The Modular Structure of the Inner-Membrane Ring Component PrgK Facilitates Assembly of the Type III Secretion System Basal Body

Julien R.C. Bergeron,1,2 Liam J. Worrall,1,2 Soumya De,1 Nikolaos G. Sgourakis,3 Adrienne H. Cheung,1 Emilie Lameignere,1,2 Mark Okon,1 Gregory A. Wasney,1,2 David Baker,4 Lawrence P. McIntosh,1,5 and Natalie C.J. Strynadka1,2,*

1Department of Biochemistry and Molecular Biology, The University of British Columbia, 2350 Health Sciences Mall, Vancouver, BC, V6T 1Z3, Canada
2Centre for Blood Research, University of British Columbia, 2350 Health Sciences Mall, Vancouver, BC, V6T 1Z3, Canada
3National Institute of Diabetes & Digestive & Kidney Diseases, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA
4Department of Biochemistry, Department of Genome Sciences, and Howard Hughes Medical Institute, University of Washington, 4000 15th Avenue NE, Seattle, WA 98195, USA
5Department of Chemistry, University of British Columbia, 2350 Health Sciences Mall, Vancouver BC, V6T 1Z3, Canada
*Correspondence: ncjs@mail.ubc.ca
http://dx.doi.org/10.1016/j.str.2014.10.021

SUMMARY

The type III secretion system (T3SS) is a large macromolecular assembly found at the surface of many pathogenic Gram-negative bacteria. Its role is to inject toxic “effector” proteins into the cells of infected organisms. The molecular details of the assembly of this large, multimembrane-spanning complex remain poorly understood. Here, we report structural, biochemical, and functional analyses of PrgK, an inner-membrane component of the prototypical Salmonella typhimurium T3SS. We have obtained the atomic structures of the two ring building globular domains and show that the C-terminal transmembrane helix is not essential for assembly and secretion. We also demonstrate that structural rearrangement of the two PrgK globular domains, driven by an interconnecting linker region, may promote oligomerization into ring structures. Finally, we used electron microscopy-guided symmetry modeling to propose a structural model for the intimately associated PrgH-PrgK ring interaction within the assembled basal body.

INTRODUCTION

Salmonella is a Gram-negative bacterium, several strains of which are human pathogens. S. typhimurium is a major source of food-borne enterocolitis, whereas S. typhi is the etiological agent for typhoid fever, a disease that remains endemic in the developing world (Haraga et al., 2008). A common feature of all pathogenic strains of Salmonella is the presence of Salmonella Pathogenicity Islands (SPIs) in their genome. Specifically, two of these (SPI-1 and SPI-2) encode for type III secretion systems (T3SS), large macromolecular assemblies responsible for the injection of toxic “effector” proteins into the cytosol of infected cells (Coburn et al., 2007a; de Jong et al., 2012). T3SSs have been identified as essential pathogenicity components in many infectious Gram-negative bacteria, including Salmonella spp., enteropathogenic E. coli (EPEC) and enterohemorrhagic E. coli, Pseudomonas aeruginosa, Yersinia spp., Shigella spp., Chlamydia spp., and Vibrio spp. (Coburn et al., 2007b; Troistfontaines and Cornielis, 2005). This renders the T3SS a very attractive target for the development of new antibiotics and vaccines (Keyser et al., 2008; Kline et al., 2012; Marshall and Finlay, 2014).

A T3SS consists of a series of conserved genes coding for the secretion apparatus (“injectisome”) and of a set of genes coding for the effector proteins, which vary between systems and species (Böttner, 2012; Kosarewicz et al., 2012). Effector proteins are channeled through the injectisome to the cytosol of target cells (Dohlich et al., 2014; Radics et al., 2014). A number of groups have employed an integrative approach (Alber et al., 2008) to obtain structural models of the prototypical Salmonella SPI-1 T3SS injectisome. Electron microscopy (EM) studies have revealed its global organization (Marlovits et al., 2004; Schraadt et al., 2010; Schraadt and Marlovits, 2011), and the atomic structures of several isolated domains have been reported (Bergeron et al., 2013; Lunelli et al., 2011; Spreter et al., 2009; Wang et al., 2007). Cross-linking in combination with mass spectrometry was employed to obtain information on the interaction between the various components (Sanowar et al., 2009; Schraadt et al., 2010). Finally, computational methods have been developed to combine these data in order to obtain an atomic model of the T3SS injectisome architecture (Bergeron et al., 2013; Demers et al., 2013).

However, the mechanistic details for the assembly of this large complex (~3.5 MDa) remain poorly understood. Genetic, biochemical, and EM data have demonstrated that the inner-membrane components of the injectisome can assemble independently, with the proteins PrgK and PrgH (Figure 1A) forming intimately nested 24-mer ring structures around the inner-membrane components SpaPQRS (Kimbrough and Miller, 2000; Wagner et al., 2010). Biochemical data also showed that in parallel the secretin InvG forms a pore in the outer membrane (Figure 1A), facilitated by the action of the pilin lipoprotein InvH.
RESULTS

Elucidating the Roles of PrgK’s Membrane-Embedded Domains

PrgK is an inner-membrane component of the *Salmonella* SPI-1 T3SS basal body (Kimborough and Miller, 2000) (Figure 1A). At its N terminus, PrgK possesses a canonical lipoprotein signal sequence, with a conserved cysteine residue (Cys 18) forming a predicted site of lipidation (Juncker et al., 2003). Lipidation has been experimentally confirmed in vivo for the closely related *Shigella* ortholog MxiJ (44% sequence identity) (Allaoui et al., 1992). In addition, a C-terminal transmembrane (TM) helix is predicted in all PrgK orthologs, except the EPEC T3SS component EscJ (Figure 1B; Figure S1A available online). Nanogold labeling experiments have shown that the C terminus of PrgK is located in the cytoplasm, confirming that PrgK traverses the inner membrane (Schraidt et al., 2010).

To assess the role of the putative N-terminal lipidation and C-terminal TM helix for basal body assembly, we monitored secretion of effector proteins in a *S. typhimurium* strain containing a chromosomal deletion of the *prgK* gene and complemented with plasmids encoding WT *prgK* or mutants. Surprisingly, removal of the C-terminal 52 residues, which include the cytoplasmic tail and TM helix (PrgK1–200), does not alter effector secretion (Figure 1C). This observation demonstrates that the TM helix of PrgK is not essential for inner-membrane localization or basal body assembly. In contrast, secretion is abrogated when the putative N-terminal lipidation site is mutated (C18A), suggesting a requirement for the specific anchoring of PrgK to the inner membrane via lipidation of Cys 18.

Following from the above observations, we were able to purify intact needle complexes from *Salmonella* strains lacking the TM helix of PrgK, as observed by negative-stain EM (Figures S1B and S1C). We note that the yield of complexes obtained with the PrgK1–200 truncation is approximately 10-fold lower than with the WT PrgK. This suggests that the TM helix does contribute to the overall stability of the needle complex.

Structures of the PrgK Globular Domains

Existing homology models of the PrgK ring structure (Sanowar et al., 2009; Schraidt and Marlovits, 2011; Worrall et al., 2011) are based on the experimental observation of a superhelical array of 24-mer containing successive turns found in the crystal structure of the EPEC ortholog, EscJ (26% identity) (Yip et al., 2005). However, we showed previously that ring models of basal body domains based on the structures of distant homologs lead to inaccuracies, including domain misorientation and clashes, and interfered with model convergence in our EM-guided symmetry modeling procedure (Bergeron et al., 2013). Therefore, in order to obtain a more accurate PrgK ring model, we experimentally determined the structure of the PrgK monomer was required. To this end, we expressed a recombinant construct of its periplasmic domain lacking the lipidation site (PrgK19–200). This protein also show that a linker region of PrgK may regulate oligomerization through multiple interactions with the flanking globular domains. Based on these results, we propose a molecular model for the stepwise assembly of the PrgK–PrgH inner-membrane ring.

