

Structural Characterization of the *Yersinia pestis* Type III Secretion System Needle Protein YscF in Complex with Its Heterodimeric Chaperone YscE/YscG

Ping Sun, Joseph E. Tropea, Brian P. Austin, Scott Cherry and David S. Waugh*

Macromolecular
Crystallography Laboratory,
Center for Cancer Research,
National Cancer Institute at
Frederick, P.O. Box B, Frederick,
MD, USA

Received 15 August 2007;
received in revised form
11 December 2007;
accepted 21 December 2007
Available online
5 January 2008

The plague-causing bacterium *Yersinia pestis* utilizes a type III secretion system to deliver effector proteins into mammalian cells where they interfere with signal transduction pathways that mediate phagocytosis and the inflammatory response. Effector proteins are injected through a hollow needle structure composed of the protein YscF. YscG and YscE act as “chaperones” to prevent premature polymerization of YscF in the cytosol of the bacterium prior to assembly of the needle. Here, we report the crystal structure of the YscEFG protein complex at 1.8 Å resolution. Overall, the structure is similar to that of the analogous PscEFG complex from the *Pseudomonas aeruginosa* type III secretion system, but there are noteworthy differences. The structure confirms that, like PscG, YscG is a member of the tetratricopeptide repeat family of proteins. YscG binds tightly to the C-terminal half of YscF, implying that it is this region of YscF that controls its polymerization into the needle structure. YscE interacts with the N-terminal tetratricopeptide repeat motif of YscG but makes very little direct contact with YscF. Its function may be to stabilize the structure of YscG and/or to participate in recruiting the complex to the secretion apparatus. No electron density could be observed for the 49 N-terminal residues of YscF. This and additional evidence suggest that the N-terminus of YscF is disordered in the complex with YscE and YscG. As expected, conserved residues in the C-terminal half of YscF mediate important intra- and intermolecular interactions in the complex. Moreover, the phenotypes of some previously characterized mutations in the C-terminal half of YscF can be rationalized in terms of the structure of the heterotrimeric YscEFG complex.

Published by Elsevier Ltd.

Edited by G. Schulz

Keywords: type III secretion; plague; *Yersinia*; tetratricopeptide repeat; needle protein

Introduction

Yersinia pestis, the causative agent of plague, is arguably the deadliest pathogen in history and a likely instrument of bioterrorism.^{1,2} Like many other

Gram-negative bacterial pathogens of plants and animals, *Y. pestis* relies on a contact-dependent type III secretion system (T3SS) to deliver a small number of proteins, termed effectors, through a hollow needle extending across the inner and outer membranes of the bacterium and into the cytosol of eukaryotic cells.^{3,4} These effector Yops (*Yersinia* outer proteins) interfere with signal transduction pathways that regulate the actin cytoskeleton, phagocytosis, apoptosis, and the inflammatory response, thereby favoring survival of the bacteria.⁵ More than 20 proteins are involved in the assembly of the syringe-like type III secretion “injectisome,” a so-called nanomachine.⁶ The structure of the injecti-

*Corresponding author. E-mail address:
waughd@ncifcrf.gov.

Abbreviations used: T3SS, type III secretion system; Yops, *Yersinia* outer proteins; TPR, tetratricopeptide repeat; CD, circular dichroism; EDTA, ethylenediaminetetraacetic acid; TEV, tobacco etch virus.

some is similar in some respects to that of the bacterial flagellum, confirming previous phylogenetic analyses that revealed substantial sequence similarities between several of their core components.⁷

The assembly and operation of the injectisome are assisted by a number of small bacterial proteins termed T3SS chaperones. Several different types of secretion chaperones have been recognized to date.⁸ Class I chaperones, which are small acidic polypeptides that exhibit little sequence similarity but adopt similar dimeric folds, generally interact with one unique effector (or, less frequently, with two or three effectors) and stimulate its secretion via the type III pathway. Class II chaperones, which interact with the translocator proteins of the T3SS that breach the plasma membrane of mammalian cells (e.g., *Yersinia* YopB and YopD), are somewhat larger and composed in part of tetratricopeptide repeats (TPRs). Flagellar chaperones, exemplified by FliS, have an entirely different structure⁹ and have therefore been designated Class III. Because the CesA chaperone that interacts with the EspA filament protein of the T3SS in enteropathogenic *Escherichia coli* is structurally distinct from the other classes,¹⁰ it has been designated Class IV. Finally, the unique structure of *Yersinia* YscE,¹¹ the cochaperone (along with YscG) for the needle protein YscF, prompted it to be categorized as yet another type of secretion chaperone (Class V).

The structural diversity of secretion chaperones may reflect their varying roles in type III secretion and flagellar assembly. It is believed that the principal function of Class I chaperones and CesA is to mediate targeting of their cognate binding partners to the type III secretion apparatus.¹² However, recent evidence suggests that Class I chaperones may also serve to mask aggregation-prone membrane localization domains in their cognate effectors.¹³ In addition to stimulating the secretion of their binding partners, at least some Class II chaperones also appear to play a regulatory role in type III secretion.^{14,15} The main function of the Class III chaperone FliS and the heterodimeric chaperone YscE (Class V)/YscG, on the other hand, evidently is to prevent premature polymerization of their cognate binding partners (FliC and YscF, respectively) prior to secretion.^{9,16}

The needles of the T3SSs in *Yersinia*, *Salmonella*, *Shigella*, *E. coli*, *Burkholderia*, and *Pseudomonas* are each composed primarily of a single protein (YscF, PrgI, MxiH, EscF, BsaL, and PscF, respectively) that polymerizes to form a hollow conduit^{17–22} with an inner diameter of 25 Å and a length of approximately 40–60 nm.³ Efforts to elucidate the three-dimensional structures of T3SS needle proteins have been hampered by their tendency to polymerize into oligomers of varying sizes *in vitro*. However, it was possible to create and crystallize monomeric forms of two needle proteins, MxiH²⁰ and BsaL,²¹ by deleting several residues from their C-terminus. The crystal structures of both MxiH and BsaL monomers revealed similar α -helical hairpin conformations.

Quinaud *et al.* demonstrated that in *Pseudomonas aeruginosa*, PscE and PscG trap the needle protein PscF in a ternary 1:1:1 complex, blocking the polymerization of PscF.¹⁶ An interaction between YscE and YscG, the orthologs of PscE and PscG in *Y. pestis*, had previously been identified by yeast two-hybrid technology, suggesting the existence of a similar complex in this organism.²³ To investigate the detailed molecular architecture of the complex between the cochaperones YscE/YscG and the *Yersinia* needle protein YscF, we crystallized the heterotrimer and determined its structure at a resolution of 1.8 Å.

Results and Discussion

Similarities and differences between the structures of YscE/YscF/YscG and PscE/PscF(55–85)/PscG

The crystal structure of the YscE/YscF/YscG multiprotein complex (YscEFG) was solved by the SAD (single-wavelength anomalous dispersion) method using incorporated selenomethionine (Table 1). The crystallographic asymmetric unit contains one heterotrimer, which is consistent with a 1:1:1 stoichiometry (Fig. 1). The final YscEFG model consists

Table 1. Data collection and refinement statistics

| Parameter | YscEFG complex ^a |
|---|----------------------------------|
| <i>Data collection</i> | |
| Space group | P2 ₁ 2 ₁ 2 |
| Cell dimensions | |
| <i>a</i> , <i>b</i> , <i>c</i> (Å) | 74.5, 94.5, 31.3 |
| Wavelength (Å) | 0.97942 |
| Resolution (Å) ^b | 50–1.8 (1.86–1.80) |
| <i>R</i> _{merge} ^b | 10.1 (36.9) |
| <i>I</i> / σ <i>I</i> ^b | 11.9 (4.3) |
| Completeness (%) ^b | 98.6 (95.1) |
| Redundancy ^b | 6.5 (4.6) |
| <i>Refinement statistics</i> | |
| Resolution (Å) | 50–1.8 |
| Unique reflections | 20,372 |
| <i>R</i> _{work} / <i>R</i> _{free} (%) | 19.03/22.13 |
| No. of atoms | |
| Protein | 1632 |
| Water | 227 |
| <i>B</i> -factors (Å ²) | |
| Protein ^c | 27.1/24.4/26.5 |
| Water | 40.5 |
| RMSDs | |
| Bond lengths (Å) | 0.01 |
| Bond angles (°) | 1.4 |
| Ramachandran plot statistics | |
| Residues in most favored regions | 177 |
| Residues in additional allowed regions | 7 |
| Residues in generously allowed regions | 0 |
| Residues in disallowed regions | 0 |

^a Se-Met protein was used in a SAD experiment for structure determination.

^b Values in parentheses refer to the highest-resolution shell.

^c Average *B*-factors for YscE, YscF, and YscG, respectively.

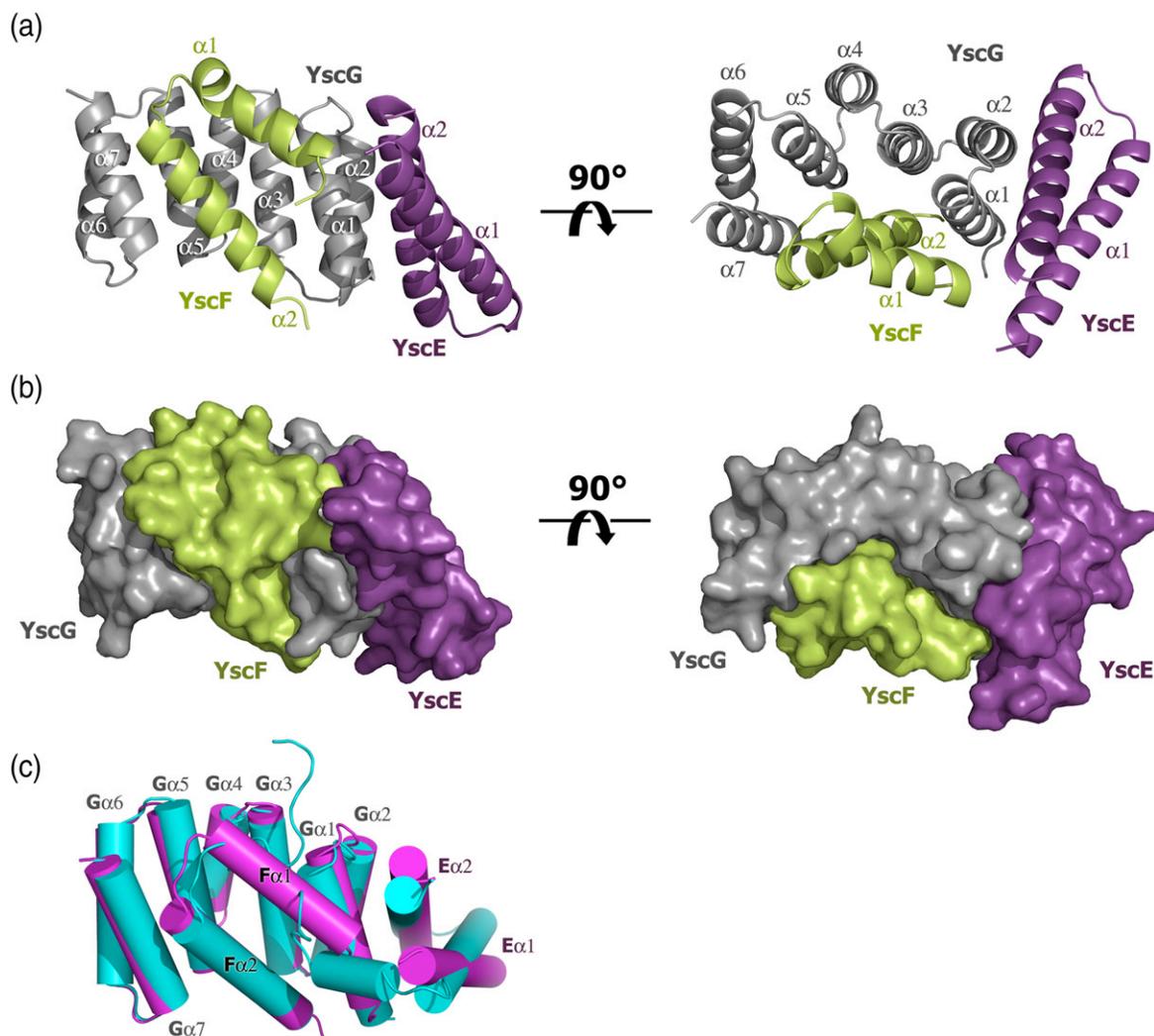


Fig. 1. (a) Ribbon representation of the overall structure of the YscEFG complex. YscE, YscF, and YscG are colored purple, lime, and gray, respectively. (b) Surface representation of the YscEFG complex. (c) Superposition of YscEFG (magenta) and PscEFG (cyan) complexes. α -Helices are depicted as cylinders and labeled according to YscEFG.

of 204 amino acids, including residues 10–63 of YscE, residues 50–87 of YscF, and residues 3–114 of YscG.

During the preparation of this article, the structure of the analogous PscEFG complex from the *P. aeruginosa* T3SS was reported.²⁴ The degree of sequence identity between YscE, YscF, YscG, and their respective counterparts in *P. aeruginosa* is 25%, 58%, and 53%. The overall protein sequence identity between YscEFG and PscEFG is 48%, while the RMSD between these two structures is about 6 Å. As one might expect, the two structures are similar in many respects (Fig. 1c), particularly the conformations of the YscG and PscG subunits, but there are also some significant differences between them.

The YscE subunit of the complex is composed of two α -helices, $\alpha 1$ (residues 10–33) and $\alpha 2$ (residues 37–60), which are connected by a three-residue loop, as reported previously.¹¹ RMSD values between C_{α} atoms in the structure of YscE in the YscEFG complex and that of the YscE dimer determined previously¹¹ range from 1.64 to 3.57 Å, with the

predominate differences being in the relative position of the two helices and the conformation of the loop between them.

The conformations of YscE and PscE in their respective complexes are also somewhat different. The first nine residues of YscE are disordered in the complex, whereas they form part of a contiguous α -helix in the crystal structure of YscE by itself.¹¹ In the structure of the PscEFG complex, the N-terminal residues of PscE adopt yet another conformation, this time in the form of an additional α -helix that angles toward the inner face of PscG. The biological significance of these different conformational states is unclear and may be an artifact of crystal packing since further truncation of the PscE polypeptide, to remove the additional N-terminal α -helix, had no impact on its ability to form a complex with PscF and PscG, nor did it lead to a defect in type III secretion.²⁴

Of the three pairs of homologous proteins in the two complexes, YscG and PscG share the greatest degree of structural similarity (RMSD=1.5 Å). The

crystal structure of the PscEFG complex revealed that PscG is a member of the TPR family of proteins,²⁴ a relationship that was not recognized by sequence comparisons.²⁵ The TPR motif is an imperfect 34-amino-acid repeat that commonly functions as a protein–protein interaction module and is found in proteins from all three domains of life.²⁶ Although the amino acid sequences of TPRs are not highly conserved, canonical positions at 8, 20, and 27 are usually occupied by small and sometimes hydrophobic residues. A sequence alignment between YscG and a canonical TPR protein PP5²⁷ (RMSD = 3.9 Å) reveals that the TPR domains in YscG (and PscG) are not strictly composed of 34 amino acids, but all three canonical positions are occupied by either alanine or leucine (Fig. 2).

Like PscG, YscG is composed of seven antiparallel α -helices that form a right-handed superhelix with a continuous helical groove. Its fold is similar to that of the TPR protein LcrH,²⁸ the Class II secretion chaperone for the translocators YopB, YopD, and LcrV, with an RMSD of 3.5 Å (Fig. 2). The solvent-exposed faces of helices α 3– α 7 of YscG are composed primarily of charged residues, whereas residues facing the interior of the protein are mainly nonpolar, generating a hydrophobic groove for YscF binding

(discussed below). The two N-terminal helices of YscG, α 1 and α 2, are packed against YscE through hydrophobic interactions. There is very little direct interaction between YscF and YscE in the complex. The total buried surface area between YscE/F, YscE/G, and YscF/G is 276, 1548, and 2724 Å², respectively.

YscG acts as a scaffold to organize the assembly of YscE and YscF into the heterotrimeric complex. Hydrophobic residues on the concave surface of YscG (Val7, Ala10, Leu14, Ala40, Leu43, Leu70, Pro72, Trp73, Leu76, Tyr79, Met109, and Phe105) interact with hydrophobic residues on YscF (Ile66, Ile71, Met75, Met78, Met79, Ile82, Leu83, and Phe86) (Fig. 3a). Similarly, hydrophobic interactions also mediate contact between YscG (Leu5, Leu8, Leu9, Ile12, Ile15, Ile28, and Trp31) and YscE (Ile17, Leu21, Ala24, Val43, Ala50, Leu51, Ala54, Ile57, Ile58, and Leu61) (Fig. 3b and c). An alignment of the amino acid sequence of YscG with those of related proteins reveals a pattern of conservation among these hydrophobic residues (Fig. 3d).

Crystal structures of Class I secretion chaperones bound to their cognate effectors have revealed that the chaperone-binding domains of the latter molecules are mostly nonglobular. Rather, they wrap around the surface of the chaperones in an extended

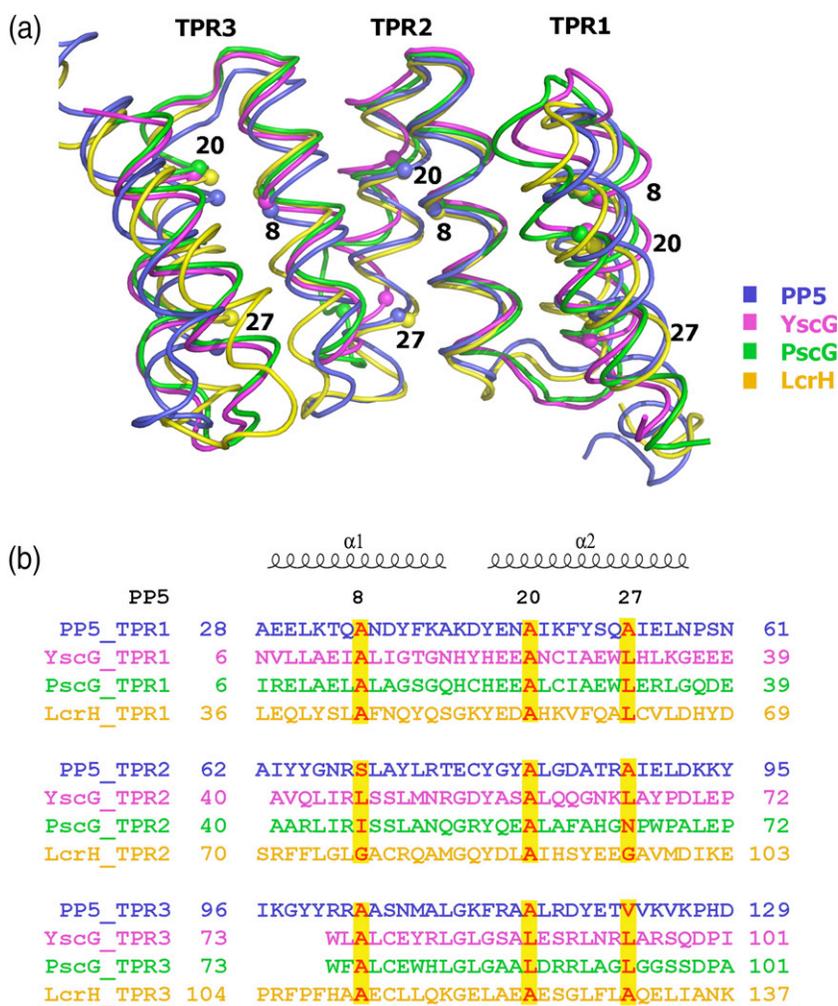


Fig. 2. Comparison of YscG with PP5 [Protein Data Bank (PDB) ID: 1A17], PscG (PDB ID: 2UWJ) and LcrH (PDB ID: 2VGJ). (a) Superposition of YscG (magenta), PP5 (blue), PscG (green), and LcrH (yellow) with canonical residues (8, 20, and 27) from each TPR domain represented as spheres. (b) Sequence alignment of three TPR domains in YscG, PP5, PscG, and LcrH. The three conserved positions are colored red and shaded yellow.

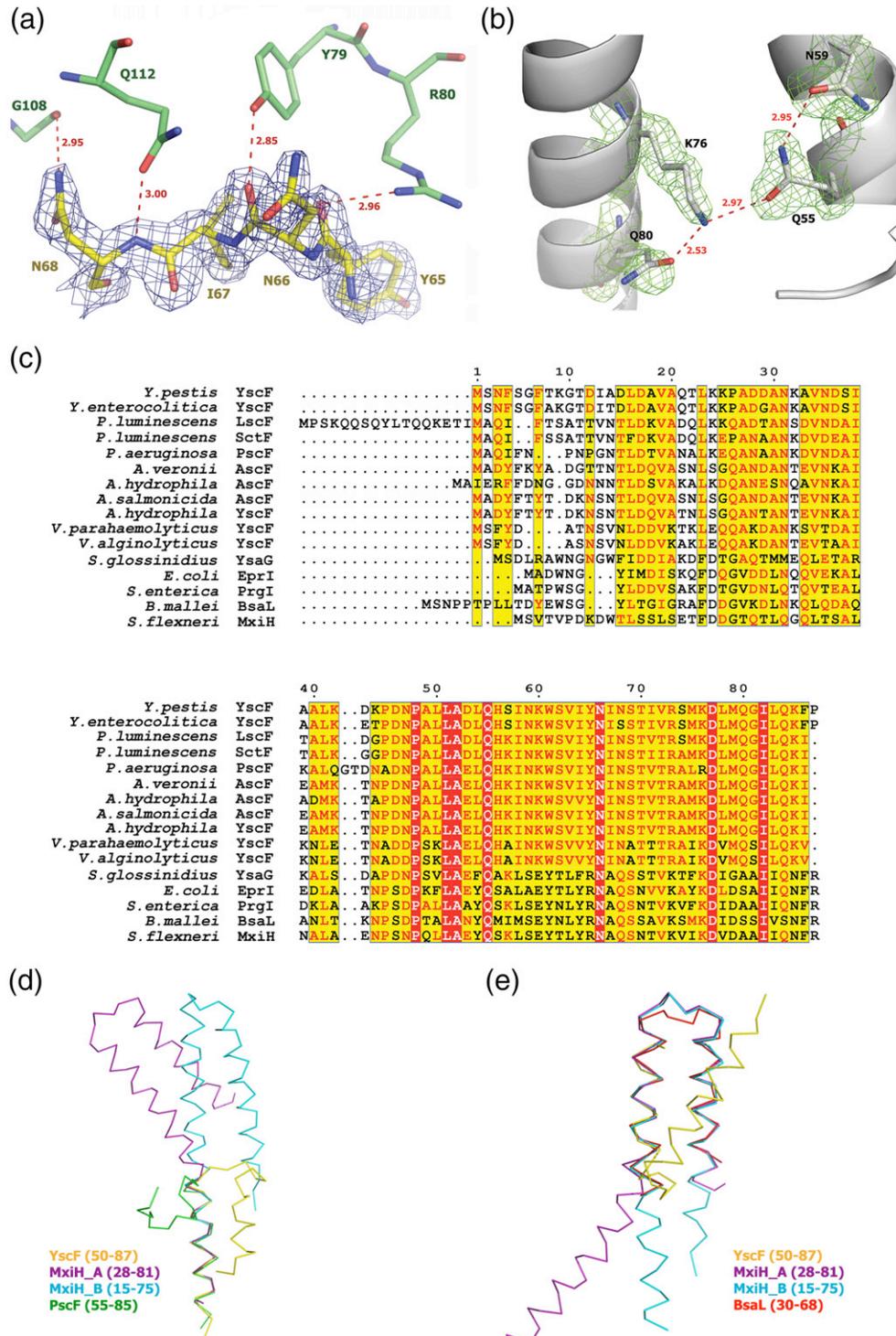


Fig. 4. YscF is captured in a partly helical, monomeric conformation. (a) $2F_o - F_c$ electron density for the loop region in YscF is contoured at 1.2σ . YscF and YscG are shown in yellow and green, respectively. Hydrogen bonds are represented as broken lines in red. (b) Hydrogen-bond network between $\alpha 1$ and $\alpha 2$ in YscF. Gln55, Asn59, Lys76, and Gln80 are depicted in gray ball-and-stick format. $2F_o - F_c$ electron density around these four residues is contoured at 1.0σ in green. Hydrogen bonds are represented as broken lines in red. (c) Sequence alignment of YscF with related open reading frames from other bacteria, with residues numbered according to YscF. Identical residues and highly similar residues are highlighted in red and yellow, respectively. (d) Superposition of MxiH chain A (purple), MxiH chain B (cyan), and PscF (green) with the C-terminal α -helix of YscF (yellow). Because the C-terminus of YscF does not have a stable conformation, it is not aligned. (e) Superposition of MxiH chain A (purple), MxiH chain B (cyan), and BsaL (red) onto the N-terminal α -helix of YscF (yellow). Because the corresponding residues of PscF were observed to be in an extended conformation, they were not included in the alignment.

proteins from *Shigella flexneri* (MxiH) and *Burkholderia pseudomallei* (BsaL), which were crystallized in isolation (Fig. 4d and e). The structure of MxiH (with four residues truncated from its C-terminus) consists of two extended helices connected by a short Pro-Ser-Asn-Pro turn.²⁰ The NMR structure of BsaL (with five residues deleted from its C-terminus) adopts a similar hairpin conformation with a much shorter C-terminal helix than MxiH.²¹ The structure of PscF in the PscEFG complex differs from that of the three other needle proteins in that it forms just a single α -helix (residues 68–85) that is similar to the C-terminal helix observed in YscF.²⁴ The region of PscF (residues 55–67) that is analogous to the N-terminal helix and hairpin turn observed in YscF adopts an extended conformation instead, with the N-terminal half of this region protruding away from the compact complex. The intramolecular hydrogen-bond “bridge” between highly conserved Gln55 and Lys76 in YscF may play an important role in maintaining the helical hairpin fold of YscF. We note that the residue equivalent to Gln55 in PscF was removed and replaced by a methionine during genetic engineering of the truncated polypeptide for crystallization purposes. The absence of the corresponding intramolecular hydrogen bond in PscF may be one of the reasons why its N-terminus was observed to be in an extended (i.e., nonhelical) conformation in the PscEFG complex. In addition, the chelate geometry of the Ni^{2+} ion that is bound to the N-terminus of PscF, which was introduced from the well solution during crystallization of the complex, is also worth considering. The Ni^{2+} ion is chelated by N of Met54,

N δ and N of His55 from PscF, and N δ of His80 from a symmetry-related PscG molecule in the crystal lattice (Fig. 5). Hence, the presence of this Ni^{2+} ion may also have had an impact on the conformation of the N-terminal region of the truncated PscF protein in the complex. Finally, we note that although several potentially favorable sites for thermolysin cleavage (hydrophobic residue in the P1' position) exist within the region of YscF and PscF that was observed to be in an extended conformation in the latter molecule, no fragments with these endpoints were detected after digestion of the YscEFG complex with thermolysin (see below). Taken together, these observations lead us to conclude that the α -helical hairpin conformation of YscF in the YscEFG complex may be a more accurate representation of the biologically relevant structure than is the conformation of PscF that is observed in its respective complex. It is also worth noting that whereas the manner in which the YscEFG complexes pack together in the $P2_12_12$ crystal lattice would allow enough room to accommodate the disordered N-terminal residues of YscF, the packing of PscEFG complexes in the $P6_222$ lattice would not. This may explain why it was not possible to obtain crystals of the full-length PscF protein in complex with PscE and PscG.

Biochemical and biophysical characterization of the YscEFG complex

Because they are not observed in the electron density maps of the YscEFG complex, either the 49

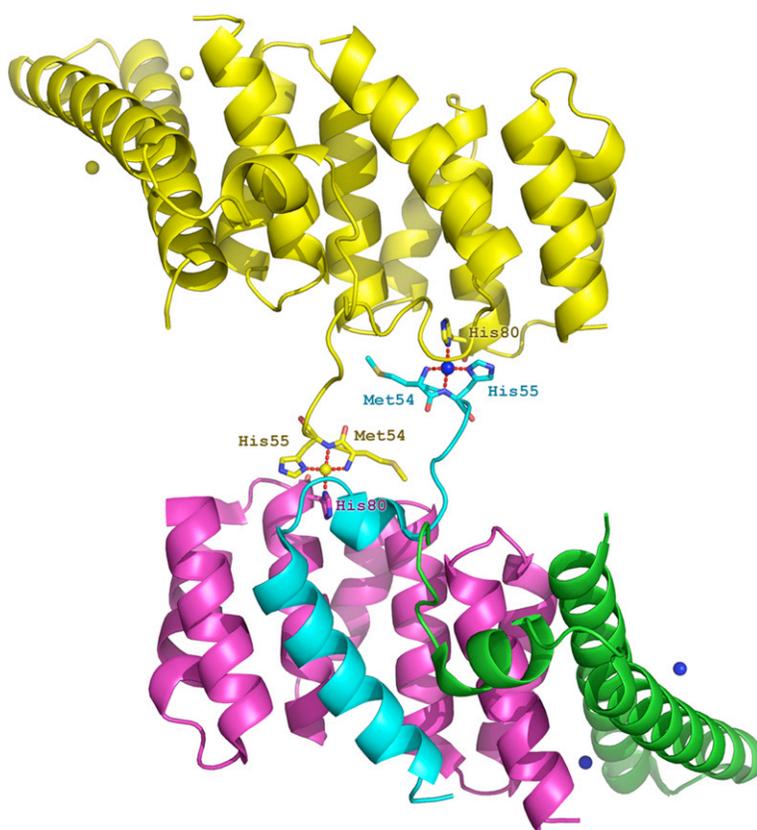


Fig. 5. Two symmetry-related PscEFG complexes are linked together by Ni^{2+} ions. The lower complex is colored with green (PscE), cyan (PscF), magenta (PscG), and blue (Ni^{2+} ions). The symmetry-related complex (upper) is shown in yellow. Metal-coordinating residues are represented in ball-and-stick format.

N-terminal residues of YscF are disordered in the structure of the heterotrimer or the N-terminal half of YscF is absent from the crystal due to proteolytic degradation. Several experiments were performed in an effort to distinguish between these possibilities.

First, proteins extracted from crystals of the complex were analyzed by electrospray ionization mass spectrometry. Although strong signals were obtained for YscG and YscE, no mass corresponding to YscF or a fragment of YscF could be detected, presumably due to poor ionization (data not shown). Crystals of the YscEFG complex were grown over the course of 10 days prior to freezing in liquid nitrogen for data collection, yet when the complex was incubated under crystallization conditions for 10 days and analyzed by SDS-PAGE, no degradation of YscF or the other two polypeptides was observed. Next, we subjected the complex to limited proteolysis using thermolysin (Fig. 6a). YscE and YscG proved to be remarkably resistant to proteolysis whereas YscF was rapidly degraded to yield a stable homogeneous fragment. This fragment was unambiguously identified as consisting of residues 52–87 of YscF by electrospray mass spectrometry and N-terminal amino acid sequencing (data not shown). It is unclear why this fragment of YscF could be readily detected by mass spectrometry whereas the full-length protein could not. A new expression vector was subsequently constructed to produce a truncated form of YscF corresponding to

the product of the thermolysin digest (residues 52–78), which was termed sYscF, and the ternary complex of sYscEFG was purified by a method similar to the one that was used to purify the full-length YscEFG complex. No crystals of the sYscEFG complex could be obtained. However, copurification of sYscF with YscE and YscG demonstrated that residues 52–87 of YscF are sufficient to bind tightly to YscEG (data not shown).

Next, we compared the far-UV circular dichroism (CD) spectra of the YscEFG and sYscEFG complexes (Fig. 6b). The two spectra were very similar, lending additional support to the hypothesis that the N-terminus of YscF is disordered in the YscEFG complex. Finally, we compared the stability of the two complexes by conducting thermal melting experiments monitored by CD (Fig. 6c). Both complexes exhibited virtually identical melting transition temperatures: 66.7 °C for YscEFG and 67 °C for sYscEFG. Taken together, these results strongly suggest that the 49 N-terminal residues of YscF are present but disordered in crystals of the YscEFG complex. Moreover, the manner in which the molecules pack in the crystal lattice creates gaps that could accommodate the disordered residues.

On the role of YscE in the YscEFG complex

YscE is required for the secretion of Yops in *Yersinia enterocolitica*³⁴ and *Y. pestis*,²³ yet virtually all of the

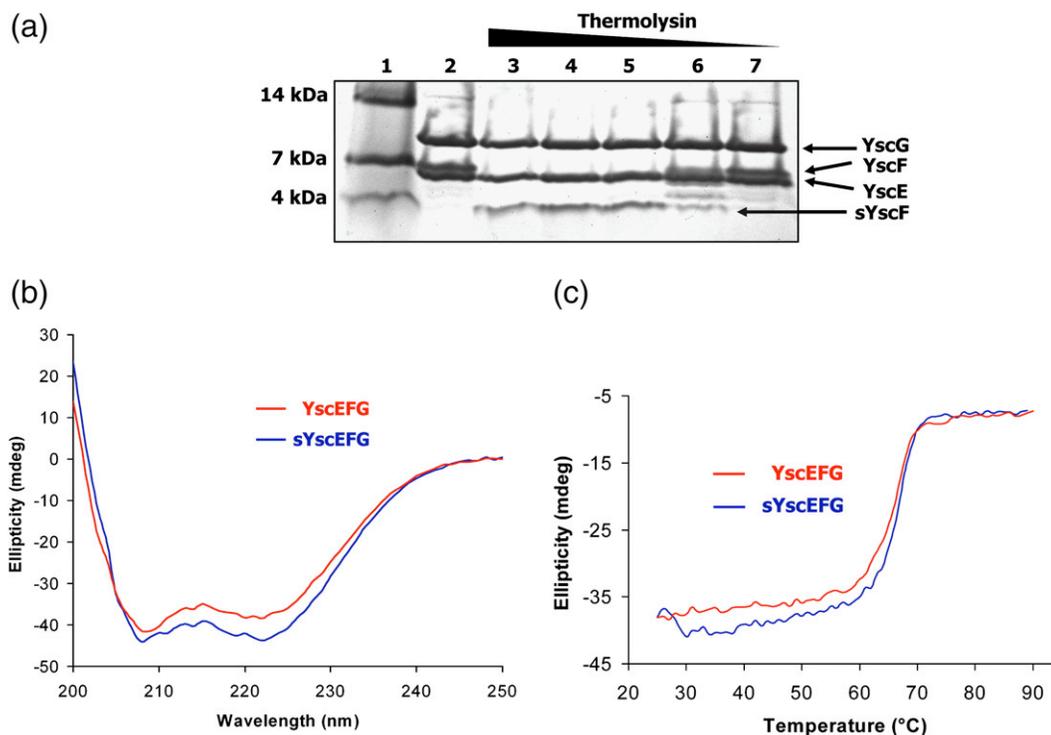


Fig. 6. Biochemical and biophysical characterization of the YscEFG complex. (a) Limited proteolysis of the YscEFG complex by thermolysin. Lane 1, molecular weight standards; lane 2, undigested YscEFG complex; lanes 3–7, YscEFG complex treated with 1:4, 1:16, 1:64, 1:256, and 1:1024 dilutions of the thermolysin stock solution, respectively. (b) CD spectra of YscEFG and sYscEFG complexes indicate that they have very similar α -helical content. (c) Melting curves of YscEFG and sYscEFG recorded at 222 nm.

intermolecular contacts with YscF in the YscEFG complex are mediated by the cochaperone YscG. It is not immediately obvious, therefore, what the function of YscE is. One possibility is that it serves to stabilize the fold of YscG, as PscE has been proposed to do for PscG.¹⁶ Consistent with this hypothesis, a *Y. pestis* mutant that lacks YscE fails to accumulate YscG.²³ A second possibility is that the role of YscE is to ensure that YscF is recognized by the translocation apparatus and assimilated into the growing needle structure at the proper time. However, there are no obvious patches of highly conserved residues on the surface of YscE that might mediate its interaction with components of the secretion apparatus.

The phenotypes of some mutations in YscF can be rationalized in terms of its interactions with YscG

Two studies have examined the impact of individual amino acid substitutions in YscF on type III secretion. In one study, a genetic screen was devised to identify mutations that are capable of secreting Yops but which are defective in translocating them into eukaryotic cells.³⁵ Because secretion requires a functional needle structure, which, in turn, depends on the ability of YscF to form a complex with its cochaperones YscE and YscG, none of these mutations would be expected to disrupt the YscEFG complex. Consistent with this notion, nearly all of the mutations that exhibit a translocation-defective phenotype and which are located in the portion of YscF that is visible in the electron density map of the complex occur at solvent-exposed positions (Ile67, Thr70, and Lys85). Only the side chain of the latter residue engages in any hydrogen-bonding interactions. Two mutations, I58V and S74G, resulted in a reduction in secreted YscF but were evidently still able to form functional needle structures. Unlike the other residues, the side chains at these two positions are partially buried in the complex and may therefore be mildly destabilizing.

In a second study, Torruellas *et al.* examined the effect of amino acid substitutions on the regulation of type III secretion by YscF.³⁶ Targeting conserved and acidic residues for replacement by alanine and/or cysteine, they were able to distinguish three phenotypes: (1) mutations with no discernable effect (i.e., those that exhibited normal, regulated secretion), (2) mutations that resulted in constitutive or unregulated secretion, and (3) mutations that abolished secretion altogether. None of the type 2 mutations (constitutive secretion) were located in the C-terminal portion of YscF that is visible in the complex with YscE and YscG. Once again, the majority of type 1 mutations (wild-type phenotype) located within the C-terminal half of YscF occurred at solvent-exposed positions (Asp53, Asn59, Ser69, Arg73, and Ser74). (Paradoxically, as discussed above, replacement of Ser74 with glycine reduced the secretion of YscF in the other study.) Alanine or cysteine substitutions at two positions (Asp77 and Ile82) abolished the secretion of YscF. The highly conserved Ile82 residue

is located at the interface between YscG and YscF (Fig. 3a), and the mutant may therefore destabilize the interaction between them. Asp77, on the other hand, is a solvent-accessible residue on the outer surface of YscF in the heterotrimeric complex. Therefore, the secretion defect exhibited by this mutant is not the result of the failure to properly interact with its chaperones YscE/G in cytosol. Instead, most likely, mutation in D77 will influence the architecture between needle monomers.

Evolutionary conservation of YscE, YscF, and YscG orthologs in T3SSs

The sequencing of microbial genomes has revealed the presence of T3SSs in a wide variety of animal and plant pathogens.³⁷ Most of these have open reading frames that exhibit significant sequence homology with YscF and other known T3SS needle proteins (Fig. 4c). However, only a handful of T3SSs possess recognizable orthologs of YscE and YscG (e.g., Fig. 3d). This suggests that those organisms with T3SSs, which clearly encode needle proteins but lack analogs of YscE and YscG, such as *S. flexneri* and *Salmonella typhimurium*, evidently employ a different mechanism to control premature polymerization of their needle proteins prior to assembly of the secretion apparatus.

Conclusions

The structural description of the YscEFG complex represents another incremental step forward in the ongoing effort to unravel the molecular mechanism of type III secretion. Yet, we are still a long way from understanding how different classes of T3SS substrates are recognized by the nascent secretion apparatus at the proper time and in the proper order. Clearly, the structures of additional multiprotein complexes, especially complexes between mobile and membrane-bound components of the T3SS, will be required to achieve this goal. The structure of the *Yersinia* needle protein YscF in complex with YscE and YscG is not the same as those of the monomeric needle protein variants crystallized in isolation,^{20,21} and the conformation of the needle proteins in the context of the polymerized needle structure is unlikely to be identical with either of these. The dynamic nature of the T3SS, a nanomachine composed of many moving parts, presents a formidable challenge for structural biology.

Methods

Protein expression and purification

A multicistronic expression vector encoding YscG and YscE was assembled by Gateway multisite recombinational cloning. Ribosome-binding sites and the appropriate attB recombination sites were added by PCR. YscG and YscE were inserted into pDONR208 and pDONR209,

respectively (Invitrogen, Carlsbad, CA, USA). The genes were sequence-verified and subsequently recombined into pDEST-HisMBP³⁸ to create the multicistronic expression vector pBA1578. The open reading frame encoding YscF (residues 2–87) was amplified from *Y. pestis* genomic DNA and inserted first into pDONR201 and then into the destination vector pDEST-3 (Invitrogen) to generate expression vector pYscF2. The recombinant proteins were expressed in *E. coli* BL21(DE3) CodonPlus-RIL cells (Stratagene, La Jolla, CA), which were induced at mid-log phase with 1 mM IPTG for 4 h. Cells containing the expression vector pBA1578 were resuspended in buffer A (50 mM sodium phosphate, pH 8.0, 150 mM NaCl, and 25 mM imidazole) with ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor cocktail tablets (Roche Molecular Biochemicals, Indianapolis, IN), lysed using an APV-1000 homogenizer (Invensys, Røhølsvej, Denmark) at 69 MPa, and centrifuged at 30,000g for 30 min at 4 °C. The supernatant was loaded onto a 20 mL His-Trap affinity column (GE Healthcare) and eluted with a linear gradient from 25 to 250 mM imidazole in buffer A. Fractions containing recombinant His₆-MBP-YscG/YscE were pooled, and the concentration of imidazole was adjusted to 25 mM using an Amicon stirred cell. The free YscE/YscG complex was recovered after digestion overnight with 5 mg His₆-tagged tobacco etch virus (TEV) protease at 4 °C. Next, a 20 mL His-Trap affinity column was employed to remove the His₆-MBP fusion partner, the His₆-tagged TEV protease, and residual undigested fusion protein. The concentrated flow-through fractions containing the YscE/YscG complex were subsequently loaded on a 320 mL XK26/60 Sephacryl S-100 sizing column (GE Healthcare) in buffer B (50 mM Tris, pH 7.2, and 150 mM NaCl).

The cell pellet containing expression vector pYscF2 was resuspended in buffer C (50 mM Tris, pH 7.2, and 200 mM NaCl) with EDTA-free protease inhibitor cocktail tablets, 1 mM benzamidine HCl, and 5 mM DTT and lysed using the same method as described above. The supernatant was applied to a GSTPrep FF 16/10 column (GE Healthcare). The column was washed with 10 column volumes of buffer C; 5 column volumes of 50 mM Tris, pH 7.2, and 2 M NaCl buffer; and 10 column volumes of buffer B. Next, a substantial molar excess of the YscE/YscG complex was applied to the GST-YscF-charged GSTPrep column, washed to baseline with buffer B, and treated overnight with His₆-tagged TEV protease. The YscEFG complex was released from the column by the action of TEV protease and recovered in the flow-through fraction. This material was concentrated and applied to a HiPrep 26/60 Sephacryl S-100 HR column equilibrated in buffer D [25 mM Tris, pH 7.2, 150 mM NaCl, 2 mM tris(2-carboxyethyl)phosphine]. The peak fractions containing the YscEFG complex were pooled and concentrated to 35 mg/mL. The final product was >95% pure by SDS-PAGE electrophoresis, and the molecular weights of the individual constituents of the YscEFG complex were confirmed by electrospray mass spectroscopy. The sYscEFG complex, which contains a truncated YscF protein (residues 52–78), was expressed and purified by a similar method. Selenomethionine was incorporated into the YscEFG complex by the method of Doublet.³⁹

CD spectroscopy

CD spectra were measured on an Aviv Model 202 spectrometer at 25 °C in a 1 mm cell in 25 mM Tris, pH 7.2, 150 mM NaCl, and 2 mM tris(2-carboxyethyl)phosphine at

protein concentrations of 0.35 mg/mL for both YscEFG and sYscEFG samples. The thermodynamic stability was recorded at 222 nm by monitoring the CD signal at 25–90 °C with a scan rate of 1 °C/min for increasing temperature assay. The spectra were corrected against those of the buffer reference.

Limited proteolysis analysis of YscEFG complex

A 1 mg/mL stock solution of thermolysin (Roche Molecular Biochemicals) in 1× thermolysin buffer (10 mM Tris-HCl, pH 8.0, 2 mM CaCl₂, 150 mM NaCl, and 5% glycerol) was used for the limited proteolysis experiments. The YscEFG stock solution contains the protein at 2 mg/mL in buffer D. The five individual reactions were composed of 25 µL of YscEFG stock solution, 25 µL of 2× thermolysin buffer, and 0.5 µL of serial 1:4 dilutions of the thermolysin stock solution. The reactions were allowed to proceed for 1 h at 37 °C until termination by adding 2 µL of 0.5 M EDTA. The reaction products were initially analyzed by SDS-PAGE. The precise molecular weights of the fragments were measured by electrospray mass spectrometry.

Crystallization and data collection

Crystals of the YscEFG complex were grown at 18 °C using sitting drop vapor diffusion with protein to well solution (0.2 M ammonium fluoride and 20% polyethylene glycol 3350) at 1:3 (v/v). Crystals appeared within 3 days. The crystals belong to the space group *P2₁2₁2*. The solvent content of the crystal was estimated to be approximately 33% (v/v) with a Matthews coefficient (V_M) of 1.8 Å³ Da⁻¹ for one heterotrimer in the asymmetric unit. The unit cell has dimensions of $a=74.50$ Å, $b=94.54$ Å, and $c=31.30$ Å. Crystals were equilibrated in a cryoprotectant composed of reservoir buffer plus 10% (v/v) glycerol and were flash-frozen in a cold nitrogen stream at -170 °C. One peak data set was collected at the SER-CAT beamline 22-ID (Advanced Photon Source, Argonne National Laboratory). Data were processed and scaled using the HKL2000 program suite.⁴⁰ Data collection and processing statistics are summarized in Table 1.

Structure determination and refinement

The structure of Se-YscEFG was solved initially by the Se-SAD method at 1.8 Å. Initial phases were obtained at 1.8 Å with a figure of merit of 0.261 using the program CNS.⁴¹ Eight of the 10 selenium atoms per asymmetric unit were located. Phases were improved by density modification using CNS,⁴¹ and this led to a figure of merit of 0.799. An experimental map was then generated for model building. Nearly the entire backbone and most of the side chains could be traced in the program O.⁴² The model was completed by alternating between manual building and refinement with the program CNS. Residues 1–9 and 64–66 of YscE, residues 2–49 of YscF, and residues 1–2 and 115 of YscG were not visible in the electron density. The final 1.8 Å refined structure of the YscEFG complex consists of 204 amino acid residues and 227 water molecules. Model quality was verified with PROCHECK.⁴³ All nonglycine residues reside either in the most favorable regions or in the allowed regions of the Ramachandran plot. Model refinement statistics are listed in Table 1. All figures were generated with the graphics program PyMOL.⁴⁴

Coordinates

The atomic coordinates and structure factors for the YscEFG complex were deposited in the Research Collaboratory for Structural Bioinformatics PDB with accession code 2P58.

Acknowledgements

We thank Dr. Sreedevi Nallamsetty, Dr. George Lountos, and Mi Li for technical assistance and helpful discussions. Electrospray mass spectrometry experiments were conducted on the liquid chromatography/electrospray mass spectrometry instrument maintained by the Biophysics Resource in the Structural Biophysics Laboratory, Center for Cancer Research, National Cancer Institute at Frederick. X-ray diffraction data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at <http://www.ser-cat.org/members.html>. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract no. W-31-109-Eng-38.

References

- Rollins, S. E., Rollins, S. M. & Ryan, E. T. (2003). *Yersinia pestis* and the plague. *Am. J. Clin. Pathol.* **119**, S78–S85.
- Ligon, B. L. (2006). Plague: a review of its history and potential as a biological weapon. *Semin. Pediatr. Infect. Dis.* **17**, 161–170.
- Cornelis, G. R. (2006). The type III secretion injectosome. *Nat. Rev., Microbiol.* **4**, 811–825.
- Galan, J. E. & Wolf-Watz, H. (2006). Protein delivery into eukaryotic cells by type III secretion machines. *Nature*, **444**, 567–573.
- Navarro, L., Alto, N. M. & Dixon, J. E. (2005). Functions of the *Yersinia* effector proteins in inhibiting host immune responses. *Curr. Opin. Microbiol.* **8**, 21–27.
- Mota, L. J. & Cornelis, G. R. (2005). The bacterial injection kit: type III secretion systems. *Ann. Med.* **37**, 234–249.
- Aizawa, S. I. (2001). Bacterial flagella and type III secretion systems. *FEMS Microbiol. Lett.* **202**, 157–164.
- Willharm, G., Dittmann, S., Schmid, A. & Heesemann, J. (2007). On the role of specific chaperones, the specific ATPase, and the proton motive force in type III secretion. *Int. J. Med. Microbiol.* **297**, 27–36.
- Evdokimov, A. G., Phan, J., Tropea, J. E., Routzahn, K. M., Peters, H. K., Pokross, M. & Waugh, D. S. (2003). Similar modes of polypeptide recognition by export chaperones in flagellar biosynthesis and type III secretion. *Nat. Struct. Biol.* **10**, 789–793.
- Yip, C. K., Finlay, B. B. & Strynadka, N. C. (2005). Structural characterization of a type III secretion system filament protein in complex with its chaperone. *Nat. Struct. Mol. Biol.* **12**, 75–81.
- Phan, J., Austin, B. P. & Waugh, D. S. (2005). Crystal structure of the *Yersinia* type III secretion protein YscE. *Protein Sci.* **14**, 2759–2763.
- Lee, S. H. & Galan, J. E. (2004). Salmonella type III secretion-associated chaperones confer secretion-pathway specificity. *Mol. Microbiol.* **51**, 483–495.
- Letzelter, M., Sorg, I., Mota, L. J., Meyer, S., Stalder, J., Feldman, M. *et al.* (2006). The discovery of SycO highlights a new function for type III secretion effector chaperones. *EMBO J.* **25**, 3223–3233.
- Broms, J. E., Edqvist, P. J., Carlsson, K. E., Forsberg, A. & Francis, M. S. (2005). Mapping of a YscY binding domain within the LcrH chaperone that is required for regulation of *Yersinia* type III secretion. *J. Bacteriol.* **187**, 7738–7752.
- Francis, M. S., Lloyd, S. A. & Wolf-Watz, H. (2001). The type III secretion chaperone LcrH co-operates with YopD to establish a negative, regulatory loop for control of Yop synthesis in *Yersinia pseudotuberculosis*. *Mol. Microbiol.* **42**, 1075–1093.
- Quinaud, M., Chabert, J., Faudry, E., Neumann, E., Lemaire, D., Pastor, A. *et al.* (2005). The PscE–PscF–PscG complex controls type III secretion needle biogenesis in *Pseudomonas aeruginosa*. *J. Biol. Chem.* **280**, 36293–36300.
- Hoiczky, E. & Blobel, G. (2001). Polymerization of a single protein of the pathogen *Yersinia enterocolitica* into needles punctures eukaryotic cells. *Proc. Natl Acad. Sci. USA*, **98**, 4669–4674.
- Kubori, T., Sukhan, A., Aizawa, S. I. & Galan, J. E. (2000). Molecular characterization and assembly of the needle complex of the *Salmonella typhimurium* type III protein secretion system. *Proc. Natl Acad. Sci. USA*, **97**, 10225–10230.
- Pastor, A., Chabert, J., Louwagie, M., Garin, J. & Attree, I. (2005). PscF is a major component of the *Pseudomonas aeruginosa* type III secretion needle. *FEMS Microbiol. Lett.* **253**, 95–101.
- Deane, J. E., Roversi, P., Cordes, F. S., Johnson, S., Kenjale, R., Daniell, S. *et al.* (2006). Molecular model of a type III secretion system needle: implications for host-cell sensing. *Proc. Natl Acad. Sci. USA*, **103**, 12529–12533.
- Zhang, L., Wang, Y., Picking, W. L., Picking, W. D. & De Guzman, R. N. (2006). Solution structure of monomeric BsaL, the type III secretion needle protein of *Burkholderia pseudomallei*. *J. Mol. Biol.* **359**, 322–330.
- Wilson, R. K., Shaw, R. K., Daniell, S., Knutton, S. & Frankel, G. (2001). Role of EscF, a putative needle complex protein, in the type III protein translocation system of enteropathogenic *Escherichia coli*. *Cell. Microbiol.* **3**, 753–762.
- Day, J. B., Guller, I. & Plano, G. V. (2000). *Yersinia pestis* YscG protein is a Syc-like chaperone that directly binds yscE. *Infect. Immun.* **68**, 6466–6471.
- Quinaud, M., Ple, S., Job, V., Contreras-Martel, C., Simorre, J. P., Attree, I. & Dessen, A. (2007). Structure of the heterotrimeric complex that regulates type III secretion needle formation. *Proc. Natl Acad. Sci. USA*, **104**, 7803–7808.
- Pallen, M. J., Francis, M. S. & Futterer, K. (2003). Tetratricopeptide-like repeats in type-III-secretion chaperones and regulators. *FEMS Microbiol. Lett.* **223**, 53–60.
- D'Andrea, L. D. & Regan, L. (2003). TPR proteins: the versatile helix. *Trends Biochem. Sci.* **28**, 655–662.
- Das, A. K., Cohen, P. W. & Barford, D. (1998). The structure of the tetratricopeptide repeats of protein

- phosphatase 5: implications for TPR-mediated protein-protein interactions. *EMBO J.* **17**, 1192–1199.
28. Buttner, C. R., Sorg, I., Cornelis, G. R., Heinz, D. W. & Niemann, H. H. (2007). Structure of the *Yersinia enterocolitica* type III secretion translocator chaperone SycD. *J. Mol. Biol.* **375**, 997–1012.
 29. Birtalan, S. C., Phillips, R. M. & Ghosh, P. (2002). Three-dimensional secretion signals in chaperone-effector complexes of bacterial pathogens. *Mol. Cell*, **9**, 971–980.
 30. Phan, J., Tropea, J. E. & Waugh, D. S. (2004). Structure of the *Yersinia pestis* type III secretion chaperone SycH in complex with a stable fragment of YscM2. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 1591–1599.
 31. Lilic, M., Vujanac, M. & Stebbins, C. E. (2006). A common structural motif in the binding of virulence factors to bacterial secretion chaperones. *Mol. Cell*, **21**, 653–664.
 32. Stebbins, C. E. & Galan, J. E. (2001). Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion. *Nature*, **414**, 77–81.
 33. Schubot, F. D., Jackson, M. W., Penrose, K. J., Cherry, S., Tropea, J. E., Plano, G. V. & Waugh, D. S. (2005). Three-dimensional structure of a macromolecular assembly that regulates type III secretion in *Yersinia pestis*. *J. Mol. Biol.* **346**, 1147–1161.
 34. Allaoui, A., Schulte, R. & Cornelis, G. R. (1995). Mutational analysis of the *Yersinia enterocolitica* virC operon: characterization of yscE, F, G, I, J, K required for Yop secretion and yscH encoding YopR. *Mol. Microbiol.* **18**, 343–355.
 35. Davis, A. J. & Mecsas, J. (2007). Mutations in the *Yersinia pseudotuberculosis* type III secretion system needle protein, YscF, that specifically abrogate effector translocation into host cells. *J. Bacteriol.* **189**, 83–97.
 36. Torruellas, J., Jackson, M. W., Pennock, J. W. & Plano, G. V. (2005). The *Yersinia pestis* type III secretion needle plays a role in the regulation of Yop secretion. *Mol. Microbiol.* **57**, 1719–1733.
 37. Troisfontaines, P. & Cornelis, G. R. (2005). Type III secretion: more systems than you think. *Physiology (Bethesda)*, **20**, 326–339.
 38. Nallamsetty, S. & Waugh, D. S. (2006). Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expr. Purif.* **45**, 175–182.
 39. Doublet, S. (1997). Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.* **276**, 523–530.
 40. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
 41. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **54**, 905–921.
 42. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **47**, 110–119.
 43. Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* **231**, 1049–1067.
 44. DeLano, W. L. (2002). *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.