Berendsen, H. J. (1998) Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins* **30**, 144-154
21. Korotkova, N., Piton, J., Wagner, J. M., Boy-Rottger, S., Japaridze, A., Evans, T. J., Cole, S. T., Pojer, F., and Korotkov, K. V. (2015) Structure of EspB, a secreted substrate of the ESX-1 secretion system of Mycobacterium tuberculosis. *J Struct Biol* **191**, 236-244
22. Solomonson, M., Setiাপutra, D., Makepeace, K. A. T., Lameignere, E., Petrotchenko, E. V., Conrady, D. G., Bergeron, J. R., Vuckovic, M., DiMaio, F., Borchers, C. H., Yip, C. K., and Strynadka, N. C. J. (2015) Structure of EspB from the ESX-1 type VII secretion system and insights into its export mechanism. *Structure* **23**, 571-583
23. Teutschbein, J., Schumann, G., Mollmann, U., Grabley, S., Cole, S. T., and Munder, T. (2009) A protein linkage map of the ESAT-6 secretion system 1 (ESX-1) of Mycobacterium tuberculosis. *Microbiol Res* **164**, 253-259
24. Korotkov, K. V., Delarosa, J. R., and Hol, W. G. J. (2013) A dodecameric ring-like structure of the N0 domain of the type II secretin from enterotoxigenic Escherichia coli. *J Struct Biol* **183**, 354-362
25. Kabsch, W. (2010) Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132
26. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674
27. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132
28. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221
29. Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L.

PE5-PPE4-EspG₃ structure from ESX-3

- N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., 3rd, Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S., and Richardson, D. C. (2018) MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293-315
29. Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D., and Lopez, R. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* **47**, W636-W641
30. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-324
31. Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H., Wilson, L., Chen, J., Liles, K., Chun, M., Li, P., Gohara, D. W., Dolinsky, T., Konecny, R., Koes, D. R., Nielsen, J. E., Head-Gordon, T., Geng, W., Krasny, R., Wei, G. W., Holst, M. J., McCammon, J. A., and Baker, N. A. (2018) Improvements to the APBS biomolecular solvation software suite. *Protein Sci* **27**, 112-128
32. Holm, L. (2019) Benchmarking Fold Detection by DaliLite v.5. *Bioinformatics*
33. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* **80**, 2946-2953
34. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of fab initio shape determination in small-angle scattering. *Journal of Applied Crystallography* **36**, 860-864

PE5-PPE4-EspG₃ structure from ESX-3

Table 1. Data collection and refinement statistics.

	PE5_{mt}-PPE4_{mt}-EspG_{3mm} (6UUJ)	PE5_{mt}-PPE4_{mt}-EspG_{3mm} (6VHR)
Data Collection		
Wavelength (Å)	1.000	1.000
Space group	P2 ₁ 2 ₁ 2 ₁	I422
Cell Dimensions:		
<i>a, b, c</i> (Å)	72.26, 158.63, 209.31	219.14, 219.14, 104.44
<i>α, β, γ</i> (°)	90, 90, 90	90, 90, 90
Resolution (Å)	39.51 – 3.00 (3.08 – 3.00) ^a	35.73 – 3.30 (3.39 – 3.30)
R _{sym}	0.131 (1.56)	0.087 (2.18)
CC _{1/2} ^b	0.998 (0.590)	0.999 (0.597)
<i>I</i> /σ	9.45 (1.14)	15.70 (1.31)
Completeness (%)	99.1 (99.5)	99.8 (100)
Multiplicity	4.2 (4.4)	10.3 (9.5)
Refinement		
Resolution (Å)	39.51 – 3.00	35.73 – 3.30
No. reflections (total/free)	48463/2462	19335/928
R _{work} /R _{free}	0.266/0.303	0.248/0.266
Number of atoms:		
Protein	14535	3643
Ligand/ion	0	0
Water	4	0
<i>B</i> -factors:		
Protein	101.6	173.2
Water	70.6	
All atoms	101.6	173.2
Wilson <i>B</i>	87.9	147.3
R.m.s. deviations:		
Bond lengths (Å)	0.002	0.002
Bond angles (°)	0.53	0.503
Ramachandran distribution ^c (%)		
Favored	96.42	91.34
Allowed	3.58	8.04
Outliers	0	0.62
Rotamer outliers ^c (%)	0.28	0
Clashscore ^d	7.02	6.94
MolProbity Score ^e	1.62	1.89

^aValues in parentheses are for the highest resolution shell.

^bCC_{1/2} correlation coefficient is defined in (Karplus and Diederichs) and was calculated with *XSCALE* (Kabsch).

^cCalculated with the MolProbity server (<http://molprobity.biochem.duke.edu>) (Chen or Williams).

^dClashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.

^eMolProbity Score combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution (Chen or Williams)

PE5-PPE4-EspG₃ structure from ESX-3

Table 2. Structural Variations in copies of PE5_{mt}-PPE4_{mt}-EspG_{3mm} structure in r.m.s.d. (Å).

	PE5_{mt}	PPE4_{mt}	EspG_{3mm}
Aligned to 6UUJ copy 1			
6UUJ copy 2	0.2	0.3	0.4
6UUJ copy 3	0.4	0.2	0.4
6UUJ copy 4	0.3	0.2	0.4
6VHR	0.4	0.5	0.6
Aligned to 6UUJ copy 2			
6UUJ copy 1	0.2	0.3	0.4
6UUJ copy 3	0.3	0.3	0.3
6UUJ copy 4	0.3	0.3	0.4
6VHR	0.4	0.5	0.6
Aligned to 6UUJ copy 3			
6UUJ copy 1	0.4	0.2	0.4
6UUJ copy 2	0.3	0.3	0.3
6UUJ copy 4	0.4	0.3	0.4
6VHR	0.5	0.5	0.6
Aligned to 6UUJ copy 4			
6UUJ copy 1	0.3	0.2	0.4
6UUJ copy 2	0.3	0.3	0.4
6UUJ copy 3	0.3	0.3	0.4
6VHR	0.4	0.5	0.7

PE5-PPE4-EspG₃ structure from ESX-3

Table 3. Structural variations between individual components of ESX-3 trimer and the previously published ESX-5 trimers in r.m.s.d. (Å)

	PE25 _{mt} -PPE41 _{mt} -EspG _{5mt} (4KXR)	PE25 _{mt} -PPE41 _{mt} -EspG _{5mt} (4W4L)	PE8 _{mt} -PPE15 _{mt} -EspG _{5mt} (5XFS)
PE5 _{mt}	2.3	2.3	1.4
PPE4 _{mt}	3.3	3.4	2.5
EspG _{3mm}	2.4	2.7	2.3

PE5-PPE4-EspG₃ structure from ESX-3

Table 4. Summary of analysis of PE5_{mt}-PPE4_{mt}-EspG_{3mt} interactions *in vitro*.

PPE4 mutations	
N127D	+
F128R	-
F129E	-
N132E	+
E140R	+
EspG₃ mutations	
R87E	+
R102E	+
R208E	+
E212R	-
S231Y	-

PE5-PPE4-EspG₃ structure from ESX-3

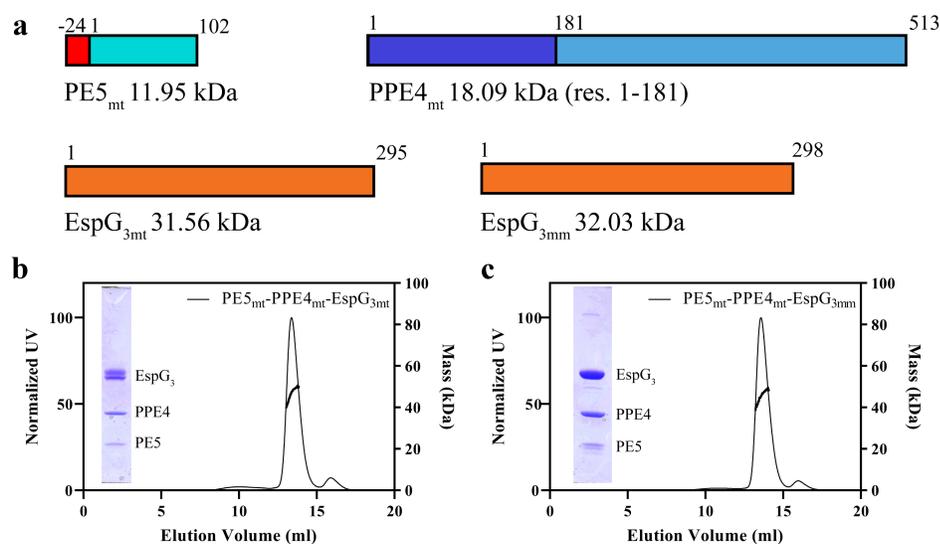


Figure 1. Solution characterization of the PE5-PPE4-EspG₃ trimer. *a*, schematic showing design and molecular weights for constructs used in this study. PE5 from *M. tuberculosis* (*Rv0285*) contains an N-terminal His₆-tag that is connected to the gene via a TEV protease cleavable linker. PPE4 (*Rv0286*) was truncated after its N-terminal PPE domain. Full-length copies of both *M. tuberculosis* (*Rv0289*) and *M. marinum* (*MMAR_0548*) EspG₃ were used. *b* and *c*, elution profile of PE5_{mt}-PPE4_{mt}-EspG_{3mt} (*b*) and PE5_{mt}-PPE4_{mt}-EspG_{3mmm} (*c*), with the right Y-axis showing the MALS-measured molecular weight. Inset shows an SDS-PAGE image of the peak fraction.

PE5-PPE4-EspG₃ structure from ESX-3

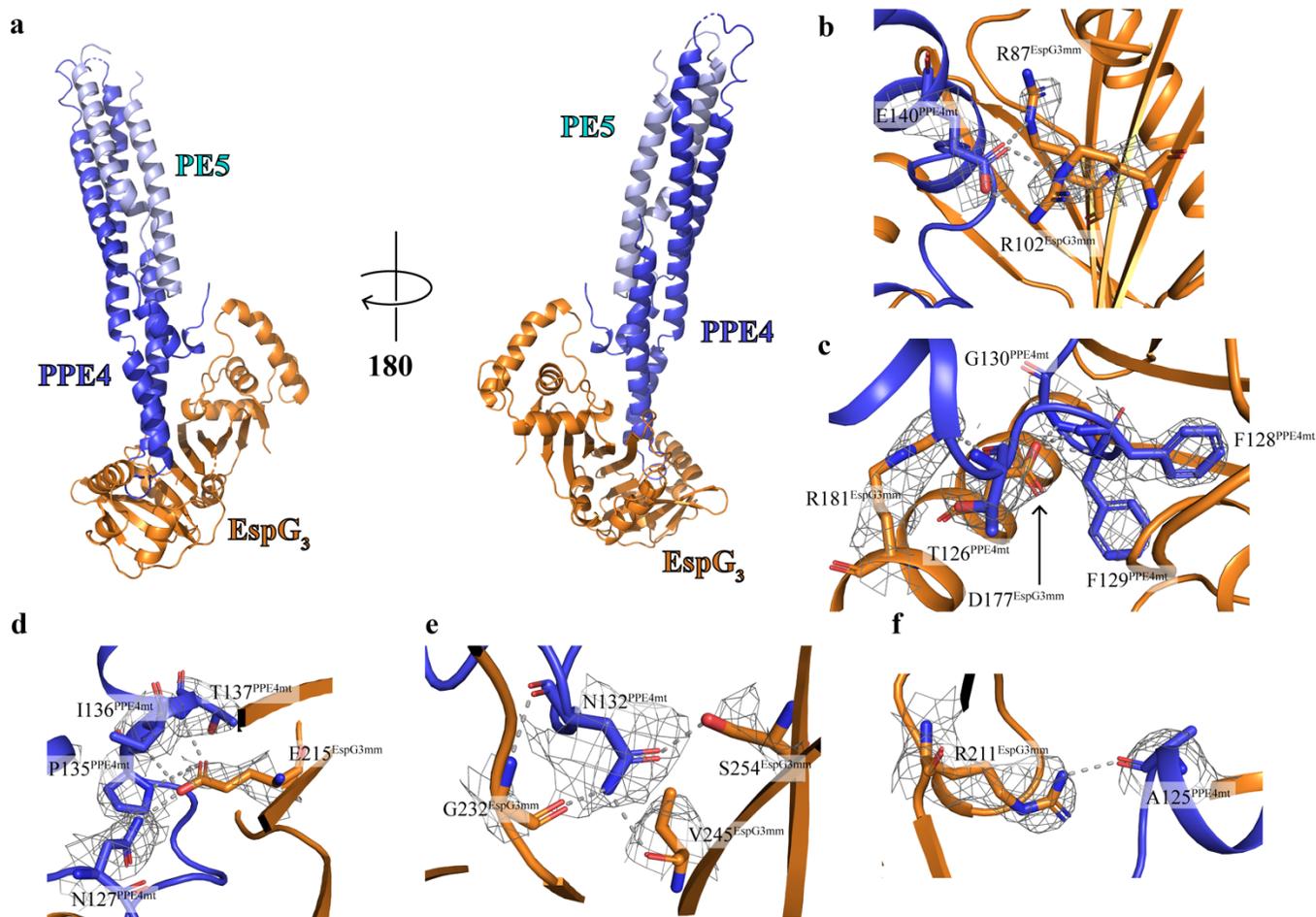


Figure 2. Crystal structure of the PE5_{mt}-PPE4_{mt}-EspG_{3mm} trimer and selected interactions in PPE4_{mt}-EspG_{3mm} interface. *a*, Crystal structure of the PE5_{mt}-PPE4_{mt}-EspG_{3mm} trimer in a cartoon representation with two views related by a rotation of 180°. EspG₃ interacts exclusively with the tip PPE4, distal to PE5. *b-f*, Interacting residues are shown with main chain and side chain in stick form with electron density map (2Fo-Fc shown at 1.0 σ) covering side chains and hydrogen bonding in grey dashed lines.

PE5-PPE4-EspG₃ structure from *ESX-3*

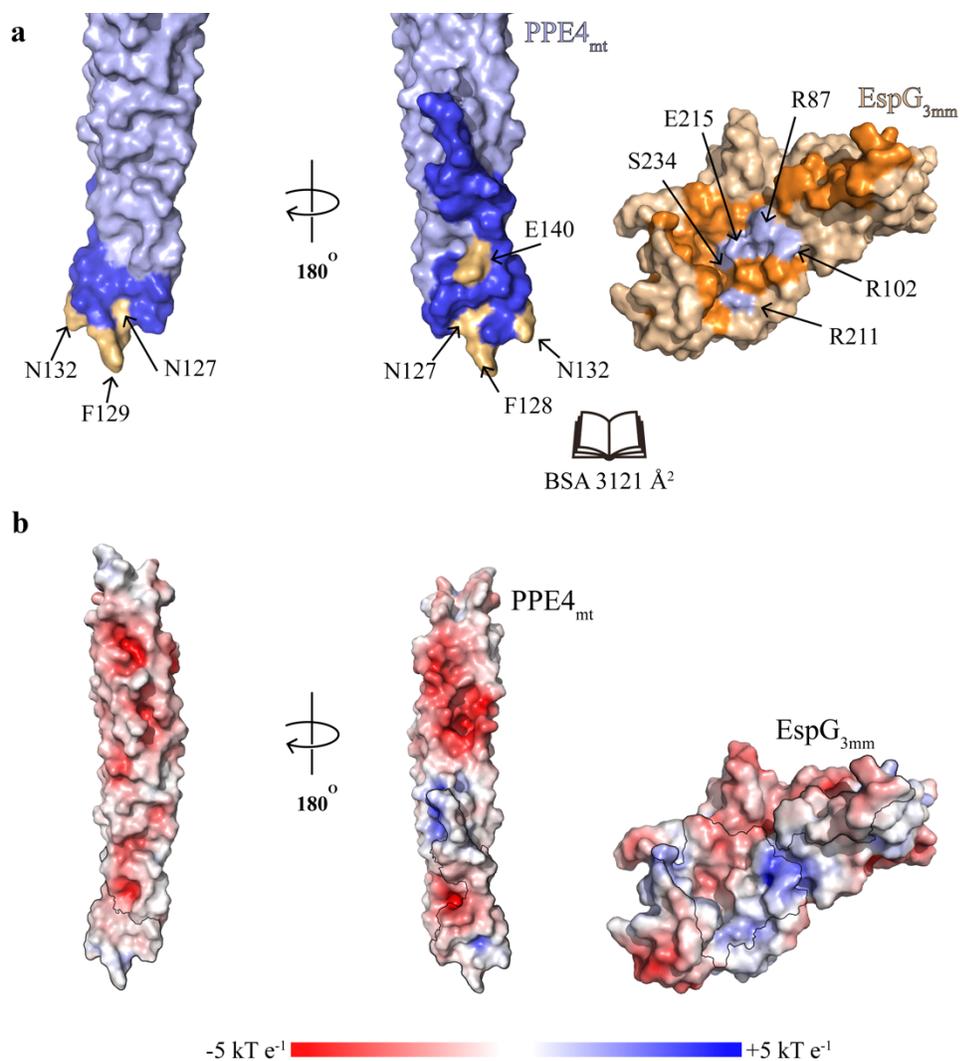


Figure 3. Interface between PPE4_{mt} and EspG_{3mm}. *a*, Surface representation of PPE4_{mt} and EspG_{3mm} shown in an ‘open book’ view. Interacting residues are colored in blue (PPE4_{mt}) and orange (EspG_{3mm}), with mutated residues highlighted in light orange (PPE4_{mt}) and light blue (EspG_{3mm}). Another view of PPE4_{mt}, related by a 180° rotation, is also shown. *b*, The same orientations as panel *a* but with the surface colored according to surface potential as calculated by APBS (ref). The interacting residues are circled within a black line.

PE5-PPE4-EspG₃ structure from ESX-3

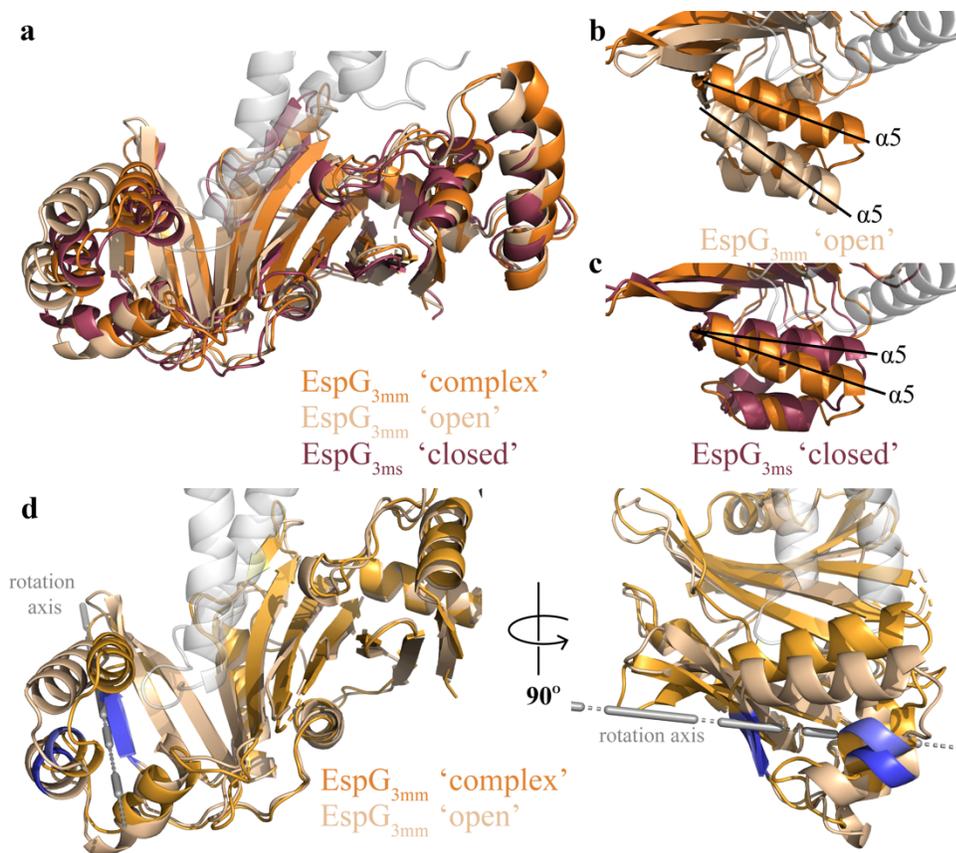


Figure 4. EspG₃ exists in multiple structural forms. *a*, An 'open' (5DLB, EspG_{3mm}, sand) and 'closed' (4RCL, EspG_{3ms}, maroon) conformation of EspG₃ were aligned to EspG_{3mm} (6UUJ, orange) as it is bound to PPE4_{mt}. Overall the different conformations align well to the bound conformation of EspG_{3mm} with r.m.s.d. of 2.1 Å ('open') and 1.9 Å ('closed'). *b and c*, Closeups highlighting the different orientations of the $\alpha 5$ helices in the 'open' (*b*) and 'closed' (*c*) EspG₃ structures as compared to EspG_{3mm} bound to PPE4_{mt}. *d*, Movement regions defined in EspG₃ as it moves from the 'open' conformation to the bound conformation in two different views related by a 90° rotation. The rotation axis for the moving domain is shown in gray. Each conformation maintains the same coloring as in panel *a*, with the hinge between the moving and fixed domains colored blue ('open') and light blue (bound).

PE5-PPE4-EspG₃ structure from ESX-3

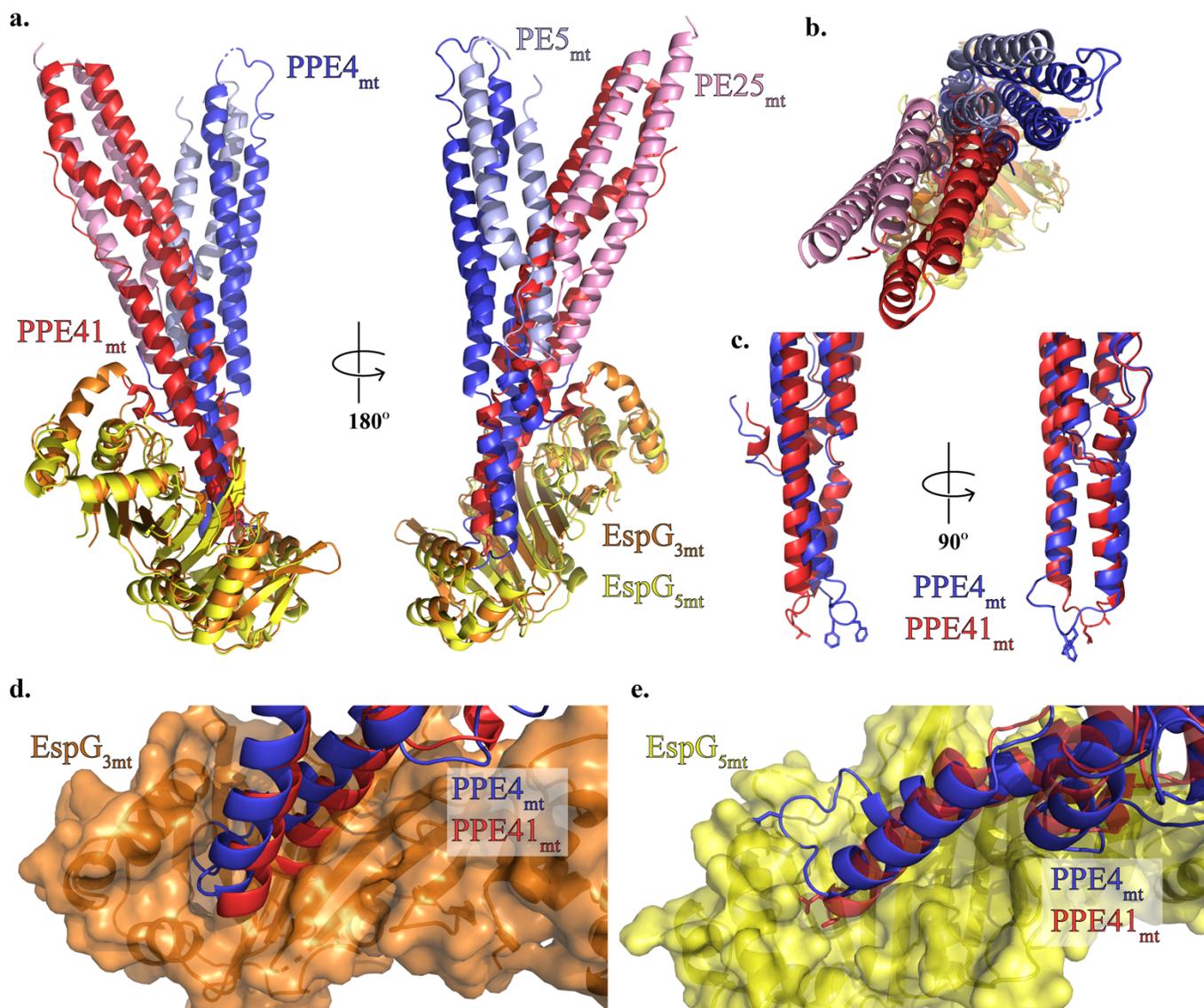


Figure 5. PE5_{mt}-PPE4_{mt} interacts with EspG_{3mt} chaperone in a unique mode compared to ESX-5 PE-PPE dimers. *a*, Structural alignment of the ESX-3 and ESX-5 trimers via the EspG chaperones (Dali ref) reveals a difference in the angle of interaction between the PE-PPE heterodimers with their respective chaperone. *b*, Top view of alignment from panel *a*. *c*, Alignment of PPE41_{mt} and PPE4_{mt} highlights difference in hh loop conformations between ESX-3 (PPE4_{mt}) and ESX-5 (PPE41_{mt}). *d-e*, Superposition of PPE alignment from panel *c* in context of EspG_{3mt} interaction (*d*) and EspG_{5mt} interaction (*e*) shows the incompatibility of each PPE protein with non-cognate chaperone binding.