See also Figure S1.
could be purified to high yield, but underwent proteolytic degradation at 20°C (hours) and 4°C (days), leading to two products of approximately 14 and 9 kDa size, respectively (data not shown). Crystal and solution structures of EPEC EscJ (Crepin et al., 2005; Yip et al., 2005) have revealed that this homologous protein possesses two globular domains (7 and 10 kDa) joined by an extended linker (Figures 1B and S1A). We therefore purified the two corresponding globular domains of PrgK independently. We could not obtain crystals for the N-terminal domain (PrgK19–76, henceforth referred to as D1). However, this protein yielded high-quality nuclear magnetic resonance (NMR) spectra (Figure 2A), allowing us to solve its solution structure (Figure 2B; Table 1). Further, we were successful in obtaining crystals of the C-terminal domain (PrgK98–200, henceforth D2) and could solve its structure by X-ray crystallography (Figure 2C; Table 2; see Supplemental Experimental Procedures for details).

Modeling of the PrgK 24-mer Ring
We next applied our NMR and crystallographic structures to an EM-guided symmetry modeling protocol (Bergeron et al., 2013) to model the PrgK 24-mer ring oligomer. Starting from the monomeric structures of D1 and D2, we applied the procedure to the two domains independently (Figure S3; Supplemental Experimental Procedures for details), allowing us to obtain a refined model for PrgK (Figure 3A). Notably, despite this independent starting set, the final model is globally similar in intermolecular packing to the intact EscJ ring model derived from the crystallographically determined superhelical structure (Figures S3C and S3D), with both domains in each of these orthologs adopting similar orientations. We further note that the diameter of the D2 ring is larger in the PrgK model, in agreement with the dimension of the region of density assigned to PrgK in the EM map. It has been reported that the EPEC T3SS possesses a narrower basal body compared with the Salmonella SPI-1 T3SS (Sekiya et al., 2001), suggesting that the larger PrgK D2 ring may correspond to structural differences between the two systems.

We then observed that in the Rosetta-based PrgK model, the D1 and D2 domains are in close proximity, forming a number of notable direct interactions, including between residues 70–75 at the surface of helix $\alpha_2$ of D1, and residues 111–126 on $\alpha_4$ and strand $\beta_3$ of D2 (Figure 3B). This prompted us to investigate whether the D1 and D2 domains of PrgK interact in vitro. Using NMR spectroscopy, we monitored spectral perturbations in the $^{15}$N-labeled D1 domain upon titration with unlabeled D2. As shown in Figure 3C, we could observe a number of amide $^{1}H$-$^{15}$N with progressively perturbed chemical shifts, implicating residues 70–74 of helix $\alpha_2$ and to a lesser extent, those in the spatially adjacent helix $\alpha_1$ (residues 38–42). These are largely hydrophobic residues, which cluster on the surface of D1 formed...
by the two helices of the domain (Figure 3D) and match with the interface generated in the PrgK 24-mer model (Figure 3B). In parallel, a weak endothermic reaction was observed when we used isothermal titration calorimetry (ITC) to monitor the titration of D1 against D2 (Figure 3E). However, over the available concentration range, we were unable to saturate the ITC- and NMR-monitored titration, indicating that it is a weak interaction with an affinity likely in the millimolar range (Figure 3C). Nevertheless, we observed that this interaction occurs between adjacent molecules, with D1 of molecule i interacting with D2 of molecule i+1 along the 24-mer oligomeric ring (we refer to this arrangement as a “domain-swapped” conformation). This is likely driven by the linker region, which is well ordered and forms a number of interactions with D2 (Figure 4B), but not with D1 or with linker regions from adjacent molecules. As we have modeled the two domains D1 and D2 of PrgK independently, we cannot directly distinguish whether the analogous D1-D2 interaction occurs in an intermolecular complex, as shown using size-exclusion chromatography (SEC-MALS) (Figure 5A, brown curve), whereas in the absence of the linker, D2 is monomeric (purple curve). Importantly, the fold of D2 is likely globally conserved in the PrgK82–200 construct, as shown by CD spectroscopy (Figure S4A), supporting that this oligomer is not constituted of unstructured, aggregated protein. A closer analysis of the EscJ superhelix structure revealed that a conserved Phe at position 89 in the linker region docks into a hydrophobic pocket formed by two adjacent D2 subunits (Figure 5B). We therefore

we observed that this interaction occurs between adjacent molecules, with D1 of molecule i interacting with D2 of molecule i+1 along the 24-mer oligomeric ring (we refer to this arrangement as a “domain-swapped” conformation). This is likely driven by the linker region, which is well ordered and forms a number of interactions with D2 (Figure 4B), but not with D1 or with linker regions from adjacent molecules. As we have modeled the two domains D1 and D2 of PrgK independently, we cannot directly distinguish whether the analogous D1-D2 interaction occurs in an intermolecular complex, as shown using size-exclusion chromatography (SEC-MALS) (Figure 5A, brown curve), whereas in the absence of the linker, D2 is monomeric (purple curve). Importantly, the fold of D2 is likely globally conserved in the PrgK82–200 construct, as shown by CD spectroscopy (Figure S4A), supporting that this oligomer is not constituted of unstructured, aggregated protein. A closer analysis of the EscJ superhelix structure revealed that a conserved Phe at position 89 in the linker region docks into a hydrophobic pocket formed by two adjacent D2 subunits (Figure 5B). We therefore

The Linker Region of PrgK Promotes Oligomerization

We further noted that the residues involved in the D1-D2 interaction identified in PrgK (Figure 3) are also present in the EscJ superhelix crystal structure (Figure 4A). However, in the latter,
hypothesized that this residue may also stabilize the PrgK82–200 oligomer. Indeed, we observed that a mutation of Phe89 to Ala abrogates the oligomerization of PrgK82–200 (Figure 5A, green curve) in vitro, confirming that the linker promotes D2 oligomerization in a similar fashion to that observed in the EscJ superhelix structure. We note that this mutation does not compromise T3SS secretion in vivo, nor does the Y70A mutation (which disrupts the oligomeric interaction between D1 and D2, see above). However, the double-mutant F89A/Y70A does abrogate secretion (Figure S4C), which suggests that PrgK oligomerization is likely stabilized by a number of interactions, including those involving residues Phe89 and Tyr70. The NMR spectrum of PrgK19–200 with the Y70A/F89A mutations (Figure S4D) confirmed that the overall fold of the protein is conserved, and its stability was confirmed by differential fluorescence calorimetry (Figure S4B). Further structural characterization of this protein was not possible due to its propensity to undergo degradation (see above).

Finally, we used negative-stain EM to analyze the PrgK82–200 oligomer and observed that it forms ring-shaped/tubular structures, with a conserved diameter of ~15–20 nm (Figure 5C). This diameter is similar to that observed in both the EscJ superhelix structure (Yip et al., 2005) and in the density attributed to the PrgK ring in the high-resolution cryo-EM map of the basal body (Schraidt and Marlovits, 2011). However, the structures vary in length from a few nm to over 100 nm (Figure 5C). Based on these observations, we propose that these filaments consist of a superhelical assembly of PrgK D2, which putatively forms in the absence of the membrane-embedded N-terminal lipidation that would restrict its assembly to a 2D plane. We note that the heterogeneous nature of these structures prevented any high-resolution characterization, as would be required to formally demonstrate their biological significance.

From these results, we conclude that in the EPEC EscJ and Salmonella PrgK inner-membrane rings D1 and D2 likely form a domain-swapped interface between adjacent molecules. We hypothesize that this allows for a more extensive set of interactions within the 24-mer rings, including with the linker region, to promote oligomerization.

Interaction between D1 and Linker Prevents Oligomerization
We have demonstrated that the linker is sufficient to promote oligomerization of D2 in vitro, in the absence of D1. However, no oligomerization was observed for the PrgK19–200 construct, which includes D1, D2, and the linker (Figure 5A, blue curve). This suggested the possibility that D1 may prevent spontaneous
and potentially premature oligomerization of PrgK in the inner membrane.

To validate this hypothesis, we engineered a protein construct including D1 followed by the linker region. Using ITC, we observed that the D1-D2 interaction is abrogated in the context of the D1-linker construct. Unsurprisingly, we noted that many residues yield two distinct sets of NMR signals, indicative of two conformations in slow exchange, referred to as “population A” and “population B,” with the peaks yielded by population A consistently more intense than those yielded by population B. We mapped chemical shift differences between these two populations along the sequence, which inferred that structural differences occur primarily in the α helices and the following linker. However, the chemical shift-derived secondary structure prediction suggests that population A and population B possess similar secondary structure elements. The Random Coil Index also reveals that residues 70–82 (end of helix 2 and beginning of the linker) are less flexible in population A than in population B. We note that most peaks found in the spectrum of PrgK D1 overlay onto the peaks assigned to population B of PrgK19–92, indicating that in population B the linker does not interact significantly with the D1 domain. These results suggest that in population A, residues 76–82 of the linker interact with D1, while this interaction is not present in population B. Importantly, the spectrum of PrgK19–92 overlay well onto that of the construct that included both PrgK domains (PrgK19–200), with the two peaks formed by population A and population B clearly observable in the larger construct. This demonstrates that the interaction between D1 and the linker is also present in the purified, monomeric full-length protein (lacking the N-terminal lipidation site and C-terminal TM region), which could not be characterized further due to its instability (see above).

Next, we were able to solve the crystal structure of PrgK19–92 (Figure 6D; Table 2; see Supplemental Experimental Procedures for details). In this structure, the linker (residues 75–82) is well ordered and clearly located in the D2-binding groove of D1, in agreement with the NMR results described above. Notably Tyr 75, Leu 77, and Pro 78 form hydrophobic interactions with loop 1, helix α1, and helix α2 of D1. We note that only the linker-bound conformation, corresponding to population A, was found in the crystal. It is also noteworthy that this protein forms a superhelical arrangement in the crystal lattice, in which intermolecular contacts support the D1 ring model reported above. Together, these results demonstrate that D1 interacts with the linker, through the groove formed in between the two helices of the domain. This corresponds to the surface of D1 that interacts with D2, as shown by our modeling and NMR results (see Figures 6C and 3D). From this observation, we conclude that the linker interacts with the D2 binding pocket of D1, thereby sterically blocking the D1-D2 interaction. Furthermore, by sequestering the linker, D1 likely also prevents spontaneous oligomerization of PrgK.

To provide additional support for this hypothesis, we used NMR spectroscopy to monitor the titration of PrgK D1 with an isolated polypeptide corresponding to the linker sequence (PrgK76–97) fused to SUMO. As summarized in Figure S5E, a number of D1 residues showed amide chemical shift changes upon peptide addition. These residues clustered around the helical regions of D1 (Figure S5G). This confirms that the linker binds the same region of D1, whether present as a separate peptide or in the intramolecular context of PrgK19–92. Furthermore, chemical shift analysis (Figure S5F) allowed us to obtain a Kd value of 100 μM (±40 μM) for the intermolecular D1-linker interaction. This is at least an order of magnitude tighter than the intramolecular D1-D2 interaction. The preferential D1-linker interaction is consistent with the observations by NMR and SEC-MALS that PrgK19–200 is monomeric.

Structural Basis for the PrgH-PrgK Interaction

We next combined the 24-mer models of PrgK D1 and D2 described above (Figure 3A), with the PrgH oligomeric model reported previously (Bergeron et al., 2013), and applied the ring

![Image](image_url)
modeling procedure on the collective set (see Figures S6A and S6B; Supplemental Experimental Procedures for details). From this we obtained a model for the PrgH-PrgK 48-mer ring complex (Figure 7A), which agrees with the EM map (Figure S6C), and previously published biochemical data (Table S1). We emphasize that the model includes a number of assumptions, such that the subunits are related by strict symmetry, and that the structures of individual domains are not significantly modified in the assembled complex. Higher resolution EM maps, crystal structures of oligomerized domains, and additional mass spectrometry/crosslinking and mutagenesis interaction data should allow for generation of models with greater accuracy in the future. In particular, while the agreement with the NMR titration and EscJ crystallographic contacts supports the PrgK model, additional experimental restraints would be necessary to better define the PrgH-PrgK interaction.

In this model, we observe that there are predicted contacts between the periplasmic domains of PrgH and PrgK monomers (Figure 7B). However, this intermolecular interface is relatively limited (collectively 1,440 Å³ of buried surface for the modeled complex) and with the major interactions occurring between the C-terminal region of PrgH and PrgK D2. This is in agreement with the highest resolution (~10 Å) EM map (Schraidt and Marlovits, 2011), in which a large cavity is observed between the sections of the density attributed to PrgH and PrgK (Figure 7B). Specifically, in the PrgH-PrgK model, the conserved Asp 333 (located on helix 2 in the third ring-building motif domain of PrgH) is buried in a well-ordered, positively charged pocket formed by helix α4 of two adjacent PrgK D2 subunits (Figure 7C). Consistent with this arrangement, we were not able to detect an interaction between the periplasmic domain of PrgH and PrgK in vitro (data not shown), likely because this construct does not form the PrgH binding pocket in its monomeric state. The highly oligomeric and heterogeneous state of the PrgK82–200 construct did not allow for further investigation of its interaction with PrgH. As shown in Figure 7D, mutation of Asp 333 to Arg in PrgH abrogates secretion (without affecting the structure of the isolated domain; Figures S6D and S6E), supporting the essential role of this residue. In PrgK, mutation of the conserved Lys 168 to Glu does not alter secretion, but it is abrogated when both Lys 168 and Arg 169 are mutated to Ala. CD spectroscopy and differential scanning calorimetry (DSC) suggest that PrgK D2 is largely folded in the presence of these mutations (Figures S6F and S6G). We also observe that the loop located between strand β5 and helix α5 of D2, which is poorly ordered in the PrgK crystal structure (see Supplemental Experimental Procedures), is located in the cavity described above (Figure 7B) in our PrgK ring model. Neither this loop nor the extended loop between strands β4 and β5 is accounted for by EM map density. It is therefore possible that flexible regions of PrgK could potentially form additional contacts with PrgH upon complex formation.

DISCUSSION

In this study, we characterize structural and functional features of the modular ring building domains of the inner-membrane
protein PrgK, from the prototypical Salmonella SPI-1 T3SS injectisome. We demonstrate that its C-terminal TM helix is not essential for injectisome assembly/effector secretion and report the crystallographic and NMR structures of its two globular domains in isolation. Using our previously developed molecular modeling method combining atomic structures of individual domains, EM maps, and symmetry, we have obtained a model of the PrgK periplasmic ring. Further analysis in the context of its binding partner PrgH provides a starting point for understanding the intimate molecular interaction between these two ring-forming proteins. Finally, we propose that the linker region between the two globular ring-building domains of PrgK may play a regulatory role in oligomerization, by promoting an intermolecular, domain-swapped disposition of D1 and D2 that stabilizes the 24-mer ring complex.

Previous studies have shown that coexpression of PrgK and PrgH leads to the spontaneous formation of the inner-membrane ring structures. In contrast, expression of PrgK or PrgH alone forms oligomers, but lacks the ring-like conformation observed in the fully assembled injectisome (Kimbrough and Miller, 2000). In this light, the data reported here suggest a molecular mechanism for the assembly of the PrgH-PrgK rings. We propose that PrgK can adopt two conformations: an assembly-incompetent conformation with D1 sequestering the linker and an assembly-competent conformation where the linker promotes oligomerization. This conformation is likely stabilized by the insertion of a PrgH molecule in between two adjacent PrgK.

Figure 6. D1 Sequesters the Linker to Prevent Oligomerization

(A) ITC isotherms for the titration of PrgK D1 against D2. A schematic representation of the experiments performed is shown on the right. In the presence of the linker region fused to D1 (PrgK19–92), the two domains do not interact. The top is a repeat of the experiment shown on Figure 3E (left).

(B) Assigned 15N-HSQC spectrum of PrgK19–92 (insert for the crowded region). Black lines connect peaks arising from the same residue in the two populations (A and B) of this protein. The asterisk indicates peaks assigned to the two cloning remnant residues at the N terminus of the PrgK sequence.

(C) The weighted combined amide 15N and 1H chemical shift differences (Δ(ΔH, N)) between corresponding peaks for population A and population B is plotted along the protein sequence (top). The perturbed amides cluster in helices α1, α2 and the linker region. Secondary structures for the populations A and B of PrgK19–92, obtained from their assigned 1H, 15N, 13Cα, and 13Cβ chemical shifts using the program SSP (Marsh et al., 2006), is shown in the middle. Scores of +1, 0, and −1 correspond to helices, random coils, and strands, respectively. The Random Coil Index profiles for the populations A and B of PrgK19–92, obtained from their assigned 1H, 15N, 13Cα, and 13Cβ chemical shifts using the RCI server (Berjanskii and Wishart, 2007), are shown at the bottom. Residues 72–84 are less disordered in population A than in population B.

(D) Crystal structure of PrgK19–92, in ribbon representation, with corresponding sequence numbering shown. Secondary structure elements are labeled as in Figure S1A, and blue to red rainbow coloring indicates N- to C-terminal directionality (see Table 2 for statistics).

(E) Close-up view of the linker region (in red) interaction with D1 (green). The composite 2Fo-Fc omit map (0.05% atoms omitted, generated by Phenix) is shown, contoured at 1 σ around residues 75–82. Residues that form hydrophobic interactions between D1 and the linker are shown as sticks. See also Figure S5.
molecules, which would explain why PrgK or PrgH alone do not spontaneously form ring-like oligomers. Propagation of this ring-initiation step leads to the formation of a stable PrgH-PrgK 24-mer ring pair (Figure 8). We argue that such a coordinated oligomerization of these two conserved luer lock rings would be essential to allow their formation only upon encompassing the several inner-membrane spanning export apparatus proteins that they are presumed to contain (Wagner et al., 2010).

We acknowledge that this proposed mechanism is derived largely from the biochemical and structural behavior of a soluble, monomeric fragment of PrgK and may therefore not reflect the behavior of the full-length membrane-embedded protein in vivo. However, the correlation between the impact of PrgK mutations in vitro and in vivo, as well as the ability to isolate a PrgH/PrgK complex (Kimbrough and Miller, 2000), supports the validity of the proposed model. Although perhaps intuitively sensible that a large complex such as the PrgK-PrgH oligomer would assemble in a coordinated, step-wise manner, additional experimental validation, such as the isolation and molecular characterization of assembly intermediates, will be required to confirm the model proposed in Figure 8.

A second aspect of this study is the proposed model for the PrgH-PrgK interaction from our Rosetta-based analysis. Perhaps surprising is the somewhat limited nature of this interface that, although bolstered by stoichiometry, suggests additional functionalities, such as the lipidated and/or membrane-spanning regions of PrgH and PrgK, or other structural components in the periplasmic region. Further, both the EM map density and our symmetry modeling suggest the presence of a large cavity between PrgH and PrgK (Figure 7), with few interactions occurring in the periplasmic region (although it is possible that in situ this cavity is occupied by other components of the injectisome that perhaps are lost during needle complex purification in the EM analysis). This cavity could...
Structural and Functional Characterization of PrgK

Figure 8. Molecular Model for the Assembly of the PrgH-PrgK Ring
(A) We propose that initially PrgK exists in a monomeric state, while localized in the inner membrane via nature of the N-terminal lipidation site and the C-terminal TM helix. D1 interacts with the linker, preventing oligomerization. The C-terminal TM helix, which is not essential for assembly (as shown in Figure 1C) is omitted for clarity.
(B) The linker then dissociates from D1, exposing the Tyr 70-containing hydrophobic pocket, as well as Phe 89.
(C) PrgK then recruits adjacent molecules via a range of intermolecular interactions involving both domains and the linker region, indicated by dotted black circles.
(D) PrgH molecules insert in the pocket formed by two adjacent PrgK molecules, forming a heterotrimeric intermediate.
(E) PrgH-PrgK heterotrimers further oligomerize to obtain an assembled PrgH-PrgK complex.

allow for a degree of structural plasticity in the basal body, as observed recently in situ for the Yersinia T3SS (Kudryashev et al., 2013). The presence of a cavity into which solvent can diffuse between PrgH and PrgK could also have favorable implications in the design of T3SS assembly inhibitors.

In conclusion, in this study, we have obtained the structures of isolated domains of PrgK and used interaction studies and computational methods to propose a structural model for the PrgH-PrgK 48-mer periplasmic rings. This provides insights into the interaction between PrgH and PrgK. We also report biochemical, structural, and functional data suggesting a stepwise assembly for the PrgH-PrgK complex, promoted by structural rearrangement in PrgK. While integrative structural biology approaches enable the piecing together of the architecture of large macromolecular assemblies, the addition of a temporal dimension, including assembly, disassembly, and/or functional changes, such as that probed here, is an important further element toward understanding and targeting complex nanomachines such as the T3SS at the molecular level.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification
All constructs were cloned in a PET28a plasmid (Novagen), with an N-terminal His10-tag fusion followed by a thrombin cleavage site. Plasmids were transformed into E. coli BL21(DE3) cells, and transformants were grown to log phase at 37°C. Expression was induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside at 20°C for 16 hr. Harvested cells were resuspended in lysis buffer (50 mM HEPES [pH 8.0], 150 mM NaCl) with protease inhibitors (Roche). Cells were lysed by sonication. Clarified lysate was run through Zn-chelating sepharose, and the proteins were eluted with 500 mM imidazole. The His-tag was cleaved with Thrombin (Roche), and proteins were further purified by gel filtration with a Superdex 75 column (GE Healthcare).

For NMR experiments, 15N- and 13C-labeled proteins were expressed in E. coli BL21(DE3) in M9 minimal media supplemented with 1 g/l 15NH4Cl and 161–172, January 6, 2015

NMR Spectra Acquisition, Assignment, and Structure Determination
For all NMR experiments, protein samples were dialyzed into 50 mM HEPES (pH 6.8), 10% D2O, and concentrated to ~0.5–1 mM. Standard 2D and 3D spectra for backbone and side-chain assignments were collected using either a 600 or 850 MHz Brucker Avance III spectrometer, equipped with TCI cryoprobe. All spectra were processed using NMRPipe (Delaglio et al., 1995) and analyzed with SPARKY (T.D. Goddard and D.G. Kneller, University of California, San Francisco) (see Supplemental Experimental Procedures for details).

Crystallization and X-Ray Crystallographic Structure Determination
Crystals of PrgK19–200 and PrgK19–200 were obtained in 100 mM sodium acetate (pH 5.5), 20% polyethylene glycol (PEG) 6000, 50 mM NaCl, 50 mM MgCl2, and 80 mM phosphate buffer (pH 4.9), 20 mM Tris pH 7.0, 25% PEG 300, 20 mM MgCl2, 20 mM NaCl, respectively, by vapor diffusion, using the sitting drop method. X-ray diffraction data were collected at the Canadian Light Source beamline 08B1-1. Phases were obtained by molecular replacement with the program Phaser (McCoy et al., 2007). The PrgK19–200 structure was refined to 2.65 Å resolution and the PrgK19–200 to 3.2 Å resolution (Table 2; see Supplemental Experimental Procedures for details).

EM-Guided Symmetrical Modeling
Ring modeling for PrgK D1 and PrgK D2 was performed as described previously (Bergeron et al., 2013; see Supplemental Experimental Procedures for details).

For the PrgK D1-D2-PrgH model, individual lowest energy ring models of PrgK D1, PrgK D2, and PrgH were combined into a single coordinate file and used to generate the input for the phase II all atom refinement procedure (see Supplemental Experimental Procedures for details).

In Vivo Assays
Salmonella typhimurium LT2 strains containing a deletion of the prgH or prgK gene were complemented with plasmids containing the corresponding gene or mutants, all with a C-terminal 6xHis tag, which does not affect T3SS function (Schraud et al., 2010), and effector protein secretion was monitored as described previously (Kimbrough and Miller, 2000). Needle complexes were purified as described previously (Bergeron et al., 2013; Kubori et al., 1998; Schraud et al., 2010).

Electron Microscopy
Samples of PrgK19–200 were diluted to ~1 mg/ml in 50 mM HEPES (pH 7.0). 150 mM NaCl. Samples of purified needle complex particles were diluted to ~0.2 mg/ml in 10 mM Tris (pH 8.0), 500 mM NaCl, 5 mM EDTA, and 10 mM lauryldimethylamine oxide. All samples were applied on glow-discharged carbon grids and stained using 0.75% uranyl formate. Images were collected on a Tecnai G2 transmission electron microscope (FEI) operating at 200 kV and equipped with a high-speed AMT 2K side-mount CCD camera.
Structure

Structural and Functional Characterization of PrgK

ACCESSION NUMBERS

The atomic coordinates for the various structures have been deposited in the PDB ID (2MKY, 4OYC, and 4W4M). The lowest all-atom energy PrgH-PrgK Rosetta model has been deposited in the PDB (3J6D).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.str.2014.10.021.

AUTHOR CONTRIBUTIONS

J.R.C.B. cloned, expressed, and purified all proteins. J.R.C.B. crystallized and solved the crystal structures, with help from L.J.W. for the refinement. J.R.C.B. collected all NMR data with help from M.O. and analyzed them with help from S.D. and L.P.M. J.R.C.B. performed the ITC and circular dichroism experiments with help from L.J.W. J.R.C.B. performed the modeling experiments with N.G.S. and D.B. E.L. and G.A.W. performed the DLS and DSF experiments, respectively. J.R.C.B. performed all in vivo and EM assays with help from A.H.C. J.R.C.B. and N.C.J.S. wrote the manuscript, with comments from all coauthors.

ACKNOWLEDGMENTS

We thank Samuel Miller for providing the S. typhimurium prgK- and fliC/prgK-deletion strains. We are also grateful to Kelvin Lau, Bradford Ross, Gerd Preinha, Matthew Solomonson, and Fred Rossel for technical advice. X-ray diffraction data were collected at beamline 6BB1-1 at the Canadian Light Source, which is supported by the Natural Sciences and Engineering Research Council of Canada, the National Research Council Canada, the Canadian Institutes of Health Research, the Province of Saskatchewan, Western Economic Diversification Canada, and the University of Saskatchewan. Instrument support was provided by the Canada Foundation for Innovation, the British Columbia Knowledge Development Fund, the UBC Blusson Fund, and the Michael Smith Foundation for Health Research. N.G.S. acknowledges funding by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases, NIH. J.R.C.B. is a MSFHR postdoctoral fellow. N.C.J.S. is a Canada Research Chair Tier 1 in Antibiotic Discovery. This work was supported by operating grants from the Canadian Institute of Health Research and HHMI International Scholar Program (to N.C.J.S.).

REFERENCES


Supplemental Information

The Modular Structure of the Inner-Membrane Ring Component PrgK Facilitates Assembly of the Type III Secretion System Basal Body

Julien R.C. Bergeron, Liam J. Worrall, Soumya De, Nikolaos G. Sgourakis, Adrienne H. Cheung, Emilie Lameignere, Mark Okon, Gregory A. Wasney, David Baker, Lawrence P. McIntosh, and Natalie C.J. Strynadka
SUPPLEMENTAL FIGURES:

A

Signal sequence

Lipidation

Linker

TM helix

B

WT  1-200  C18A

InvG

PrgH

PrgK

α-His

C

WT

1-200

1-200

C18A

WT

1-200

C18A
Figure S1, related to figure 1: Impact of PrgK TM domain mutants on injectisome assembly

(A) The sequences of *Salmonella typhimurium* SPI-1 PrgK (PrgK_SalSPI1), EHEC ETT2 EprK (EprK_ETT2), *Shigella flexneri* MxiJ (MxiJ_Shi), *Yersinia enterocolica* YscJ (YscJ_Yer), *Pseudomonas aeruginosa* PscJ (PscJ_Pse), EPEC LEE EscJ (EscJ_LEE) and *Salmonella typhimurium* SPI-2 SsaJ (SsaJ_SalSPI2) were aligned using ClustalW. The resulting alignment figure was generated with ESPript. Conserved residues are in red boxes, similar residues in red characters. The secondary structure of PrgK and EscJ are shown at the top (green) and bottom (blue), respectively. The location of the signal sequence (for export to the inner-membrane), the lipidation signal, the C-terminal transmembrane (TM) helix, and the linker region are indicated. (B) SDS-PAGE gel of purified needle complexes containing PrgK (WT, left lane), a deletion of PrgK’s TM helix (PrgK1-200, middle lane), or a mutation of the lipidation site (C18A, right lane). Bands corresponding to the basal body components InvG, PrgH and PrgK are indicated, along with those of known contaminants (*) often found in needle complex purification procedures. A western blot of the corresponding samples, done with anti-His antibody to detect C-terminally labeled PrgK (PrgK1-200 is not tagged in this experiment), is shown at the bottom. This confirms that the C18A mutation still allows for protein expression and membrane localization. (C) Negative-stain EM micrographs of purified needle complexes containing various PrgK species: WT (top panel), PrgK1-200 (middle panel), or C18A (bottom panel). Fully assembled complexes, found in the WT construct and the PrgK1-200 truncation, but not the C18A mutant, are indicated with a red arrow.
**Figure S2, related to figure 2: Overlay of the PrgK and EscJ structures.**
The structures of PrgK$_{19-76}$ (D1) and PrgK$_{96-200}$ (D2) in green, are overlayed onto the corresponding domains in an EscJ monomer (PDB: 1YJ7, in grey).
**Figure S3, related to figure 3: Illustration of the PrgK ring modeling**

(A) Results of the fixed-backbone symmetry docking step for D1 (left) and D2 (right). The energy of each model is plotted versus the RMSD to backbone of the lowest-energy model, and color-coded for the fit to the EM map density (EMDB-1875). The two obtained clusters for D1, indicated as Cl-1 and Cl-2, are shown in green in (B), with the position of the D2 subunit in cyan. The residues of D1 most perturbed in the NMR-monitored titration with D2 (Figure 3D, E) are in sticks. These are consistent with the models in Cl-2. (C) The PrgK 24-mer models for PrgK D1 (from Cl-2) and D2 are shown in green, from the top and from the side, overlaid to the EscJ ring model (in grey) derived from the crystallographic super-helix. (D) A single monomer for the PrgK D1 and PrgK D2 ring models are overlayed onto an EscJ monomer, with colors as in (C). Both D1 and D2 are in similar orientation, although the D2 ring is slightly larger in PrgK, as supported by the EM map. In addition, the position of D1 is slightly rotated in the PrgK model compared to EscJ.
Figure S4

A) spectra of PrgK<sub>96-200</sub>, PrgK<sub>92-200</sub>, and PrgK<sub>92-200</sub> F89A

B) Guanidine titration of PrgK<sub>19-200</sub> and PrgK<sub>19-200</sub> Y70A/F89A

C) SDS-PAGE of SipA (WT, 1-92, F89A, Y70A, Y70A/F89A) and FliC

D) 1<sup>5</sup>N-H HSQC spectra of PrgK<sub>19-200</sub> and PrgK<sub>19-200</sub> Y70A/F89A
Figure S4, related to figure 5: Functional and biochemical characterization of the PrgK Y70A/F89A mutant

(A) CD spectra of purified PrgKD2 (blue), D2 with the linker region (red) or D2 with the linker region including the F89A mutation (green). Ellipticity [θ] is plotted in deg cm² dmol⁻¹. All three proteins have similar secondary structure, confirming that the mutation does not alter the overall fold of the domain. We note a slightly higher β-strand content for the D2 + linker construct, likely reflecting the oligomeric state of this protein. (B) Midpoint unfolding temperature (T_m) of PrgK_{19-200} WT (blue) and Y70A/F89A mutant (red), plotted against increasing concentration of guanidinium-HCl. Both proteins show similar T_m values and sensitivity to guanidinium-induced denaturation, as measured by differential scanning fluorimetry. The Tm value was set to 0 when no unfolding transition was observed. (C) SDS-PAGE gel of proteins secreted by *S. typhimurium* strains containing mutations of PrgK. FliC is used as a loading control. SipA is secreted when the wild-type PrgK is expressed, but not with a deletion of PrgK D2 (PrgK_{1-92}). Mutation of Tyr 70 or Phe 89 to alanine, which disrupts the D1-D2 interaction (Figure 3E), or PrgK_{82-200} oligomerization (Figure 5A) respectively, do not affect secretion. However, mutating both residues results in secretion-incompetent strains. This is consistent with a cumulative role for these two interfaces in *vivo*. (D) ¹⁵N-HSQC spectrum of PrgK_{19-200} (blue) overlaid on that of the Y70A/F89A mutant (red). The majority of the amide peaks have the same chemical shifts, confirming the overall fold is not affected by the mutations. We note that some peaks appear shifted, suggesting localized perturbations.
Figure S5

**E**

- A 2D contour map with labeled residues:
  - Thr 74
  - Ile 72
  - Trp 71
  - His 42
  - Leu 39
  - Met 41
  - Tyr 70
  - Val 38
  - Lys 73

**F**

- A graph showing the ratio of PrgK_{19-74} : PrgK_{76-97} vs. Δ(H,N) for various residues:
  - L39
  - Q40
  - H42
  - W71
  - I72
  - K73

**G**

- A histogram showing Δ(H,N) (ppm) for residue number 19 to 76:
**Figure S5, related to figure 6: The linker region of PrgK interacts with D1**

(A) $^{15}$N-HSQC spectrum of PrgK$_{19-92}$ (blue), overlaid on that of PrgK$_{19-76}$ (red). The linker region induces numerous changes in the NMR spectrum of PrgK$_{19-92}$ relative to PrgK$_{19-76}$, with many residues in the larger construct yielding two signals (Figure 4A). Of these, signals from PrgK$_{19-76}$ overlap with those of the B population of PrgK$_{19-92}$. (B) $^{15}$N-HSQC spectrum of PrgK$_{19-92}$ (blue), overlaid on the spectrum of PrgK$_{19-200}$ (black). Most peaks from PrgK$_{19-92}$ align well with corresponding peaks in the PrgK$_{19-200}$ spectrum. In particular, in most cases, the two populations identified for PrgK$_{19-92}$ are also present in the PrgK$_{19-200}$ construct. (C) Molecular packing of the PrgK$_{19-96}$ crystal structure, shown in ribbon representation, with four adjacent superhelices in red, blue, orange and pink. The superhelical arrangement extends along the entire length of the crystal. (D) Adjacent molecules from the PrgK$_{19-96}$ superhelix (blue), overlayed onto the PrgK D1 ring model reported on figure 3A (green). In both contexts, the same interface is used for oligomerization, providing additional support for the PrgK D1 ring model (Figure 3A). (E) Section of the $^{15}$N-HSQC spectra of PrgK D1 alone (red), or with increasing amount of PrgK linker peptide (residues 76-97) fused to the C-terminus of the small protein SUMO (up to 2:1 ratio in the blue spectrum). The residues for which a significant chemical shift change is observed are identified. (F) Chemical shifts of the residues used for Kd calculation of the PrgK linker interaction with PrgK D1, plotted against protein ratio. (G) The amide chemical shift differences between peaks for free and bound peaks are plotted along the PrgK D1 sequence, with the secondary structure elements indicated at the bottom. The perturbations cluster in helix 1 and helix 2, corresponding to the region of interaction with D2 (see figure 3).
A

Cl-a

Cl-b

EM score

RMSD to lowest energy model (Å)

All-atom energy

B

Cl-a

Cl-b

Molecule i

Molecule i+1

Molecule i

Molecule i+1
Figure S6, related to figure 7: Illustration of the PrgH-PrgK ring modeling

(A) Result of the all-atom refinement step for the PrgH-PrgK composite model. The energy of each model is plotted versus the RMSD relative to the initial model, and color-coded for the fit to the EM map density. The two clusters of models (Cl-a and Cl-b) are indicated with a black circle. (B) Adjacent PrgK monomers from the lowest-energy model of Cl-a (left) and Cl-b (right) are shown. The D1-D2 interaction interface, identified in the NMR-monitored titration (Figure 3C), is indicated with a red arrow. This corresponding to an intra-molecular interaction in the Cl-a cluster of models, and an inter-molecular (domain swapped) interaction in the Cl-b cluster of models. (C) Side-view of individual domains from the PrgH-PrgK model, with PrgH in orange, and PrgK D1 and D2 in green, fitted into the EM density map (EMDB 1875) in grey. The contour level is 0.044 for PrgH and 0.014 for PrgK. All domains show strong fit to EM map, with only the two loops of PrgK D2 not accounted for by EM density. (D) CD spectra of purified PrgH_{170-392} (blue), or with the D333R mutation (red) which abrogates secretion. Ellipticity [θ] is plotted in deg cm^2 dmol\(^{-1}\). Both proteins have similar secondary structure, confirming that the mutation does not alter the overall fold of the domain. (E) Midpoint unfolding temperature of PrgH_{170-392} WT (blue) and D333R mutant (red), plotted against increasing concentration of guanidinium-HCl. Both proteins show similar T\(_m\) values and sensitivity to guanidinium-induced denaturation, as measured by differential scanning fluorimetry. The T\(_m\) value was set to 0 when no unfolding transition was observed. (F) CD spectra of PrgK_{96-200} (blue), or with the K168A/R169A mutation (red). Ellipticity [θ] is plotted in deg cm^2 dmol\(^{-1}\). The K168A/R169A mutation does not affect the overall fold of the protein. (G) Midpoint unfolding temperature of PrgK_{96-200} WT (blue), K168E (green) and K168A/R169A mutant (red), plotted against increasing concentration of guanidinium-HCl. Both mutations show reduced T\(_m\) values compared to wild type, but only the K168A/R169A mutant abrogates secretion (Figure 7D), confirming that the decreased thermal stability is not functionally significant in this assay.
Table S1, related to figure 7: Published experimental data supporting the PrgH-PrgK ring model

<table>
<thead>
<tr>
<th>Buried residues (biotinylation assay)</th>
<th>Prgh</th>
<th>Prgk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lys 230 (Yip et al., 2005)</td>
<td>Lys 48 (Yip et al., 2005)</td>
</tr>
<tr>
<td></td>
<td>Lys 320 (Yip et al., 2005)</td>
<td>Lys 102 (Yip et al., 2005)</td>
</tr>
<tr>
<td></td>
<td>Lys 144 (Yip et al., 2005)</td>
<td>Lys 168 (Yip et al., 2005)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue 1</th>
<th>Residue 2</th>
<th>Cross-Linker</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prgh Lys 278</td>
<td>Prgh Asp 251</td>
<td>SDA</td>
<td>3.9 Å</td>
</tr>
<tr>
<td>Prgh Lys 278</td>
<td>Prgh Glu 252</td>
<td>SDA</td>
<td>3.9 Å</td>
</tr>
<tr>
<td>Prgh Lys 278</td>
<td>Prgh Lys 255</td>
<td>BS2G</td>
<td>7.7 Å</td>
</tr>
<tr>
<td>Prgk Thr 200</td>
<td>Prgk Lys 203</td>
<td>SDA</td>
<td>3.9 Å</td>
</tr>
<tr>
<td>Prgk Lys 203</td>
<td>Prgk Lys 203</td>
<td>SDA, BS2G</td>
<td>3.9 Å, 7.7 Å</td>
</tr>
<tr>
<td>Prgh Ala 171</td>
<td>Prgh Lys 203</td>
<td>SDA</td>
<td>3.9 Å</td>
</tr>
<tr>
<td>Prgh Glu 172</td>
<td>Prgh Lys 203</td>
<td>EDC</td>
<td>12:00 AM</td>
</tr>
<tr>
<td>Prgh Leu 176</td>
<td>Prgh Lys 203</td>
<td>SDA</td>
<td>3.9 Å</td>
</tr>
<tr>
<td>Prgh Glu 180</td>
<td>Prgh Lys 203</td>
<td>EDC</td>
<td>12:00 AM</td>
</tr>
<tr>
<td>Prgh Lys 181</td>
<td>Prgh Lys 203</td>
<td>BS2G</td>
<td>7.7 Å</td>
</tr>
<tr>
<td>Prgh Gly 214</td>
<td>Prgh Lys 203</td>
<td>SDA</td>
<td>3.9 Å</td>
</tr>
<tr>
<td>Prgh Asp 215</td>
<td>Prgh Lys 203</td>
<td>EDC</td>
<td>12:00 AM</td>
</tr>
<tr>
<td>Prgh Lys 218</td>
<td>Prgh Lys 203</td>
<td>BS2G</td>
<td>7.7 Å</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Crystal contacts</th>
<th>Prgh170-392 (Bergeron et al., 2013)</th>
<th>EscJ21.150 (Yip et al., 2005)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PrgkK19-92 This study</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutagenesis</th>
<th>Prgh Y239 (Schraidt and Marlovits, 2011)</th>
<th>Prgh E252 (Bergeron et al., 2013)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prgh G322 (Bergeron et al., 2013)</td>
<td>Prgh D333 This study</td>
</tr>
<tr>
<td></td>
<td>Prgh R169 This study</td>
<td></td>
</tr>
</tbody>
</table>
SUPPLEMENTAL EXPERIMENTAL PROCEDURES:

NMR spectra acquisition, assignment and structure determination:

Secondary structure was predicted from $^{15}$N, $^1$H$_N$, $^{13}$CO, $^{13}$C$_\alpha$ and $^1$H$_\alpha$ chemical shifts using the program SSP (Marsh et al., 2006). Backbone dihedral angle restraints were calculated using TALOS+ (Shen et al., 2009). Interproton distance restraints were obtained from a simultaneous $^{15}$N/$^{13}$C-edited NOESY spectrum ($\tau_{mix} = 100$ ms) and a $^{13}$CH$_3$-edited NOESY spectrum ($\tau_{mix} = 100$ ms), both collected using an 850 MHz spectrometer. NOE assignments and structure calculations were done using Cyana (Guntert, 2004), followed by water refinement with CNS (Brunger, 2007). Chemical shifts have been deposited to the BRMB database (BMRB ID: 19800), and the 20 lowest-energy structures coordinates have been deposited on the PDB (PDB ID: 2MKY).

NMR titrations were performed by adding 50 μl aliquots of unlabeled proteins at 1 mM into 500 μl $^{15}$N-labelled PrgK$_{19-76}$ at 0.25 mM. Weighted combined $^{15}$N and $^1$H$_N$ chemical shift differences ($\Delta$ (H, N)) in ppm were calculated using the following formula:

$$\Delta(H, N) = \sqrt{(\Delta H)^2 + (0.154\Delta N)^2}$$

For residues of PrgK$_{19-92}$ where only one population was identified (Figure 6), this value was set to 0.

Equilibrium dissociation constants (Kd values) were determined by fitting, with GraphPad Prism, $\Delta$(H,N) to the following equation for a 1:1 binding isotherm:

$$\Delta(H, N) = \frac{([A] + [B] + Kd) - \sqrt{([A] + [B] + Kd)^2 - 4[A][B]}}{2[A]}$$

where [A] and [B] are the total, dilution-adjusted concentrations of labeled and unlabeled species, respectively.

PrgK D2 crystal structure:

PrgK$_{98-200}$ crystallized in a range of conditions using commercial screens, all with pH < 6.0, and including PEG as a precipitant. However, most crystals gave
limited X-ray diffraction. Optimization of the crystallization conditions allowed us to produce crystals that diffracted to \( \sim 3 \, \text{Å} \), although the resulting diffraction patterns showed a number of pathologies, including split/smeary spots and anisotropy. Nevertheless, extensive screening of over 50 crystals at synchrotron facilities allowed us to collect a single dataset, which had acceptable statistics to 2.6 Å (see Table 2). This dataset was merged with XDS (Kabsch, 2010), and scaled with Scala (Winn et al., 2011). We note that upon scaling, severe anisotropy was identified. In addition, Phenix Xtriage (Adams et al., 2011) reported the presence of translational pseudosymmetry (18 % origin). Nonetheless, we were able to solve the structure by molecular replacement (MR), using the structure of the EscJ domain as a search model (residues 98 to 190 of PDB ID: 1YJ7; z-score = 7.8, LLG = 128). Of note, a MR solution could only be obtained after anisotropic scaling (Strong et al., 2006) of the structure factors. The MR solution showed the presence of two molecules in the asymmetric unit, and confirmed that they are close to the crystallographic symmetry orientation, leading to translational pseudo-symmetry in the crystal (not shown).

Extensive rounds of building using Coot (Emsley et al., 2010), and refinement using Refmac5 (Murshudov et al., 2011), Phenix refine (Afonine et al., 2012), and Buster (Smart et al., 2012), with TLS parameters (Painter and Merritt, 2006), allowed us to obtain a structural model with acceptable refinement statistics (\( R=22.7\%; \ R_{\text{free}}=25.9\%; \) Clashscore of 1.88 (100\textsuperscript{th} percentile); 92.7 % residues in the most favorable region of the Ramachandran plot, and three residues (1.8 %) in the disallowed region, as measured by MolProbity (Chen et al., 2010), corresponding to poorly-defined loop regions of the structure). Although the refinement statistics are satisfactory and an all atom simulated annealing composite omit map calculated with Phenix (Afonine et al., 2012) corroborates the validity of the model (data not shown), we note that final model contains few visibly ordered water molecules. We ascribe this to the pathologies intrinsic to the data described above.

**PrgKD1-linker crystal structure:**
PrgK_{19-96} formed very thin needle-like crystals in 100 mM Phosphate buffer pH 4, 30% PEG300. Larger, single crystals suitable for data collection were obtained by optimizing the crystallization conditions. Most crystals diffracted very poorly (~10 Å), but extensive screening of over 50 crystals at synchrotron facilities led to a single dataset with good statistics to 3.2 Å. The Matthews coefficient obtained for this dataset (Kantardjieff and Rupp, 2003) indicated that the asymmetric unit contains 10-16 molecules.

Initial attempts to obtain phase information by molecular replacement, based on the PrgK_{19-76} solution structure (Figure 2B), or the crystal structure of Escj (Yip et al., 2005), were unsuccessful. Postulating that this may be caused by the fact that the crystals contain the linker-bound conformation of the protein (Figure 6C), we used CS-Rosetta (van der Schot et al., 2013) to generate a model for population A of PrgK_{19-96}, based on the chemical shift assignment (Figure 6B). Using this model, a MR solution was found, when searching for 10 molecules in the ASU (z-score = 7.5, LLG = 757). An additional 4 molecules were manually placed in the density during rounds of building with Coot (Emsley et al., 2010), and refinement with Refmac5 (Murshudov et al., 2011) and Phenix refine (Afonine et al., 2012), using NCS restraints. The obtained model has good statistics: R=24.7%, Rfree=27.6%, Clashscore of 9.75 (95th percentile); 98.6% residues in the most favorable region of the Ramachandran plot, and one residue in the disallowed region, as measured by MolProbity (Chen et al., 2010).

Residues 19-79 (including the D1-interacting region of the linker) were well resolved in all 14 molecules of the ASU, and additional linker residues could be confidently modeled in 9 of the molecules. We also note that strand β0 is well defined in this structure, which could be attributed to the interaction between Pro 78 of the linker, and Leu 24 and Leu 27 of β0, which may help stabilize it.

**ITC, SEC-MALS, CD and DSF:**

ITC experiments were performed using a Microcal ITC_{200} system (GE Healthcare). All samples were in lysis buffer; 20 x 3 μl aliquots of PrgK_{19-76}, PrgK_{19-76} Y70A, or PrgK_{19-92} at 2 mM were injected into 250 μl PrgK D2 at 0.1 mM.
SEC-MALS experiments were performed using an Agilent 110 HPLC (Agilent technologies) connected in-line to a Dawn HeleosII 18-angle MALS light scattering detector followed by an Optilab T-rEX differential refractometer detector (Wyatt), in lysis buffer. Bovine serum albumin (Sigma) was used to normalize the light scattering detectors. 100 μl of PrgK82-200, PrgK82-200 F89A, PrgK19-200 and PrgK95-200 proteins, at 0.31, 0.2, 0.27 and 0.22 mM respectively, were injected on a Superdex 75 HR size-exclusion chromatography column (GE Healthcare) at 0.2 ml/min. Data were collected and analyzed with the Astra 6 software package (Wyatt).

Circular dichroism spectra were recorded with a nitrogen-flushed Jasco J-810 spectro-polarimeter, at 25°C. Proteins at 0.1 mg/ml concentration were dialyzed against 25 mM Phosphate pH 7.0 buffer prior to analysis. All spectra are averages of three scans, and were acquired using a quartz cell of 0.2 cm optical path length. Spectra were corrected for buffer.

Differential scanning fluorimetry experiments were carried out using 0.3 mg/ml protein in 50mM HEPES pH 8.0, 150 mM NaCl and SYPRO Orange (Life Technologies, 5× final concentration). A 25-μl assay volume mixture was transferred into a MicroAmp Fast optical reaction plate (Applied Biosystems) and thermodenaturation monitored in an Applied Biosystems StepOnePlus RT-PCR system using the ROX filter set (excitation, 488 nm; emission, 620 nm) at a rate of 1°C per minute from 25°C to 95°C with fluorescence measured every 0.5°C.

All ITC and SEC-MALS experiments were repeated twice, using independently purified protein samples, with one representative example shown.

**EM-guided symmetrical modeling in Rosetta:**

The monomeric structures of PrgK D1 and D2 were placed in the EM map density (Schraidt and Marlovits, 2011) (EMDB ID: 1875) according to their putative global localization and relative topology (such that the N and C-termini are localized towards the inner membrane region), and fitted into the EM map using the “fit in map” command in Chimera (Goddard et al., 2007). A 24-fold symmetry file was generated from the PrgH 24-mer model reported previously (DiMaio et al.,
2011), and was used in the phase I rigid body Rosetta EM-guided symmetry modeling procedure; 1000 models were generated for each structure. For each run, the energy was plotted versus the RMSD to the lowest energy model and color-coded according the fit to EM map density. The lowest total docking energy model was then used as the input for the phase II all-atom, EM-guided symmetry modeling procedure with backbone and side-chain refinement. The weight of the EM map restraint was half that used in the first, rigid-backbone step. The energy was plotted versus the RMSD to the starting model and color-coded according the fit to EM map density. Clusters of models were identified using the cluster module of Rosetta.

**PrgK D1 ring modeling:**

The Rosetta modeling of PrgK D1 yielded two distinct clusters, denoted as Cl-1 and Cl-2 (Figure S4A, left). Both the Cl-1 and Cl-2 conformations of D1 are in close proximity of D2, albeit with distinct regions of interface presented in the 2 conformations of the former (Figure S3B). Specifically, in the Cl-1 conformation, helix α1 of D1 faces D2, with α2 directed away, whereas in the Cl-2 conformation, the surface formed by α1 and α2 faces D2. The Cl-2 conformation is very similar to the equivalent domain in the 24-mer EscJ ring reconstituted from the crystallographic super-helix (Yip et al., 2005) (Figure S3). In contrast, the D1 Cl-1 conformation does not agree with the EscJ super-helix structure (not shown). In addition, the D1 Cl-2 conformation agrees well with the NMR-monitored titration data shown on Figure 3C. Alternatively, Tyr70, which is essential for the domains to interact *in vitro*, is located away from D2 in the Cl-1 models. Collectively, these data support the Cl-2 conformation for D1, and invalidate the Cl-1 set of models.

**PrgK D2 ring modeling:**

For one of two molecules in the asymmetric unit of the PrgK D2 crystal (chain B), we could build the majority of residues, with the exception of residues 137-142 in the loop between strands 1 and 2. Chain A was less well defined, with little density for residues 134-145 in that same loop. We proceeded with the modeling strategy (See Experimental Procedures section), using either chain A or
chain B. Whereas chain B gave a single cluster in the first, rigid-body step (Figure S3A, right), chain A produced a number of clusters (not shown), of which the lowest-energy corresponds to the cluster obtained for chain B. We postulate that this difference is due to the fact that chain B has most of the extended loop 133-145 modeled, for which there is clear density in the EM map, guiding the structure in the correct orientation. This illustrates the importance of the EM density restraint in the modeling procedure.

**PrgH-PrgK ring modeling:**

For the PrgH-PrgK modeling, the PrgH, PrgK D1 and PrgK D2 24-mer models were placed in the EM map density, and the coordinates for adjacent monomers of PrgH, PrgK D1 and PrgK D2 were saved as a single chain. This file was used as an input for the all-atom refinement phase of the modeling procedure. As illustrated on Figure S6A, this led to two distinct clusters of models, labeled Cl-a and Cl-b. Upon inspection, we noticed that these two clusters are actually very similar, with all three molecules oriented similarly in Cl-a and Cl-b; however, PrgK D1 has been moved to the next subunit in Cl-b compared to Cl-a (Figure S6B). In other words, the models from Cl-a and Cl-b correspond to different binding modes for a similar conformation. This is why the RMSD value appears different, but actually belongs to a unique cluster of models.

This model shows a strong fit to EM map density (Figure S6C), and agrees with additional experimental data such as crystal packing from the EscJ super-helix (Yip et al., 2005), and crystal contacts for PrgH<sub>170-392</sub> (Bergeron et al., 2013) and PrgK<sub>19-96</sub> (This work). In addition the model is consistent with biotinylation (Sanowar et al., 2009; Spreter et al., 2009; Yip et al., 2005) and cross-linking experiments (Sanowar et al., 2009; Schraidt et al., 2010), as summarized in table S1.
SUPPLEMENTAL REFERENCES:


