Dimer recognition and secretion by the ESX secretion system in *Bacillus subtilis*

Tatyana A. Sysoeva, Martha A. Zepeda-Rivera, Laura A. Huppert, and Briana M. Burton[1]

Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138

Protein secretion typically involves translocation of unfolded polypeptides or transport of monomeric folded proteins. Here we provide, to our knowledge, the first experimental evidence for secretion of an intact multimeric complex requiring a signal formed by both members of the complex. Using systematic mutagenesis of a substrate involved in early secretory antigen 6 kDa (ESX) secretion in *Bacillus subtilis*, we demonstrate that export of the substrate requires two independent motifs. Using mixed dimers, we show that these motifs must form a composite secretion signal in which one motif is contributed by each subunit of the dimer. Finally, through targeted crosslinking we show that the dimer formed in the cell is likely secreted as a single unit. We discuss implications of this substrate recognition mechanism for the biogenesis and quality control of secretion substrates and describe its likely conservation across ESX systems.

WXG protein | type VII secretion system | protein translocation | YukE

Protein secretion is critical for protein targeting in any living cell and for its communication with the environment. Bacteria use a wide range of secretion mechanisms to export proteins out of the cytoplasm. Signals for secretion are most commonly primary amino acid sequences, but in some cases also may be formed through interacting surfaces of a substrate and its delivery effector. Some secretion systems unfold their substrates to translocate them across the membrane and cell wall. Other systems export folded proteins, sometimes in complex with bound cofactors. For example, the general secretory machinery (Sec) denatures the tertiary and secondary structure of its substrates to thread the polypeptide through the narrow opening of the integral membrane translocon complex, SecYEG (1). Type III secretion system (T3SS) machinery is thought to unfold the tertiary structure of its substrates, while preserving the secondary structure elements for the substrate recognition (2, 3). In contrast, the twin-arginine transport (Tat) system exports folded substrates (4) and is hypothesized to be able to translocate protein oligomers and complexes via a "hitchhiking" mechanism (5). Overall, these and other secretion types differ in the nature of substrate recognition signal and the mode of substrate translocation.

Early secretory antigen 6 kDa (ESX, or type VII) secretion systems are widespread in actinomycetes and Gram-positive bacteria and affect a range of bacterial processes including sporulation, conjugation, and cell wall stability (6–10). In two notorious human pathogens, *Mycobacterium tuberculosis* and *Staphylococcus aureus*, ESX secretion was found to be crucial for establishing and maintaining the infection (11–15). Despite the importance of the ESX secretion for human health, the mechanism of this type of secretion is still largely unknown.

Recent characterization of the ESX system in *Bacillus subtilis* confirmed that a functional system is encoded by the *yuk/yue* operon (16, 17). Importantly, the *B. subtilis* system codes for a homolog to the prototypical virulence factor substrates from mycobacteria, EsxA and EsxB (18). These proteins all belong to the WXG-100 protein family, which are defined by a conserved WXG amino acid motif that is roughly in the middle of the typically short, otherwise poorly conserved, ~100-amino-acid polypeptide (Fig. S1*A*). All characterized WXG proteins share

helix-turn-helix hairpin structures where the conserved WXG motif forms a sharp turn between the *N*- and C-terminal helices (19–23). Two hairpins interact in an antiparallel configuration to form either homo- or heterodimers (Fig. S1*B*) (19–24). This arrangement places the *N* and C termini of one subunit in close proximity with the WXG hairpin of the interacting partner. Previous studies have demonstrated that the C-terminal residues and the WXG motif are important for secretion of some mycobacterial substrates but not the others (7, 25, 26). However, dissecting the roles of these motifs has been challenging due to confounding effects, such as their contributions to substrate stability (7, 26, 27). Further complicating the issue, in the mycobacterial system, for example, multiple substrates have been identified and their secretion is codependent (13, 28). There are also multiple closely related ESX systems in mycobacteria that could complicate the analyses (29).

Therefore, to investigate the requirements for WXG substrate recognition and the mode of substrate translocation in ESX secretion, we took advantage of the fact that under standard laboratory conditions *B. subtilis* has a single WXG substrate, YukE (17). Here we report results of a systematic mutagenesis study combined with crosslinking experiments. First, we show that the C-terminal residues of the *B. subtilis* substrate YukE are important for secretion, suggesting a general mode of recognition of the ESX substrates in firmicutes and actinobacteria. Second, tryptophan and glycine residues of the WXG motif of YukE are required for an efficient transfer of YukE outside of the cell. Third, YukE forms stable homodimers that contain two sites composed by C terminus and WXG turn, but only one intact bipartite site is required for substrate export. Fourth, we present experimental evidence that the ESX system translocates the WXG protein dimers. Together our results show that the ESX system requires a composite signal formed by two folded polypeptides for secretion and exports an intact protein complex that possesses the bipartite signal.

Significance

Bacteria use a wide range of secretion mechanisms to export proteins from the cytoplasm. These secretion pathways differ in the nature of substrate recognition signal and the mode of substrate translocation. This work elucidates fundamental properties of substrate secretion by the early secretory antigen 6 kDa (ESX) system. We use systematic mutagenesis and targeted crosslinking of an ESX substrate from *Bacillus subtilis* to provide, to our knowledge, the first experimental evidence for secretion of an intact dimeric complex requiring a composite recognition signal formed by both members of the complex.

Results

YukE Homodimerizes in Vitro and in Vivo. All WXG substrates of the ESX systems studied to date have been shown to exist as stable homo- and heterodimers in solution and as dimers or tetramers in crystals. We purified recombinant YukE from *Escherichia coli* to test whether it can dimerize. Size-exclusion chromatography of the purified YukE protein showed that it forms homodimers in vitro, consistent with recently reported YukE–His$_6$ data (16)(Fig. 1). The hydrodynamic behavior of the YukE dimer is consistent with the 3D model of this protein (Figs. S1*B* and S2). Dimerization remains unchanged in the tested pH range from 5.5 to 7.5 (Fig. S3). We also did not observe a detectable fraction of YukE monomeric species at any of the tested conditions (20–150 mM NaCl; Fig. S3) and range of the concentrations (0.1–1 mM), indicating a tight homodimeric complex.

For mycobacterial WXG substrates, it was shown that heterodimerization of the substrates is crucial not only for stability of each of the polypeptides but also for their secretion (25, 26, 28). Such codependence for secretion was also found for non-WXG substrates of ESX systems (13, 26). To test whether dimerization is important for the ESX in *B. subtilis*, we started by establishing whether the YukE substrate forms dimers in vivo using cross-linking. Exploiting the fact that the YukE substrate is a naturally cysteine-less protein (Fig. S1*A*), we created a series of variants with one or two cysteines to test the potential of YukE to oligomerize in vivo. Cellular content and secreted proteins of the strains expressing YukE^{Cys} variants were analyzed for YukE presence with and without excess of reducing agent DTT (Fig. 1). As a result, we see robust formation of spontaneous disulfide bridges in some of the YukE variants containing cysteines, for example those containing the S62C substitution. In these strains, YukE-specific antibodies detected an abundant species at about 22 kDa in size, which is consistent with predicted YukE homodimer size (Fig. 1). Other tested YukE^{Cys} variants did not form a higher molecular weight species of the size of a dimeric substrate. Modeling of the threaded YukE sequence on a 3D structure reveals that disulfide bond formation at residue 62 is consistent with YukE homodimerization. From comparison of the oxidized and reduced samples, it is clear that our anti-YukE antibodies preferentially recognize epitopes in the dimeric form of YukE protein. This is consistent with the fact that the antibodies were raised against recombinantly purified YukE, which as confirmed above exists as a tight dimer in solution. Analyses of all tested YukE^{Cys} variants with cysteine residues along the helical "arms" of the protein (residues 9–42 and 47–82) confirmed that none of these substitutions caused significant changes in secretion levels of this substrate (Fig. 1). Together these results demonstrate a strong preference for YukE to exist in a homodimeric state and suggest that the paired portion of the helix-turn-helix dimer is generally permissive for mutations.

Alterations in the WEG Motif Affect Secretion of YukE. In the WXG protein family, only the W and G residues are almost absolutely conserved, whereas the rest of the sequence and the length of the polypeptide vary significantly (Fig. S1). Thus far, importance of the WXG motif was tested for EsxA, EsxB, and EspA proteins, suggesting that it is only important for EspA (Table S3). We decided to test whether the intact WXG motif of YukE, WEG, is important for YukE secretion. Single amino acid substitutions were introduced in the W44, E45, or G46 positions of YukE that was expressed ectopically under control of a strong inducible promoter in a Δ*yukE* background. We established that all of the tested variants are indeed expressed in *B. subtilis* (Fig. 2). Different variants of YukE accumulated in the cells to different extents; compare, for example, the intensity of the YukE bands in the lanes with E45A and G46A variants. We purified recombinant versions of these variants from *E. coli* and confirmed that their stability in vitro was not affected. Therefore, protein stability issues are unlikely to account for the differential cellular accumulation and secretion of these mutants. Substitutions at the W44 position were well represented in the cell lysates, unlike the three tested G^{46} mutations that accumulated to lower levels. Analyses for the presence of YukE variants in the cells and in the spent medium showed that tryptophan and glycine residues are absolutely required for efficient transport of the substrate (Fig. 2). Only changes of the tryptophan to other bulky aromatic residues phenylalanine and tyrosine resulted in a low level of detectable YukE in the cells' supernatant. All other changes resulted in loss of YukE in the spent medium. At the same time, the glutamate residue can be substituted to different amino acids without total loss of secretion (Fig. 2). The levels of secreted YukE with modified position 45 varied depending upon identity of the residue in this position. Switching the negatively charged glutamate E45 into a positive residue such as lysine or arginine did not disrupt the secretion of YukE. Meanwhile, introduction of the small glycine residue at this position dramatically reduced export of the protein. Overall analyses of WXG substitutions in YukE show that tryptophan and glycine residues are required for secretion, whereas identity of the middle position can modulate levels of YukE secretion.

C-Terminal Residues of YukE Are Crucial for Efficient Secretion. For mycobacterial ESX substrate EsxB (CFP-10), several C-terminal residues were shown to be crucial for secretion, but there are no analogous residues in EsxA (26). Most recently, a C-terminal consensus motif YxxxD/E was found in several mycobacterial substrates (30). Furthermore, in other WXG-100 substrates, a weaker HxxxD/ExxHxxxH motif was identified (31). To date, the importance of this motif was only tested for some mycobacterial substrates. Thus, we wondered whether the C-terminal sequence of YukE is important for its secretion, similar to the studied ESX substrates, or whether the specialization of the various ESX systems has led to disparate substrate recognition mechanisms. To test this, in a Δ*yukE* background, we incorporated the *yukE* gene coding for a series of C-terminally truncated YukE variants into an ectopic site under control of an
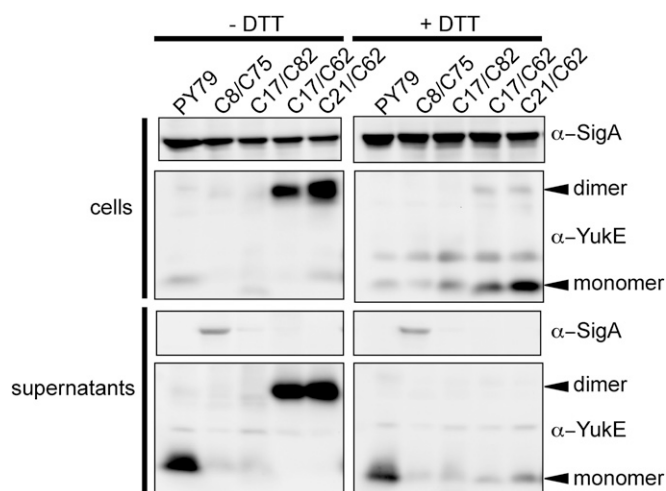
Fig. 1. YukE homodimerizes in vivo. Spontaneous disulfide bridge formation in modified YukE proteins with introduced cysteine residues. *B. subtilis* cellular proteins were separated by SDS/PAGE and analyzed by Western blot with YukE-specific antibodies. PY79 is the wild-type strain; YukE proteins with cysteines at specified positions in the Δ*yukE* background. In the oxidized samples (*Left*), a prominent 22 kDa band is apparent for some of the strains, and it is labeled with an arrow. This species is consistent with formation of the YukE dimer, which can be reduced by adding excess of a reducing agent (*Right*).
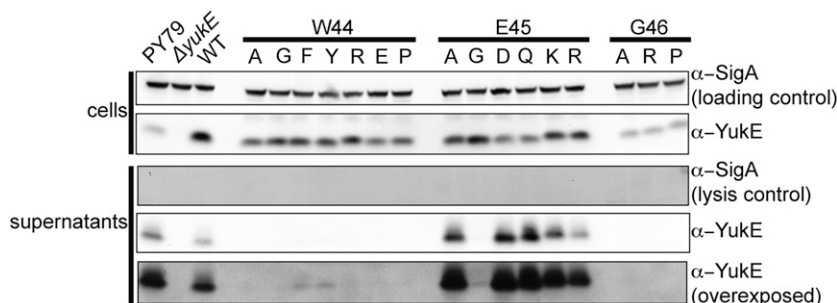
Fig. 2. Effect of single amino acid substitutions in the conserved WEG motif of the YukE substrate. YukE variants with mutations in the defining WEG motif were expressed in *B. subtilis* cells in a Δ*yukE* background. PY79 is the wild-type strain; the "WT" label defines the strain expressing wild-type YukE protein from an ectopic locus in a Δ*yukE* background. Obtained strains were tested for their ability to produce and secrete YukE.

isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible promoter, and we assessed secretion of the truncation mutants. Surprisingly, truncation of even a single amino acid from the carboxyl terminus reduced the efficiency of YukE secretion. Moreover, we could not detect in the spent medium any of the YukE missing five or more residues at the C terminus (Fig. 3). Notably, the accumulation of the YukE mutants in cells varied, which may reflect reduced stability of some of the truncations. However, despite high levels of YukE missing 7 or 12 residues, neither of these variants could be detected in the growth medium (Fig. 3). These results indicate that the C-terminal tail of YukE is crucial for substrate secretion. The data also suggest that although the absolute sequences are distinct in YukE, staphylococcal, and mycobacterial substrates, the requirement for C-terminal sequence recognition during ESX secretion is conserved.

Dimerization of the WXG-Substituted YukE as Well as of Its Truncated Variants Is Not Affected in Vitro or in Vivo. Based on the result that YukE homodimerizes in vivo, we hypothesized that inability of the YukE mutants to dimerize may explain the lack of secretion of these proteins. Therefore, we set out to check whether the dimerization of the YukE variants is disrupted. To this end, we first purified five mutant forms of YukE from *E. coli*: the three alanine substitutions of the WXG motif (W44A, E45A, and G46A) and two C-terminal truncation variants of YukE (YukE$^{C\Delta5}$ or YukE$^{C\Delta15}$). These proteins were then analyzed via gel filtration (Fig. S2). As we had seen for the wild-type YukE protein, all five variants eluted with profiles consistent with forming homodimeric species in vitro. Having in hand an in vivo assay to test dimerization of YukE in *B. subtilis*, we next analyzed whether the mutant versions of YukE dimerize inside the cells. We introduced the S62C mutation into the WXG-substituted or C-terminally truncated YukE variants (mutation^{Cys}) and tested whether each of these proteins retained the ability to form spontaneous disulfides in cellular lysates (Fig. 4). Our results revealed that none of the tested variants are disrupted in dimerization, confirming that dimerization of the substrate does not rely on an intact WXG motif nor on the C terminus alone. The observed dimerization of the mutant forms of YukE supports the notion that the defect in secretion of these variants likely stems from another attribute of the YukE substrate. Combined, our results thus far pointed to the possibility that there are two signal determinants in the YukE dimer, which are involved in substrate recognition in ESX, the WXG motif, and the C-terminal tail.

Secretion of an Irreversibly Cross-Linked YukE Dimer. Based on the signal determinants that we had identified thus far, we wondered whether after being appropriately recognized, the proteins would be disassembled from one another in order for translocation to occur. The intuitive hypothesis was that the protein complex, which had already formed, would be maintained through the

secretion process. However, to our knowledge, such a mechanism does not have a precedent.

We exploited the insensitivity of the *yuk* system to cysteines in YukE variants and tested whether YukE homodimers that are irreversibly crosslinked inside the bacterial cells could still be secreted by the ESX system. To do this, we engineered strains expressing different single- and double-cysteine YukE substitutions and tested them with three irreversible, membrane-permeable maleimide crosslinkers for sulfhydryl conjugation. We next expressed the optimal YukE$^{V21C/S62C}$ (hereafter referred to as YukE^{2Cys}) and used bismaleimidoethane (BMOE) to crosslink the YukE dimer inside the bacterial cells. The excess of the crosslinker and all of the extracellular YukE protein was extensively washed away from the cells and cells were then allowed to recover in medium without the crosslinker. We then analyzed the secreted and cytosolic fractions from each stage of this experiment to assess whether cells were able to secrete the crosslinked WXG substrate. Fig. 5 illustrates that BMOE indeed irreversibly crosslinks two YukE subunits and, unlike the disulfides in Fig. 1, cannot be reversed by excess of a reducing agent. Strikingly, comparison of the S3 fractions shows that irreversible crosslinked YukE dimer accumulates in the medium and cannot be reduced with DTT. This result is consistent with the hypothesis that a dimeric ESX substrate can be exported as a dimer unit.

Testing Dominance of YukE Variants. Given that truncation mutants dimerize in vivo but cannot be secreted, we next asked about the effect of the secretion-disrupting mutations in the context of the dimer. To do this, we assessed the dominance of the previously established mutations. We introduced the modified variants of YukE at an ectopic site into cells that contained the endogenous, wild-type copy of YukE under its native promoter and tested these strains for YukE secretion. Our experiments revealed that C-terminally truncated YukE substrates slightly reduce levels of the full-length YukE in the lysates and dramatically reduce secretion of the full-length YukE (Fig. 6*A*).
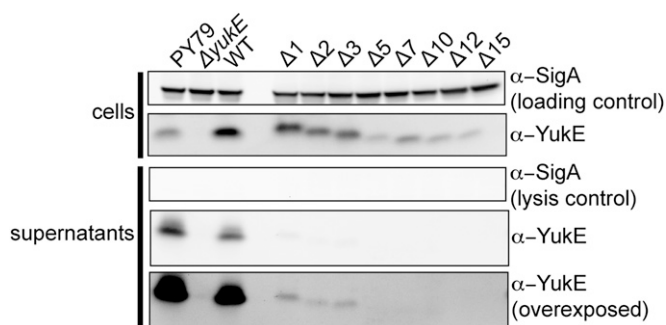
Fig. 3. Effect of C-terminal truncations of the YukE substrate on its secretion. YukE variants deleted for 1–15 C-terminal residues were expressed in a Δ*yukE* background and tested for secretion. Labeling is as in Fig. 2.
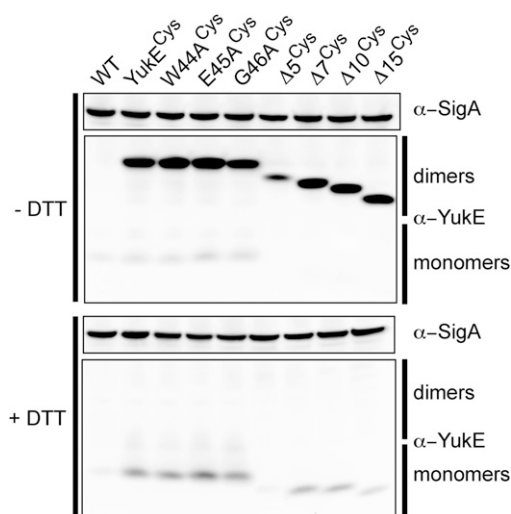
Fig. 4. WXG-substituted and C-terminally truncated YukE variants form dimers. Dimerization of the modified YukE substrate in vivo using spontaneous disulfide bridge formation between S62C of the two YukE subunits in the dimeric complex.

To understand this effect better, we titrated the inducer to vary the ratio of the truncated version to the full-length protein. This experiment produced two notable outcomes. First, as expected, increasing levels of the mutant still caused a partial dominant-negative effect on secretion of the full-length protein. However, at the highest concentrations of truncated YukE, where the stoichiometry of the two was roughly equivalent, we could now observe a rescue of secretion of the truncated YukE variant (Fig. 6*B*). Together our data indicate a heterodimer formed by the truncated and full-length substrates is proficient in secretion, suggesting that an intact composite signal at one end may be sufficient for recognition and secretion.

Discussion

Here we dissect the fundamental features of substrate recognition and secretion by the ESX system in *B. subtilis*. Analysis of site-directed and truncation mutants revealed that two sets of features, the WXG motif and the C-terminal residues, are required to form a composite recognition signal for substrate secretion. We confirm that the *B. subtilis* substrate exists as a dimer and that this dimer is secreted as a dimer, but secretion happens only when an intact composite recognition signal is present. Thus, we propose a model whereby the ESX system worksthrough recognition of a bipartite site on one end of the substrate dimer followed by translocation of the folded complex (Fig. 7).

Substrate secretion requires a signal for export. Our data support the idea that there are two major features that are critical for YukE as an ESX substrate: C-terminal residues and conserved residues of the WXG motif. YukE, like the other ESX substrates, lacks any recognizable N-terminal signal sequence. Although, unlike mycobacterial substrates, sequence analyses of YukE show presence of only a weak motif at the C terminus (31) (Fig. S1*A*), we found that C-terminal residues are required for efficient secretion of YukE. That makes YukE similar to mycobacterial EsxB and PE25 substrates but different from the EsxA-like proteins. Interestingly, truncation of as few as five amino acids from the C terminus of YukE disrupts its secretion (Fig. 3); this was, to our knowledge, the first reported series of truncations combined with tuning of the expression levels. Next our extensive mutagenesis of YukE showed that the most conserved residues in the family-defining WXG motif are required for export of this protein rather than for its stability, as was suggested earlier (Fig. 2) (25). In fact, tryptophan and glycine

positions are not permissive for substitutions, whereas changes to the middle glutamate are tolerated and they modulate the efficiency of YukE secretion. Intriguingly, the available 3D structures of the WXG motifs (Fig. S1*B*) indicate that the middle residue of this turn points away from the interhelical space, facing outward in the context of the substrate dimer. We propose that there might be differential interactions with the three residues or that this particular segment is involved at distinct stages of the translocation through the ESX system. Further investigation is required to test these hypotheses.

From several studies it is clear that dimerization of WXG family substrates is an important feature of their biology. Mycobacterial and staphylococcal substrates of ESX systems tend to dimerize similarly to the heterodimerization of prototypical *M. tuberculosis* substrates EsxA and EsxB (19–24, 32). In *S. aureus* both EsxA and EsxB homologs are present, but EsxA was shown to homodimerize similarly to the GBS1074 protein from *Streptococcus agalactiae* (21, 22). Our experiments show that YukE, the substrate of the *B. subtilis yuk* system, homodimerizes both in cellular lysates and in medium. This is consistent with recently reported dimerization of YukE in the supernatants of an undomesticated *B. subtilis* strain (16). In vitro tests of recombinantly purified YukE show that similar to other WXG dimers, YukE dimers are extremely stable under a variety of solution conditions.

Our tests in vivo and in vitro ruled out disrupted dimerization as a cause for the secretion defect of truncated YukE variants. This in turn led us to structural considerations for the secretion signal. In the context of the wild-type YukE dimer, the two recognition determinants are in close proximity to one another; the WXG motif of one subunit is immediately proximal to the C-terminal sequence of the other subunit. Within a dimer, two of these composite recognition signals exist; they are symmetrically located at the "ends" of each elongated homodimer (Fig. S1*B*). Our tests with strains carrying mixed populations of full-length and truncated YukE substrates revealed that wild-type protein can partly rescue secretion of the C-terminally truncated copy (Fig. 6*B*). Therefore, we propose the ESX substrate in *B. subtilis* is recognized through an intact bipartite recognition site. This
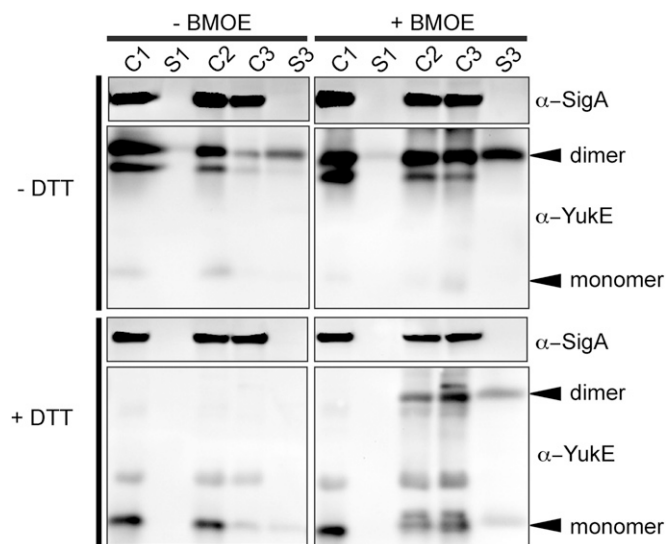
Fig. 5. Irreversibly crosslinked YukE dimer can be secreted. YukE Western blot analysis of the control cell lysates (C1) and medium supernatant (S1) samples, BMOE-treated cells (C2), cells (C3) recovered in the absence of crosslinker, and newly secreted proteins (S3). The dimeric YukE species, crosslinked while inside the cells, is present in the final supernatant (S3) sample even after adding the excess of reducing agent DTT.
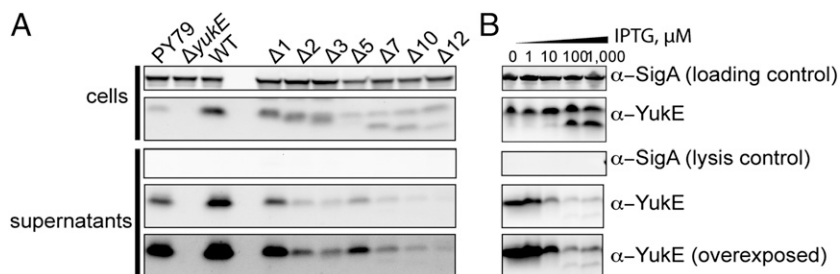
Fig. 6. Secretion of truncated YukE protein can be rescued by the presence of full-length YukE. (*A*) C-terminal truncations of 5–15 amino acids interfere with the secretion of the full-length protein. Secretion of the truncated substrate can be partially rescued by the wild-type YukE. (*B*) Accumulation of the truncated forms of YukE in the presence of the full-length substrate rescues the secretion of the truncated version. Titration of YukECΔ7 is shown.

mode of recognition would explain our results, as well as accumulated data about other ESX substrates. For example, it is known that the C terminus of EsxB is critical for secretion, unlike the tail of EsxA (Table S3). In the case of other mycobacterial substrates such as heteropairs EsxG/EsxH and EsxS/EsxR, one of each of the tryptophans of the WXG motifs is substituted by a histidine side chain (33, 34) (Fig. S1), leaving only one intact WXG motif per pair. In the context of our model, the intact WXG motif, along with a second signal on one end of the complex, would form the necessary composite recognition site for secretion.

The major, well-studied Sec and Tat secretory mechanisms illustrate two possible modes of protein export in which the substrate is either completely unfolded during translocation or in which the substrate preserves its tertiary structure. To gain insights into the functioning of the ESX system, we addressed the following fundamental question: In what state is the WXG substrate translocated via the secretory machinery? Based on the established tight interaction both in cell lysates and supernantants within the heterodimer EsxA/EsxB and the mutual dependencies of these two proteins for stability and secretion, the prevailing idea in the field has been that the ESX system may secrete obligatory dimers. However, this idea has not been addressed by direct experimentation. Here we established that YukE behaves as a stable homodimer similar to the mycobacterial substrates, and we developed the crosslinking tools for stabilizing and trapping the dimerization in vivo. With this at hand, we challenged the *yuk* machinery to secrete irreversibly crosslinked YukE dimers and showed that these dimers indeed accumulate in the spent medium. Such results are suggestive of the ability of the *yuk* secretory apparatus to translocate intact substrate dimers. This is, to our knowledge, the first experimental evidence that the ESX system secretes complex folded structures and, to our knowledge, the first example of a secretory system that secretes a quaternary structure depending upon a composite signal. Whereas the Tat secretion system secretes folded proteins and has even been proposed to secrete quaternary structures via a "hitchhiker" mechanism, the Tat secretion signal is present on only one subunit in a substrate complex (30).

Several secretion systems are hypothesized to recognize a composite recognition site that does not constitute a single stretch of a substrate polypeptide. In the T3SS and Type IV Secretion System, special proteins serve as chaperones to supply an additional recognition site and deliver the partly unfolded substrate to the secretion apparatus. Based on established structures of chaperone–substrate complexes, it seems that unfolding of the substrate exposes the substrate recognition site for the secretory apparatus. In ESX secretion, on the contrary, the two subunits are more stable in the heterocomplex as shown in multiple in vivo studies. In vitro studies of the heterocomplex formation also document that dimerization increases helicity of the subunits, signifying that the complex is more "folded" than are individual subunits in solution. Considering the "chaperone" hypothesis further, we note that ESX secretion is also unusual because both polypeptides of the WXG or PE/PEE substrates are exported to the medium, unlike chaperones of other

secretion systems. Thus, our data are, to our knowledge, the first to demonstrate functional secretion of a complex in this manner.

The proximal arrangement of the two substrate features that form the composite signal and come from different polypeptides once again brings our attention to the 3D aspect of ESX substrate recognition. Even in the absence of the sequence conservation, the ESX substrates carry distinct 3D patterns composing the bipartite stereochemical recognition site, rather than a simple signal sequence. About 10% of the genomes of some *Mycobacteria* are dedicated to encoding numerous proteins from Pro-Glu (PE) and Pro-Pro-Glu (PPE) families. Many members of the families are confirmed substrates of the mycobacterial ESX systems. The only available structure of a PE/PPE heterodimer (35) reveals that the heterodimer possesses an EsxA-like C-terminal tail closely opposed to a WXG motif from the partner protein. Moreover, there is only one such set present in this heterodimer, which according to our model would be sufficient to direct export. Therefore, we propose that recognition of the disparate families of ESX substrates may follow a similar pattern to the rules we have described here for YukE. Two flavors of ESX systems have been defined in firmicutes and actinobacteria (18). One commonality among the systems is the presence of the prototypic WXG-100 class of substrates. The present study combined with the observations discussed above suggests that both subtypes of the ESX system may share a similar mode of substrate recognition and translocation.

B. subtilis to the extent of our knowledge has only one ESX apparatus and one known substrate under standard laboratory conditions (17). This setup is particularly important at these initial steps of the investigation of the general mechanism of ESX secretion. The proposed mode of recognition of the composite site in the folded protein complex can provide an explanation on how several ESX secretion machines in mycobacterial
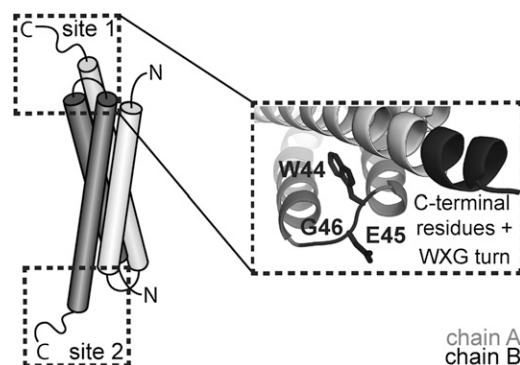
Fig. 7. WXG substrate composite recognition signal model. Wild-type YukE contains two equivalent recognition sites located on each end of the elongated dimer consisting of four α-helices. The WXG loop of one subunit and the C-terminal residues of the second subunit constitute the bipartite recognition signal. One intact composite recognition site is sufficient for secretion of the complexed, folded substrate.

species may distinguish their numerous substrates by fine-tuning their sequence and 3D structure.

Materials and Methods

Strains and Protein Purification. Plasmids were constructed and maintained using standard protocols (36). Strains and oligonucleotides are listed in Tables S1 and S2. The coding sequence of YukE was cloned into pET28b+ vector (Novagen) and *B. subtilis* into integration vector pDR111 [gift of D. Rudner (Harvard Medical School, Boston)]. Details of cloning and purification protocols are described in *SI Materials and Methods*.

Size-Exclusion Chromatography of YukE Protein. YukE variants were run on a pre-equilibrated Superdex200 (GE Pharmacia) column using 25 mM Hepes pH 7.5, 150 mM NaCl, 5% (wt/vol) glycerol at a flow rate of 0.5 mL/min. Where indicated, the proteins were also analyzed under low-salt (20 mM NaCl) or at different acidity (25 mM Mes pH 6.5 and pH 5.5) buffer conditions. Protein elution was monitored by absorbance at 280 nm, and fractions were analyzed by SDS/PAGE gel and immunoblotting or Coomassie staining if needed. The BioRad gel filtration standards used for column calibration were as follows: thyroglobulin (670 kDa), γ-globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa), and vitamin B (1.35 kDa).

YukE Secretion Assay. *B. subtilis* strains expressing *yukE* variants from an ectopic locus were grown in LB medium supplemented with 0.3–1 mM IPTG for induction of the ectopic expression. For titration experiments, concentrations from 10 μM to 1 mM were used. Cells were inoculated at an OD_{600} of 0.05 and grown at 37 °C to an OD_{600} of 1.5–2.5. Total protein from cell-free conditioned *B. subtilis* supernatants was precipitated by adding 10% of the initial volume of 100% TCA (BD Chemicals) and incubation at 4 °C overnight. Precipitated proteins were collected via centrifugation at 15,000 rpm for

20 min at 4 °C in a SS34 rotor (Sorvall 5C). Pellets were washed with 10 mL of cold acetone, collected by the second spin (15,000 rpm, 20 min, using an SS34 Sorvall 5C rotor), and air dried to remove the excess of acetone. Proteins were resuspended in 1× SDS/PAGE loading buffer with or without added 100 mM DTT for subsequent SDS–Tris–Tricine–PAGE analysis. Cell lysate samples were prepared by incubating 1.5 OD_{600} cells in the lysis buffer (0.1 mg/mL RNaseA, 0.01 mg/mL DNase I, 1 mg/mL lysozyme, 1 mM PMSF) for 20 min at 37 °C followed by adding 2× SDS/PAGE loading buffer. SigmaA was used as a lysis control for supernatant samples and as a loading control for cell lysates. Coomassie staining of supernatant samples was used for normalization of loading of precipitated samples.

Crosslinking of YukE^{2Cys} in Vivo. Strains were grown in 30 mL LB to an OD_{600} of 0.6–1 in the presence of 1 mM IPTG to allow expression of the YukE^{2Cys} protein. Growth medium (S1) and cell aliquots (C1) were collected for YukE content analyses. Cells were harvested (15 min at 4,000 × *g* at ~25 °C) and washed twice with 30 mL LB medium to remove extracellular proteins. Next, fresh LB with a 0.5 mM BMOE crosslinker (Pierce) was added to the cells, and they were incubated further with agitation for 30 min at 37 °C. Aliquots of cells (1.5 OD_{600}) were collected to confirm that the crosslinker permeated the cells and crosslinked intracellular YukE (C2). After the incubation, the cells were collected by centrifugation and washed twice with 30 mL of crosslinker-free, prewarmed LB broth. In the third fresh LB broth, cells were allowed to recover for 1 h with agitation at 37 °C. Recovered bacteria (C3) and the spent broth (S3) were collected for analyses as described above.

1. Lycklama A Nijeholt JA, Driessen AJ (2012) The bacterial Sec-translocase: Structure and mechanism. *Philos Trans R Soc Lond B Biol Sci* 367(1592):1016–1028.
2. Akeda Y, Galán JE (2005) Chaperone release and unfolding of substrates in type III secretion. *Nature* 437(7060):911–915.
3. Evdokimov AG, et al. (2003) Similar modes of polypeptide recognition by export chaperones in flagellar biosynthesis and type III secretion. *Nat Struct Biol* 10(10):789–793.
4. Lee PA, Tullman-Ercek D, Georgiou G (2006) The bacterial twin-arginine translocation pathway. *Annu Rev Microbiol* 60:373–395.
5. Rodrigue A, Chanal A, Beck K, Müller M, Wu LF (1999) Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway. *J Biol Chem* 274(19):13223–13228.
6. Akpe San Roman S, et al. (2010) A heterodimer of EsxA and EsxB is involved in sporulation and is secreted by a type VII secretion system in Streptomyces coelicolor. *Microbiology* 156(Pt 6):1719–1729.
7. Chen JM, et al. (2013) Phenotypic profiling of Mycobacterium tuberculosis EspA point mutants reveals that blockage of ESAT-6 and CFP-10 secretion in vitro does not always correlate with attenuation of virulence. *J Bacteriol* 195(24):5421–5430.
8. Garces A, et al. (2010) EspA acts as a critical mediator of ESX1-dependent virulence in Mycobacterium tuberculosis by affecting bacterial cell wall integrity. *PLoS Pathog* 6(6):e1000957.
9. Coros A, Callahan B, Battaglioli E, Derbyshire KM (2008) The specialized secretory apparatus ESX-1 is essential for DNA transfer in Mycobacterium smegmatis. *Mol Microbiol* 69(4):794–808.
10. Gray TA, Krywy JA, Harold J, Palumbo MJ, Derbyshire KM (2013) Distributive conjugal transfer in mycobacteria generates progeny with meiotic-like genome-wide mosaicism, allowing mapping of a mating identity locus. *PLoS Biol* 11(7):e1001602.
11. Burts ML, DeDent AC, Missiakas DM (2008) EsaC substrate for the ESAT-6 secretion pathway and its role in persistent infections of Staphylococcus aureus. *Mol Microbiol* 69(3):736–746.
12. Burts ML, Williams WA, DeBord K, Missiakas DM (2005) EsxA and EsxB are secreted by an ESAT-6-like system that is required for the pathogenesis of Staphylococcus aureus infections. *Proc Natl Acad Sci USA* 102(4):1169–1174.
13. Fortune SM, et al. (2005) Mutually dependent secretion of proteins required for mycobacterial virulence. *Proc Natl Acad Sci USA* 102(30):10676–10681.
14. Hsu T, et al. (2003) The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc Natl Acad Sci USA* 100(21):12420–12425.
15. Lewis KN, et al. (2003) Deletion of RD1 from Mycobacterium tuberculosis mimics bacille Calmette-Guérin attenuation. *J Infect Dis* 187(1):117–123.
16. Baptista C, Barreto HC, São-José C (2013) High levels of DegU-P activate an Esat-6-like secretion system in Bacillus subtilis. *PLoS ONE* 8(7):e67840.
17. Huppert LA, et al. (2014) The ESX system in *Bacillus subtilis* mediates protein secretion. *PLoS ONE*, 10.1371/journal.pone.0096267.
18. Pallen MJ (2002) The ESAT-6/WXG100 superfamily—And a new Gram-positive secretion system? *Trends Microbiol* 10(5):209–212.
19. Arbing MA, et al. (2010) The crystal structure of the Mycobacterium tuberculosis Rv3019c-Rv3020c ESX complex reveals a domain-swapped heterotetramer. *Protein Sci* 19(9):1692–1703.
20. Ilghari D, et al. (2011) Solution structure of the Mycobacterium tuberculosis EsxG·EsxH complex: Functional implications and comparisons with other M. tuberculosis Esx family complexes. *J Biol Chem* 286(34):29993–30002.
21. Shukla A, Pallen M, Anthony M, White SA (2010) The homodimeric GBS1074 from Streptococcus agalactiae. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 66(Pt 11):1421–1425.
22. Sundaramoorthy R, Fyfe PK, Hunter WN (2008) Structure of Staphylococcus aureus EsxA suggests a contribution to virulence by action as a transport chaperone and/or adaptor protein. *J Mol Biol* 383(3):603–614.
23. Renshaw PS, et al. (2005) Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. *EMBO J* 24(14):2491–2498.
24. Renshaw PS, et al. (2002) Conclusive evidence that the major T-cell antigens of the Mycobacterium tuberculosis complex ESAT-6 and CFP-10 form a tight, 1:1 complex and characterization of the structural properties of ESAT-6, CFP-10, and the ESAT-6*CFP-10 complex. Implications for pathogenesis and virulence. *J Biol Chem* 277(24):21598–21603.
25. Brodin P, et al. (2005) Functional analysis of early secreted antigenic target-6, the dominant T-cell antigen of Mycobacterium tuberculosis, reveals key residues involved in secretion, complex formation, virulence, and immunogenicity. *J Biol Chem* 280(40):33953–33959.
26. Champion PA, Stanley SA, Champion MM, Brown EJ, Cox JS (2006) C-terminal signal sequence promotes virulence factor secretion in Mycobacterium tuberculosis. *Science* 313(5793):1632–1636.
27. Garufi G, Butler E, Missiakas D (2008) ESAT-6-like protein secretion in Bacillus anthracis. *J Bacteriol* 190(21):7004–7011.
28. Champion PA, Champion MM, Manzanillo P, Cox JS (2009) ESX-1 secreted virulence factors are recognized by multiple cytosolic AAA ATPases in pathogenic mycobacteria. *Mol Microbiol* 73(5):950–962.
29. Abdallah AM, et al. (2007) Type VII secretion—Mycobacteria show the way. *Nat Rev Microbiol* 5(11):883–891.
30. Daleke MH, et al. (2012) General secretion signal for the mycobacterial type VII secretion pathway. *Proc Natl Acad Sci USA* 109(28):11342–11347.
31. Poulsen C, Panjikar S, Holton SJ, Wilmanns M, Song YH (2014) WXG100 protein superfamily consists of three subfamilies and exhibits an α-helical C-terminal conserved residue pattern. *PLoS ONE* 9(2):e89313.
32. Brodin P, et al. (2006) Dissection of ESAT-6 system 1 of Mycobacterium tuberculosis and impact on immunogenicity and virulence. *Infect Immun* 74(1):88–98.
33. Siegrist MS, et al. (2009) Mycobacterial Esx-3 is required for mycobactin-mediated iron acquisition. *Proc Natl Acad Sci USA* 106(44):18792–18797.
34. Serafini A, Boldrin F, Palù G, Manganelli R (2009) Characterization of a Mycobacterium tuberculosis ESX-3 conditional mutant: Essentiality and rescue by iron and zinc. *J Bacteriol* 191(20):6340–6344.
35. Strong M, et al. (2006) Toward the structural genomics of complexes: Crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. *Proc Natl Acad Sci USA* 103(7):8060–8065.
36. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY).