Structural and Functional Studies on the Interaction of GspC and GspD in the Type II Secretion System

Konstantin V. Korotkov¹, Tanya L. Johnson², Michael G. Jobling³, Jonathan Pruneda¹, Els Pardon⁴,⁵, Annie Héroux⁶, Stewart Turley¹, Jan Steyaert⁴,⁵, Randall K. Holmes³, Maria Sandkvist², Wim G. J. Hol¹

¹ Department of Biochemistry, Biomolecular Structure Center, University of Washington, Seattle, Washington, United States of America, ² Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan, United States of America, ³ Department of Microbiology, University of Colorado School of Medicine, Aurora, Colorado, United States of America, ⁴ Department of Molecular and Cellular Interactions, VIB, Brussels, Belgium, ⁵ Structural Biology Brussels, Vrije Universiteit Brussel, Brussels, Belgium, ⁶ National Synchrotron Light Source, Brookhaven National Laboratory, Upton, New York, United States of America

Abstract

Type II secretion systems (T2SSs) are critical for secretion of many proteins from Gram-negative bacteria. In the T2SS, the outer membrane secretin GspD forms a multimeric pore for translocation of secreted proteins. GspD and the inner membrane protein GspC interact with each other via periplasmic domains. Three different crystal structures of the homology region domain of GspC (GspCHR) in complex with either two or three domains of the N-terminal region of GspD from enterotoxigenic Escherichia coli show that GspC-HR adopts an all-β topology. N-terminal β-strands of GspC and the N0 domain of GspD are major components of the interface between these inner and outer membrane proteins from the T2SS. The biological relevance of the observed GspC–GspD interface is shown by analysis of variant proteins in two-hybrid studies and by the effect of mutations in homologous genes on extracellular secretion and subcellular distribution of GspC in Vibrio cholerae. Substitutions of interface residues of GspD have a dramatic effect on the focal distribution of GspC in V. cholerae. These studies indicate that the GspC-HR–GspD-N0 interactions observed in the crystal structure are essential for T2SS function. Possible implications of our structures for the stoichiometry of the T2SS and exoprotein secretion are discussed.

Citation: Korotkov KV, Johnson TL, Jobling MG, Pruneda J, Pardon E, et al. (2011) Structural and Functional Studies on the Interaction of GspC and GspD in the Type II Secretion System. PLoS Pathog 7(9): e1002228. doi:10.1371/journal.ppat.1002228

Introduction

Many Gram-negative bacteria use a multi-protein type II secretion system (T2SS) to secrete a wide variety of exoproteins from the periplasm into the extracellular milieu [1,2,3,4]. In Vibrio cholerae and enterotoxigenic Escherichia coli (ETEC), cholera toxin and the closely related heat-labile enterotoxin, in addition to other virulence factors, are secreted in their folded state across the outer membrane by the T2SS [3,6,7]. The T2SSs are composed of 12 to 15 different proteins that form three distinct subassemblies: (i) the inner membrane platform consisting of multiple copies each of GspG, GspF, GspL, and GspM with an associated cytoplasmic secretion ATPase; (ii) the pseudopilus, a filamentous arrangement of multiple copies of five different pseudopilins; and (iii) a large, pore-forming outer membrane complex, mainly consisting of the secretin GspD [8,9].

Secretins are multimeric outer membrane proteins composed of 50–70 kDa subunits and are among the largest outer membrane proteins known. The secretin superfamily has representatives in several other multi-protein complexes engaged in transport of large macromolecular substrates across the outer membrane [10] including the T2SS, the filamentous phage extrusion machinery [11], the type IV pilus system (T4PS) [12,13,14], and the type III secretion system (T3SS) [15,16]. Of these systems, the T2SS is most closely related to the T4PS which assembles and disassembles long filamentous fibers on bacterial surfaces and is responsible for diverse functions including attachment to host cells, biofilm formation, DNA uptake and twitching motility [17,18].

The T2SS secretin GspD forms a dodecameric assembly according to electron microscopy studies [19,20]. The C-terminal 300 to 400 residues of GspD contain the most conserved segments of the secretin superfamily, which form the actual outer membrane pore [21,22,23]. The N-terminal part of GspD consists of four domains: N0–N1–N2–N3 (Figure 1A) [19,24]. The crystal structure of the N0–N1–N2 domains of the ETEC secretin GspD has been solved previously with the assistance of a single-domain llama antibody fragment or nanobody [24]. Nanobodies are the antigen-binding fragments (VHH) of heavy-chain-only Fab antibodies, which have been proven as effective crystallization chaperones for
Many bacterial pathogens affecting humans, animals and plants export diverse proteins across the cell membranes into the medium surrounding the bacteria. Some of these secreted proteins are involved in pathogenesis. One example is cholera toxin secreted by the bacterium *Vibrio cholerae*, a causative agent of cholera. The sophisticated type II secretion system is responsible for moving this toxin, and several other proteins, across the outer membrane. Here, we studied the interaction between the outer membrane pore of the type II secretion system, the secretin GspD, and the inner membrane protein GspC. We have solved three crystal structures of complexes between the interacting domains and identified critical contacts in the GspC–GspD interface. We also showed the importance of these contacts for assembly of the secretion system and for secretion of proteins by *V. cholerae*. Our studies provide a major piece in the puzzle of how the type II secretion system is assembled and how it functions. One day this knowledge might allow us to design compounds which interfere with this secretion process. Such compounds would be useful in the battle against bacteria affecting human health.

**Author Summary**

Many bacterial pathogens affecting humans, animals and plants export diverse proteins across the cell membranes into the medium surrounding the bacteria. Some of these secreted proteins are involved in pathogenesis. One example is cholera toxin secreted by the bacterium *Vibrio cholerae*, a causative agent of cholera. The sophisticated type II secretion system is responsible for moving this toxin, and several other proteins, across the outer membrane. Here, we studied the interaction between the outer membrane pore of the type II secretion system, the secretin GspD, and the inner membrane protein GspC. We have solved three crystal structures of complexes between the interacting domains and identified critical contacts in the GspC–GspD interface. We also showed the importance of these contacts for assembly of the secretion system and for secretion of proteins by *V. cholerae*. Our studies provide a major piece in the puzzle of how the type II secretion system is assembled and how it functions. One day this knowledge might allow us to design compounds which interfere with this secretion process. Such compounds would be useful in the battle against bacteria affecting human health.

**Results**

**Structures of Three GspC–GspD Complexes from ETEC**

A complex of ETEC GspC^{HR} and GspD^{N0-N1-N2} could be obtained but yielded only poorly diffracting crystals. To improve the quality of these crystals, we screened the same set of GspD specific nanobodies that had been used previously to solve the structure of GspD^{N0-N1-N2} as crystallization chaperones for the GspC^{HR}–GspD^{N0-N1-N2} complex. Using nanobody Nb3, we obtained crystals of a ternary ETEC GspC^{HR}–GspD^{N0-N1-N2}–Nb3 complex, which diffracted initially only to ~5.5 Å resolution. Nevertheless, a partial molecular replacement structure revealed that the HR domain of GspC interacts with the lobe formed by the N0-N1 domains of GspD. To better characterize this interaction we also crystallized smaller complexes of GspC^{HR}–GspD^{N0-N1} with or without nanobodies. To assist in crystallographic phasing, we also engineered a lanthanide-binding tag (LBT) into the N0 domain of GspD^{N0-N1} [41]. The LBT to GspD^{N0-N1} facilitated crystal growth and the resultant crystals of the binary GspC^{HR}–GspD^{N0-N1} complex diffracted to better than 2.7 Å resolution, with the LBT engaged in multiple crystal contacts (Figure S1). The structure of this binary GspC^{HR}–GspD^{N0-N1} complex was solved by molecular replacement and refined with good crystallographic and stereochemical statistics (Table 1). In parallel, we also obtained crystals and solved the 4 Å resolution structure of a ternary GspC^{HR}–GspD^{N0-N1}–Nb3 complex, and also improved the diffraction of crystals of the GspC^{HR}–GspD^{N0-N1}–Nb3 complex to ~4 Å resolution (Table 1, Figure S2).

The three multiprotein structures obtained from different crystal forms allow a detailed description of the interactions between GspC and GspD. In all three structures, the N0 domain of GspD interacts exclusively with the HR domain of GspC. In the 2.63 Å resolution binary complex, the LBT introduced into GspD^{N0} faces away from the interface with GspC (Figure 1B). In the two low-resolution ternary complex structures, the nanobody Nb3 binds the N0 domain of GspD, opposite to the binding site of the HR domain of GspC (Figure 1C and Figure S3). In all three structures the HR domain binds in very similar orientation to GspD, relative to its N0 domain. Hence, neither the LBT nor the nanobody appears to affect the binding mode of GspC to GspD. Because the structure of the binary GspC^{HR}–GspD^{N0-N1} complex has the
highest resolution, this structure will be used below to analyze the specific features of the GspC–GspD interaction.

Structure of the HR Domain of GspC

The HR domain folds into a β-sandwich formed by six consecutive β-strands arranged as two three-stranded anti-parallel β-sheets (Figure 1B). The residues between strands β3 and β4 adopt an approximately one-turn helical conformation. In its folded structure as seen in the complex with GspD<sup>N0,N1</sup> (Figure 2), the distribution of charges on the surface of GspC<sup>HR</sup> is quite uneven with the main hydrophobic surface interacting with GspD<sup>N0</sup>. Part of the remaining HR surface that is not involved in the GspD interaction is defined by residues Val127, Ile142 and Leu157. The other side of the HR domain (lower panel Figure 2) displays a mix of positive, negative and...
hydrophobic patches. The functions of these features during assembly and action of the T2SS, if any, remain to be determined.

The Comparison of the HR Domain of GspC from the T2SS and PilP from the T4PS

The closest known structural homolog of the HR domain of ETEC GspC appears to be Neisseria meningitidis lipoprotein PilP (Nm PilP) which interacts with the secretin of the T4PS [42]. The HR domain of GspC and the core domain of Nm PilP superimpose with an r.m.s. deviation of 1.6 Å and 25% sequence identity over 59 residues (Figure 3). The structure of Nm PilP has been described as a \( \beta \)-sandwich composed of 7 \( \beta \)-strands [42]. Whereas residues 154–156 of GspC, corresponding to strand \( \beta 4 \) of Nm PilP, make some main chain hydrogen bonds to residues in strand \( \beta 4 \) of GspC (corresponding to \( \beta 5 \) of Nm PilP), the secondary structure assignment algorithm of DSSP [43] does not classify these residues as \( \beta \)-structure.

A potential binding site has been described for the core Nm PilP domain [42]. It consists of a hydrophobic crevice on the open end of the \( \beta \)-sandwich. The residues that create this hydrophobic groove appear to be conserved between these two proteins from the T2SS and the T4PS when they are superimposed (Figure 3C). The structure of Nm PilP is from the binary complex with GspD\(^{N0-N1}\); the LBT is omitted for clarity. Note the deep pocket in the front surface of GspC\(^{HR}\) in the upper panel.

The GspC–GspD Interface

The interface between GspC\(^{HR}\) and GspD\(^{N0}\) buries 1280 Å\(^2\) of accessible surface area with a calculated \( \Delta G \) of interaction of \(-5.4\) kcal\(\times\)mol\(^{-1}\) as assessed by the PISA server (Figure 4) [44]. The overall shape of the interface is relatively flat with a small concave area on the GspD surface. A total of 18 residues from GspC\(^{HR}\) differences do not appear to stem from crystal contacts in the GspC\(^{HR}\)-GspD\(^{N0-N1}\) structure. Moreover, these residues are well conserved (Figure S4A) and contribute to the hydrophobic core of the HR domain. The full implications of the global structural similarity between the core PilP domain of the T4PS and the HR domain of GspC from the T2SS remain to be established, but it is in line with several known similarities between the T2SS and T4PS [17,18].

The GspC–GspD Interface

The interface between GspC\(^{HR}\) and GspD\(^{N0}\) buries 1280 Å\(^2\) of accessible surface area with a calculated \( \Delta G \) of interaction of \(-5.4\) kcal\(\times\)mol\(^{-1}\) as assessed by the PISA server (Figure 4) [44]. The overall shape of the interface is relatively flat with a small concave area on the GspD surface. A total of 18 residues from GspC\(^{HR}\) differences do not appear to stem from crystal contacts in the GspC\(^{HR}\)-GspD\(^{N0-N1}\) structure. Moreover, these residues are well conserved (Figure S4A) and contribute to the hydrophobic core of the HR domain. The full implications of the global structural similarity between the core PilP domain of the T4PS and the HR domain of GspC from the T2SS remain to be established, but it is in line with several known similarities between the T2SS and T4PS [17,18].

The GspC–GspD Interface

The interface between GspC\(^{HR}\) and GspD\(^{N0}\) buries 1280 Å\(^2\) of accessible surface area with a calculated \( \Delta G \) of interaction of \(-5.4\) kcal\(\times\)mol\(^{-1}\) as assessed by the PISA server (Figure 4) [44]. The overall shape of the interface is relatively flat with a small concave area on the GspD surface. A total of 18 residues from GspC\(^{HR}\) differences do not appear to stem from crystal contacts in the GspC\(^{HR}\)-GspD\(^{N0-N1}\) structure. Moreover, these residues are well conserved (Figure S4A) and contribute to the hydrophobic core of the HR domain. The full implications of the global structural similarity between the core PilP domain of the T4PS and the HR domain of GspC from the T2SS remain to be established, but it is in line with several known similarities between the T2SS and T4PS [17,18].

Table 1. Data collection and refinement statistics.

|                     | GspC\(^{HR}\)–GspD\(^{N0-N1}\) | GspC\(^{HR}\)–GspD\(^{N0-N1-N2–Nb3}\) | GspC\(^{HR}\)–GspD\(^{N0-N1-N1–Nb3}\) |
|---------------------|-----------------------------|----------------------------------|----------------------------------|
| **Data collection** |                             |                                  |                                  |
| Wavelength (Å)      | 0.97946                     | 0.97973                          | 0.97946                          |
| Space group         | P2\(_1\)2\(_1\)2\(_1\)       | P6\(_3\)22                       | P6\(_3\)22                       |
| Unit cell dimensions|                             |                                  |                                  |
| a, b, c (Å)         | 45.50, 76.81, 85.77         | 142.20, 142, 57.57              | 156.57, 156.57                   |
| a, b, c (deg)       | 90, 90, 90                 | 90, 90, 120                     | 90, 90, 120                     |
| Resolution (Å)      | 42.88–2.63                 | 43.93–4.05                       | 41.69–4.00                       |
| Completeness (%)    | 99.9 (99.9)                 | 99.8 (99.7)                      | 99.9 (100)                       |
| Redundancy          | 3.9 (3.9)                  | 7.7 (7.8)                        | 7.7 (7.6)                        |
| R\(_{merge}\) (%)   | 13.5 (84.3)                | 8.5 (88.2)                       | 15.9 (94.6)                      |
| I/\(\sigma\)(I)     | 11.0 (2.1)                 | 16.7 (2.8)                       | 10.0 (2.3)                       |
| **Refinement**      |                             |                                  |                                  |
| Resolution (Å)      | 42.88–2.63                 |                                  |                                  |
| R\(_{work}/\)R\(_{free}\) (%) | 21.3/26.5           |                                  |                                  |
| No. of reflections  | 8942                       |                                  |                                  |
| No. of atoms        | 1775                       |                                  |                                  |
| \(\beta\) factor (Å\(^2\)) | 31.2                        |                                  |                                  |
| R.m.s. deviations   |                             |                                  |                                  |
| Bond lengths (Å)    | 0.013                      |                                  |                                  |
| Bond angles (deg)   | 1.314                      |                                  |                                  |
| Ramachandran values\(^a\) |                    |                                  |                                  |
| Favoured (%)        | 98.2                       |                                  |                                  |
| Allowed (%)         | 1.8                        |                                  |                                  |

\(^a\)Values in parenthesis are for the highest resolution shell.

\(^b\)Molprobity [73], http://molprobity.biochem.duke.edu/

doi:10.1371/journal.ppat.1002228.t001

Figure 2. Surface characteristics of ETEC GspC\(^{HR}\). A stereo representation of surface charge distribution of GspC\(^{HR}\). The HR domain structure is from the binary complex with GspD\(^{N0-N1}\). The LBT is omitted for clarity. Note the deep pocket in the front surface of GspC\(^{HR}\) in the upper panel.

doi:10.1371/journal.ppat.1002228.g002
and 19 residues from GspD are engaged in a combination of hydrophobic interactions and hydrogen bonds. The first three β strands of GspC and the first β strand plus the subsequent helix of GspD are the major contributors to the interface. The majority of the hydrogen bonds are formed by an antiparallel arrangement of strand β1 of GspC and strand β1 of GspD (Figure 4C). This β-strand augmentation is frequently observed in protein–protein interfaces [45]. Several nonpolar residues are engaged in intermolecular hydrophobic interactions, e.g. Ala133/Val141 from GspC, and Phe5/Phe9 from GspD. The hydrophobic nature of these interacting residues is well conserved, with GspC residue 133 being Ala, Leu, Val or Met; GspC residue 141 either a Val or Ile; GspD residue 5 a Phe or Tyr; and GspD residue 9 a Phe according to a family sequence alignment (Figure S4). Nonetheless, the GspC–GspD interface provides a species-specific connection point between outer and inner membrane assemblies of the T2SS as has been observed in genetic complementation studies [46,47].

Figure 3. Comparison of GspC from the T2SS and PilP from the T4PS. (A) Structural superposition of the ETEC GspC domain with the N. meningitidis PilP structure (PDB 2IVW) [42]. GspC and PilP are colored in green and yellow, respectively. Flexible N- and C-terminal residues of PilP are not shown for clarity. The conserved hydrophobic residues are shown as sticks. (B) Surface representation of GspC and PilP in the same orientation as in (A). The crevice on the surface of PilP is absent in GspC. (C) Sequence alignment of GspC and PilP based on the structural superposition in (A). Secondary structure elements of GspC and PilP are displayed above and below the sequences, respectively; the colored dots represent the conserved hydrophobic residues of GspC and PilP.

doi:10.1371/journal.ppat.1002228.g003
Figure 4. The interface of the GspC–GspD complex. (A) An ‘open book’ view of the GspC\(^{HR}\)–GspD\(^{N0-N1}\) binary complex in surface representation. Residues in the interface are colored according to the degree of burial upon complex formation: yellow, up to 40% reduction in accessible surface area (ASA); orange, 40–70% reduction in ASA; and brown, more than 70% reduction in ASA. Atoms participating in intermolecular hydrogen bond formation are colored in cyan. (B) Same view as in (A) with the interaction surfaces colored according to the solvent accessible electrostatic potential. The interaction surface is contoured by black lines. (C) Anti-parallel β\(^1\)HR–β\(^1\)N0 interactions in the GspC\(^{HR}\)–GspD\(^{N0-N1}\) complex. The upper strand is β\(^1\)N0. Interacting residues are shown as sticks and labeled. Hydrogen bonds are shown as dashed lines. A σ\(^A\)-weighted 2\(F_\text{O}\)–\(F_\text{C}\) electron density map contoured at 1.2 \(σ\) is shown as a dark blue mesh. (D) Interface surface of a homology model of the \(V.\) cholerae GspC–GspD complex [79]. Residues in the interface are colored according to the color of the interacting partner: GspD in cyan and GspC in green. The residues that were subjected to mutational analysis are colored in magenta and labeled. (E) Amino acid sequence alignments of the HR domains of GspC and the N0 domains of GspD from ETEC and \(V.\) cholerae. The corresponding secondary structure elements are shown above the sequences. Residues that make intermolecular Van der Waals contacts and H-bonds in the ETEC GspC\(^{HR}\)–GspD\(^{N0-N1}\) complex are labeled by triangles and stars, respectively. The residues that were subjected to mutational analysis in \(V.\) cholerae GspC and GspD are indicated by magenta triangles underneath the alignments. doi:10.1371/journal.ppat.1002228.g004
Based on the GspC\textsuperscript{HR}–GspD\textsuperscript{N0–N1} structure, we selected several well-conserved interface residues for subsequent substitutions. Ala133 and Val141 from ETEC GspC (equivalent to Val118 and Val129 from \textit{V. cholerae} GspC, respectively; Figure 4E) and Thr20 from ETEC GspD (equivalent to Thr25 of \textit{V. cholerae} GspD) makes a hydrogen bond on the main chain oxygen of ETEC GspC Arg137. We evaluated the role of these residues on complex formation of truncated forms of GspC and GspD in a bacterial two-hybrid system and in a functional \textit{V. cholerae} secretion assay \textit{in vivo}. We also assessed the effect of interface substitutions on the distribution of GspC in the cell envelope of \textit{V. cholerae}.

Tests of GspC–GspD Interactions in the Bacterial Two-hybrid System

The effect of several interface substitutions on the ability of GspD to associate with GspC was assayed in a bacterial two-hybrid system based on reconstitution of activity of the catalytic domain of \textit{Bordetella pertussis} adenylate cyclase when T18 and T25 fragments are fused to interacting proteins [see Methods] [48]. V/GspD–T18 with a conservative Asn22Gln substitution retained the ability to interact with T25–GspC and formed dark red colonies on indicator agar. In contrast, V/GspD–T18 with either an Asn22Arg substitution or an Ile18Asp substitution lost the ability to interact with T25–GspC and formed colorless colonies (Table 2). Two variants of T25–GspC, with either Val118Arg or Val129Arg substitution, also lost the ability to interact with V/GspD–T18 and formed colorless colonies in the bacterial two-hybrid system.

Mutations in the GspC–GspD Interface Interfere with Protease Secretion in \textit{V. cholerae}

The functional importance of residues involved in the GspC–GspD interface was also assessed \textit{in vivo} by monitoring the effect of the Ile18Arg and Asn22Tyr mutations in V/GspD on the extracellular secretion of protease by \textit{V. cholerae}. No protease secretion was observed when plasmid-encoded V/GspD\textsuperscript{Ile18Arg/Asn22Tyr} was produced in a \textit{V. cholerae} mutant strain lacking the gene encoding GspD (Figure 5A), indicating that the simultaneous exchange of these two amino acids prevents protein secretion by the T2SS. The singly substituted variants, however, remained functional (Figure 5A). Immunoblot analysis of cell extracts from the \textit{AoppD} mutant strain producing plasmid-encoded wild type and mutant V/GspD showed that the double V/GspD mutant protein was made at levels similar to that of wild-type V/GspD (Figure 3B).

\textbf{Table 2. Characterization of GspC–GspD interaction in the bacterial two-hybrid system.}

| GspC  | GspD     | Interaction |
|-------|----------|------------|
| wt*   | wt       | +          |
| wt    | Asn22Gln | +          |
| wt    | Asn22Arg | -          |
| wt    | Ile18Asp | -          |
| Val118Arg | wt  | -          |
| Val129Arg | wt  | -          |

*wt – wild type.

doi:10.1371/journal.ppat.1002228.t002

Distribution of GFP-VcGspC in \textit{V. cholerae} Cells

Using \textit{V. cholerae} strains producing chromosomally encoded V/GspC fused to the green fluorescent protein (GFP), we visually examined the effects of substitutions in the GspC–GspD interface on subcellular localization of GspC. GFP-V/GspC forms fluorescent foci in the \textit{V. cholerae} cell envelope, which disperse upon deletion of the gene encoding GspD and reappear when the deletion strain is complemented with plasmid-encoded V/GspD (Figure 6, first and second panels) [59]. The substitution of wild-type V/GspD with V/GspD\textsuperscript{Ile18Arg/Asn22Tyr} resulted in loss of fluorescent foci and dispersal of the fluorescence in a manner indistinguishable from cells that do not have the gene encoding GspD at all (Figure 6, fourth panel). This result suggests that residues Ile18 and Asn22 of V/GspD are critical for the incorporation of GFP-V/GspC fusion protein into fluorescent foci, and supports the suggestion that the interaction between GspC and GspD observed in the crystal structure of GspC\textsuperscript{HR} in complex with GspD\textsuperscript{N0–N1} (Figure 4) is physiologically relevant. Based on these results, it appears that V/GspD has to interact directly with V/GspC in order to support its focal distribution in \textit{V. cholerae}.

Discussion

The current paper reveals for the first time key structural features of critical interactions between the outer membrane secretin GspD and the inner membrane protein GspC of the T2SS. The crystallographic studies benefited from the set of nanobodies against the N0-N1-N2 domains of GspD from ETEC [24] and from the incorporation of a lanthanide-binding tag (LBT) into ETEC GspD\textsuperscript{N0}. The three GspC–GspD crystal structures elucidated reveal the same 1280 Å\textsuperscript{2} interface involving the HR domain of GspC and the N0 domain of GspD. The crucial role of this interface was tested and confirmed by subsequent biochemical and functional studies. These results have interesting implications for our understanding of the T2SS and related secretion systems in many bacteria as discussed below.

The Mutual Orientation of the N0 and N1 Domains of GspD

The structures of the first two domains of related secretins have been determined in two prior studies: ETEC GspD from the T2SS and EPEC EscC from the T3SS [24,49]. The relative orientations of the N0 and N1 domains in these two studies appeared to be remarkably different: when the N1 domains of the T2SS and T3SS secretins are superimposed, the N0 domains are rotated by not less than 143 degrees [10]. This raises an important question as to the actual orientation of these two domains in the T2SS and T3SS secretins.

Regarding the T2SS, the relative orientations of the N0 and N1 GspD domains can now be compared in two high resolution structures, i.e. in the current structure of the binary complex of ETEC GspC\textsuperscript{HR} and GspD\textsuperscript{N0–N1}, and in the previously determined binary complex of ETEC GspD\textsuperscript{N0–N1–N2} in complex with Nb7 [24]. The linker between the N0 and N1 subdomains is disordered in both these high resolution structures. The interface and relative orientation of the N0 and N1 subdomains, however, is essentially the same in the two structures despite the binding of either Nb7 or the presence of the LBT insertion into the N0 domain: the superposition of the two N0 domains results in a r.m.s. deviation of 0.49 Å for 72 C\textit{z} pairs (Figure S5). Taking also into account the two new low resolution structures of the ternary complexes of GspC\textsuperscript{HR}–GspD\textsuperscript{N0–N1–N2}–Nb3 and GspC\textsuperscript{HR}–GspD\textsuperscript{N0–N1–N2}–Nb3 (Figure S3), then the N0-N1 lobe in the T2SS secretin GspD is observed as the same compact unit in four different crystal
structures, independent of the presence or absence of a GspC{414R} domain, Nb molecules or crystal contacts. The available data suggest that the N0-N1 orientation in GspD is a characteristic feature in the T2SS. However, we cannot exclude the possibility that the relative orientation of the N0 and N1 domains may change as the secretin oligomerizes. Only high resolution structures of the dodecameric secretin will resolve this question.

Since the N0-N1 lobe of the T2SS secretin fits well into the cryo-electron microscopy reconstruction of VcGspD [20], and the N0 and N1 domains of the T3SS secretin fit well into a cryo-electron microscopy density of the Salmonella typhimurium needle complex [16], it might be that the N0 and N1 domains of these related secretins adopt different mutual orientations in the assembled T2SS and T3SS in vivo as observed in crystals. Obviously further studies are required to confirm this hypothesis where it also should be kept in mind that secretins are dynamic proteins and multiple orientations of N-terminal secretin domains might transiently occur during the secretion process [10].

The GspC–GspD Crystal Structure and Functional Studies

The crystal structure indicates that a number of residues are critical for the interactions of ETEC GspC and GspD (Figure 4).

---

**Figure 5. Simultaneous substitution of Ile18 and Asn22 in V. cholerae GspD results in inactivation of protease secretion by V. cholerae.** (A) V. cholerae wild-type or gspD mutant strain (JgspD) containing either pMMB or pGspD variants were grown overnight in LB. Culture supernatants were separated from cells by centrifugation and tested for the presence of extracellular protease. The rate of hydrolysis was obtained from three independent experiments, and the results are presented with standard error. Both the pMMB vector control and pGspD{I18R/N22Y} were below the limits of detection. (B) V. cholerae gspD cells containing either pMMB or pGspD variants were disrupted and subjected to SDS-PAGE and immunoblotting with anti-GspD antibodies to determine the relative level of expression. The positions of molecular mass markers are shown. Arrow indicates monomeric GspD and arrowhead indicates multimeric GspD.

doi:10.1371/journal.ppat.1002228.g005

---

**Figure 6. Differential localization in V. cholerae of GFP-GspC in the presence of GspD{I18R/N22Y}.** Localization of chromosomally expressed GFP-GspC was examined in wild-type and gspD mutant backgrounds by fluorescence microscopy. GFP-GspC displayed a continuous membrane localization in the gfp-gspC gspD{I18R/N22Y} strain (second panel) compared to the wild-type background (first panel). Punctate fluorescence was restored when the gfp-gspC gspD strain was complemented with GspD on a plasmid (third panel). Expression of GspD{I18R/N22Y} in the gfp-gspC gspD{I18R/N22Y} strain resulted in membrane localization similar to the pMMB vector control (fourth panel).

doi:10.1371/journal.ppat.1002228.g006
Moreover, these residues are conserved in the family sequence alignment (Figure S4). As many mutants and other useful reagents have already been generated and developed for studies of the T2SS in *V. cholerae*, subsequent probing of the importance of these residues for the interaction was carried out in three different ways using *V. cholerae* GspC and GspD homologues. The two-hybrid studies showed that substitutions Val118Arg and Val129Arg in *V. cholerae* GspC, and Asn22Arg in F*GspD*, abrogated the interaction between GspCHR and GspD10-N1-N2 from *V. cholerae* (Table 2). The secretion of protease by *V. cholerae* was also greatly diminished by substitutions Ile18Arg/Asn22Tyr in full length *V. cholerae* GspD (Figure 5). Finally, the same Ile18Arg/Asn22Tyr variant of *V. cholerae* GspD altered the distribution of full-length F*GspD* in the inner membrane of *V. cholerae*, possibly by interfering with normal assembly of the inner membrane platform of the T2SS (Figure 6). Taking all data together, we conclude that the substitutions altering the interface of GspC with GspD in *V. cholerae* affect the interactions of GspC with GspD as demonstrated both in a bacterial two-hybrid system and by analysis of protease secretion by the T2SS in *V. cholerae*.

Interactions between GspC and GspD from *D. dadantii* have been recently investigated [40]. This study confirmed the interactions between GspCHR and the N-terminal domains of GspD reported earlier for *V. vulnificus* homologs [38]. A GST-fusion of residues 139–158 of *D. dadantii* GspC (corresponding to residues 168–187 in ETEC GspC) co-purified with both *D. dadantii* GspD10-N1-N2 and *D. dadantii* GspD11-N2-N3 [40]. The 139–158 residues of *D. dadantii* GspC were therefore designated as secretin interacting peptide (SIP). In a homology model of *D. dadantii* GspC (corresponding to residues 168–187 in ETEC GspC) the residues 139–158 of the *D. dadantii* GspC were therefore designated as secretin interacting peptide (SIP). In a homology model of *D. dadantii* GspC with PilP of *V. vulnificus* [56] located far from the interface (Figure S6). It appears that this segment forms an anti-parallel pair of B-stands, B5 and B6, in the ETEC GspCHR crystal structure, with B6 at the surface and B5 located between strands B6 and B4 (Figure S6). Furthermore, the substitutions introduced into the *D. dadantii* GspC 139–158 fragment had no effect on the interaction with *D. dadantii* GspD, whereas one substitution, Val143Ser, prevented *D. dadantii* GspC interaction with *D. dadantii* GspD11-N2-N3 [40]. The same substitution, when introduced into full length *D. dadantii* GspC, also interfered with secretion in *D. dadantii*. We also mapped these substitutions onto the homology models of the *D. dadantii* GspC–GspD complex and it is clear that none of them are buried in the *D. dadantii* GspC–GspD interface (Figure S6). The only substitution that had an effect on secretion, Val143Ser, replaces a buried hydrophobic residue in the core of *D. dadantii* GspC with a polar residue that would likely be detrimental to the HR domain stability. This is in agreement with the finding that this substitution in GST-*D. dadantii* GspC129-272 resulted in a protein that is degraded in the cells [40]. A more conservative Val143Ala substitution in full length *D. dadantii* GspC appeared to largely support secretion of pectinases, in agreement with the less drastic change of the nature of the side chain, which could result in a larger proportion of properly folded protein than in the case of the Val143Ser variant. Therefore, the ETEC GspCHR–GspD10-N1-N2 complex structure explains several experimental results of the studies on *D. dadantii* GspC–GspD interactions [40]. The observations that a GST-fusion of the *D. dadantii* GspC 139–158 fragment is capable of interacting with fragments of the secretin in the absence of both the rest of the HR domain and the rest of the secretin, and of interfering with pectinase secretion when over-expressed in wild type *D. dadantii*, are difficult to interpret precisely. Additional studies are required to show that such interactions are not the result of non-specific interactions, possibly due to exposed hydrophobic residues of the peptide which are normally buried in the complete HR domain.

### T2SS Stoichiometry

The implications of the GspCHR–GspD10 interactions unravelled by our studies for the architecture of the T2SS are intriguing. The three new structures in the current paper show that one GspCHR domain interacts with one GspD10 domain, which suggests a 1:1 ratio of GspC and GspD in the assembled T2SS. Since the stoichiometry of full length GspC and GspD has not been established yet in the context of a functional T2SS, it is of interest to see if the current complex of GspCHR–GspD10 is compatible with the dodecameric ring of GspD10-N1 derived recently by a combination of crystallographic and electron microscopy studies [20,24]. Superimposing the GspCHR–GspD10 complex twelve times onto the N0-domains of the GspD10-N1 ring results in a double ring structure where the GspCHR subunits added do not interfere with the formation of the GspD10-N1 ring. Although this procedure does result in some clashes between the subunits of the GspCHR ring, specifically between residues of the B2–B3 loop of one subunit and residues just prior to β6 in a neighboring subunit, small conformational changes in these loops, or minor adjustments in the mutual orientation of domains in the GspD10 ring, or both, might alleviate these close contacts. If this would be the case, the GspD dodecamer would interact with twelve GspC subunits in the assembled T2SS (Figure 7A). Alternatively, only alternating GspD subunits of the dodecameric secretin might interact with GspCHR, obviously removing close contacts between the then well separated GspCHR subunits. In this case, the GspD dodecamer would interact with six GspC subunits (Figure 7B).

These two alternatives for the interface of the outer membrane complex and the inner membrane platform can be combined with previous studies on the T2SS even though the ratio between GspC and the other components of the inner-membrane platform complex is currently unknown. Yet, the following observations are of interest for the T2SS stoichiometry puzzle:

1. the secretin ATPase GspE of the T2SS has been suggested to be a hexamer [50,51];
2. the cytoplasmic domain of the inner membrane T2SS protein GspL forms a 1:1 complex with GspE [52];
3. homology of GspM and of the cytoplasmic domain of GspL from the T4PS have been reported to form heterodimers [53,54];
4. there are a few cases of gene fusion of the T4PS proteins PilP and PilO (e.g. *Pseudomonas putida* PilO-PilP, Uniprot entry Q88CU9) in the T4PS. PilP is a GspCHR homolog (Figure 3) and PilO is proposed to be a homolog of the inner membrane protein GspM from the T2SS [54,55]. The presence of PilO-PilP fusions may imply a 1:1 stoichiometry of these proteins in the T4PS and, in view of the homology between the T4PS and the T2SS, a GspM-GspC ratio of 1:1 in the T2SS as well.

These four observations suggest that GspE, GspL, GspM and GspC might be present in an equimolar ratio in the inner membrane platform. In view of the likely hexameric nature of GspE, this implies the presence of six subunits of each of these proteins in the assembled T2SS. If the GspD dodecamer would interact with six GspC subunits (Figure 7B), then this arrangement would agree well with six subunits each of GspC, GspL, GspM and GspE in the inner membrane platform. If a GspD dodecamer, however, would interact with twelve GspC subunits in the assembled T2SS (Figure 7A), then, a stoichiometry mismatch is likely to occur somewhere along the GspC–GspL–GspM–GspE chain of interactions in the inner membrane platform. This could be possible in spite of the evidence in points (i) to (iv) above for an equimolar ratio of these four proteins in the T2SS since e. g. points (ii) and (iii) are rather indirect and derived from observations on T4PS proteins. Clearly, the stoichiometry of the T2SS remains a fascinating topic for further studies, where the number of GspF subunits, the only T2SS protein which spans the inner membrane multiple times, also remains to be determined.
Implications for Exoprotein Secretion

Another major outstanding question is the recognition of exoproteins by the T2SS. Interestingly, the inner diameter of the dodecameric GspCHR–GspDN0-N1 double ring is ~68 Å, which implies that a large exoprotein like the cholera toxin AB5 heterohexamer [56] just fits into this ring (Figure 7C). This is in agreement with recent electron microscopy studies which indicate that the B-pentamer of cholera toxin can bind to the entrance of the GspD periplasmic vestibule [57]. The periplasmic domains of GspD and of GspC have been implicated in this crucial exoprotein recognition function [34,46,57,58,59], but the specific details of exoprotein–T2SS interactions remain to be uncovered. The accumulation of recent structural and biochemical data provides a platform for asking increasingly precise questions regarding the many remaining mysteries still pertaining to the architecture and mechanism of the sophisticated T2SS.

Methods

Expression and Purification of GspC and GspD for Crystallization

ETEC GspDNO-N1-N2 (residues 1–237; numbering corresponds to mature protein sequence) was expressed and purified as described [24]. The DNA sequence corresponding to residues 1–165 of ETEC GspD was PCR amplified and cloned into the pCDF-NT vector to obtain a GspDNO-N1 expression construct. pCDF-NT is a modified pCDF-Duet1 vector (Novagen) encoding an N-terminal His6-tag sequence and a TEV protease cleavage site. The DNA sequence corresponding to residues 122–186 of ETEC GspC was PCR amplified and cloned into a pCDF-NT vector to obtain a GspC HR expression construct. A lanthanide binding tag (LBT) was introduced into GspD NO-N1 construct in order to assist with crystallographic phase determination and promote crystal formation. In order to decrease the flexibility of the LBT, we introduced it into the loop between two adjacent β-strands rather than at the termini. The design was based on the crystal structure of ubiquitin with the double LBT (PDB 2OJR) [41] where two β-strands flank one of the LBT. The LBT sequence YIDTNNDGYIEGDEL was inserted between residues Met70 and Val74 of GspDN0 (Figure S4B) using the polymerase incomplete primer extension method [60]. While this manuscript was in preparation, a similar approach for the LBT insertion was successfully applied to a model protein, interleukin-1β [61].

GspDN0-N1 was expressed at 25°C in BL21(DE3) cells (Novagen) in LB medium containing 30 µg mL–1 streptomycin. Protein production was induced with 0.5 mM IPTG. Cells were harvested 3 h after induction. GspDNO-N1 variants with or without LBT were purified by Ni-NTA agarose (Qiagen) chromatography followed by His6-tag cleavage with TEV protease; a second Ni-NTA chromatography to remove His6-tag, uncleaved protein and His-tagged TEV protease; and a final size-exclusion chromatography using Superdex 75 column (GE Healthcare). GspC HR was expressed and purified under same conditions as GspD NO-N1. The proteins were concentrated, flash-frozen [62] and stored at −80°C. Se-Met-labeled proteins were expressed using metabolic inhibition of methionine biosynthesis [63] and purified using the protocols for native proteins.
Purification of Nanobodies

The nine nanobodies generated against ETEC GspD<sub>100-112</sub> were expressed and purified as described previously [24].

Crystallization, Data Collection and Structure

Determination

ETEC GspCHR<sub>1</sub>, GspD<sub>100-112</sub> and individual nanobodies were mixed at 1:1 molar ratio, concentrated to 4–8 mg/ml total protein concentration and subjected to crystallization conditions screening by the vapor diffusion method at 4 or 21°C. The crystallization conditions were identified using SaltRx (Hampton Research) and JCSG+(Qiagen) screens. The complex of GspCHR–GspD<sub>100-112</sub>–Nb3 was crystallized in 1.2 M lithium sulfate, 0.1 M Tris-HCl pH 7 at 4°C. The crystals were gradually transferred to precipitant solution supplemented with 30% glycerol and flash-frozen in liquid nitrogen. Initial crystals diffracted to 5.5 Å resolution and optimized crystals with Se-Met substituted GspD<sub>100-112</sub> showed improved diffraction to 4.6 Å. Data were processed and scaled using XDS [64]. The structure was solved by molecular replacement using Phaser [63]; the search models included the GspD<sub>100-112</sub> structure (PDB 3FZJ) [24], a camellid antibody fragment (PDB 1QD0) [66], and a homology model of GspC<sub>1</sub> obtained using the I-TASSER server [67] and the N. meningitidis PDB structure as template (PDB 2IVW) [42]. The N2 domain of GspD could not be located in the electron density maps due to statistical disorder (Figure S2).

The complex of GspCHR<sub>1</sub>, GspD<sub>100-112</sub> with an engineered LBT in the GspD<sub>100</sub> domain was crystallized in 0.9 M magnesium sulfate, 0.1 M bis-tris propane pH 7.0 at 21°C. The crystals were transferred to precipitant solution supplemented with 20% ethylene glycol and flash-frozen in liquid nitrogen. The structure of the GspCHR<sub>1</sub>, GspD<sub>100-112</sub> complex was solved by molecular replacement using Phaser and rebuilt using Buccaneer [68] and Coot [69]. The metal binding site of the LBT appears to be important for the crystal structure.

Cloning of Sequences Encoding Soluble Cytoplasmic Domains of VcGspC and VcGspD into Two Hybrid Vectors

A DNA sequence encoding residues 53–305 of VcGspC (AAA58784.1) was amplified by PCR using the primers EpsCXF and EpsCHIIIR adding XhoI (Leu-Glu frame) and a stop codon.-HindIII sites at the 5′ and 3′ ends respectively. This product was cloned in frame after the T25 domain in place of ARF6 in pXC752536f (pCT25ARF6 from [48]) with a vector XhoI site deleted) to generate pCT25VcGspC. Similarly the coding sequence residues 25–294 of VcGspD (AAA58785.1) was amplified with primers EpsDNdeIF and EpsDClA which add XhoI (and Met codon) and ClaI (Ser-Met frame) sites at the 5′ and 3′ ends respectively; this PCR product was cloned in frame after the T25 domain in place of ARF6 in pXC752536f (pCT25VcGspD from [48]) with a vector XhoI site deleted to generate pCT25VcGspD. The primers sequence information is available upon request. Specific mutations in the eps gene domains (encoding GspC or GspD) in pVcGspC and pVcGspD were performed by SOE-PCR [77] or by subcloning of a PCR fragment performed with a restriction site containing mutagenic primer and a vector primer, followed by recloning into the parental vector. All clones were verified in-frame and correct by DNA sequencing to ensure no additional PCR-generated mutations.

Two Hybrid Interaction of VcGspC and VcGspD Domain Interactions

Interaction between protein domains was detected by the ability of fusion proteins containing the enzymatically inactive T18 and T25 fragments of adenylate cyclase toxin from Bordetella pertussis to confer adenylate cyclase activity (and the ability to ferment maltose and form red colonies on maltose-MacConkey plates) to a ca17 mutant E. coli strain as described previously [48]. E. coli DC8F<sup>+</sup> was a ca17::Km<sup>+</sup> derivative of the strain MM294 (amp<sup>R</sup>, lacI<sup>R</sup>, thi<sup>R</sup>, xyl<sup>R</sup>) and the DH<sup>10</sup> strain was transformed into E. coli DC8F<sup>+</sup>, and transformants were selected on LB-Cm or LB-Ap plates, respectively. Each of the resulting transformants formed white colonies when streaked onto maltose MacConkey plates and incubated at 30°C. In contrast, when both plasmids were transformed together into E. coli DC8F<sup>+</sup>, the resulting transformants formed red colonies when streaked onto maltose MacConkey plates, demonstrating a productive protein-protein interaction between the VcGspC and VcGspD domains of the T25–VcGspC and VcGspD–T18 fusion proteins, bringing together the T18 and T25 fragments to form active cyclase.

Cloning of Sequences Encoding Soluble Cytoplasmic Domains of VcGspC and VcGspD into Two Hybrid Vectors

A DNA sequence encoding residues 53–305 of VcGspC (AAA58784.1) was amplified by PCR using the primers EpsCXF and EpsCHIIIR adding XhoI (Leu-Glu frame) and a stop codon.-HindIII sites at the 5′ and 3′ ends respectively. This product was cloned in frame after the T25 domain in place of ARF6 in pXC752536f (pCT25ARF6 from [48]) with a vector XhoI site deleted) to generate pCT25VcGspC. Similarly the coding sequence residues 25–294 of VcGspD (AAA58785.1) was amplified with primers EpsDNdeIF and EpsDClA which add XhoI (and Met codon) and ClaI (Ser-Met frame) sites at the 5′ and 3′ ends respectively; this PCR product was cloned in frame after the T25 domain in place of ARF6 in pXC752536f (pCT25VcGspD from [48]) with a vector XhoI site deleted to generate pCT25VcGspD. The primers sequence information is available upon request. Specific mutations in the eps gene domains (encoding GspC or GspD) in pVcGspC and pVcGspD were performed by SOE-PCR [77] or by subcloning of a PCR fragment performed with a restriction site containing mutagenic primer and a vector primer, followed by recloning into the parental vector. All clones were verified in-frame and correct by DNA sequencing to ensure no additional PCR-generated mutations.

Generation of VcGspD Mutants

The AggD strain of V. cholerae, a gsp-gspC AggD strain, and the complementing pMMB-gspD plasmid were constructed previously [39]. Mutations were introduced in the gspD gene of V. cholerae with the QuikChange II site-directed mutagenesis kit (Stratagene) using pBAD-gspD as a template. Primers used for the site change in gspD<sub>110R</sub> and gspD<sub>222Y</sub> were 5′-GAATT-TATCACTCGTGTTGGACCGAAATC-3′, 5′-GATTGGCGCTCCACACGATTGAAATTC-3′ and their reverse complements, respectively. gspD<sub>110R</sub>-gspD<sub>222Y</sub> was then constructed using pBAD-gspD<sub>110R</sub> as a template and the above primers specific for the gspD<sub>222Y</sub> site change. All mutations were verified by sequencing. Following sequencing, the gspD variants of V. cholerae obtained were cloned into the low-copy-vector pMMB67 using restriction enzymes BamHI and SphI.
Detection of Secreted Protease Activity

*V. cholerae* cultures were grown overnight at 37°C in Luria broth supplemented with 100 μg/ml 1 thymine, 200 μg/ml carbenicillin, and 20 μl IPTG and centrifuged to separate the supernatant and cellular material. The supernatants were centrifuged once more, and the protease activity was measured as described previously [76].

Microscopy

*Cultures of V. cholerae* were grown overnight at 37°C in M9 medium containing 0.4% casamino acids, 0.4% glucose, and 100 μg/ml thyme; diluted 50-fold into fresh medium; and grown to mid-log phase before observation. Plasmids were maintained with 50 and 200 μg/ml carbenicillin in log-phase and overnight cultures, respectively. Plasmid expression was induced with IPTG as described above. For fluorescence microscopy of live cells, cultures were mounted on 1.5% low-melting temperature agarose pads prepared with M9 glucose medium. All microscopy was performed with a Nikon Eclipse 90i fluorescence microscope equipped with a Nikon Plan Apo VC100 (1.4 numerical aperture) oil immersion objective and a Cool SNAP HQ2 digital camera. Captured images were analyzed with NIS-Elements imaging software (Nikon).

Accession Numbers

Atomic coordinates and structure factors have been deposited in the Protein Data Bank (http://www.pdb.org) with accession code 3OSS.

Supporting Information

**Figure S1** The lanthanide binding tag (LBT) in the GspCHR–GspDN0-N1 crystal structure. (A) Stereoview of the LBT in the GspCHR–GspDN0-N1 crystal structure. The Cα- weighted 2Fo-Fc electron density map is displayed as a grey mesh at the 1 σ level. The Ca2+ ion is shown as an orange sphere; a coordinating water molecule as a red sphere. (B) The LBT makes several crystal contacts in the lattice. (TIF)

**Figure S2** Crystal structure of the GspCHR–GspDN0-N1-N2–Nb3 ternary complex. (A) SDS-PAGE analysis of crystals. Lane 1, molecular weight standards; lane 2, purified GspCH; lane 3, purified GspDN0-N1-N2; lane 4, purified Nb3; lane 5, GspCHR–GspDN0-N1-N2–Nb3 complex before crystallization; lane 6, drop medium containing 0.4% casamino acids, 0.4% glucose, and 100 μg/ml thymine; diluted 50-fold into fresh medium; and grown to mid-log phase before observation. Plasmids were maintained with 50 and 200 μg/ml carbenicillin in log-phase and overnight cultures, respectively. Plasmid expression was induced with IPTG as described above. For fluorescence microscopy of live cells, cultures were mounted on 1.5% low-melting temperature agarose pads prepared with M9 glucose medium. All microscopy was performed with a Nikon Eclipse 90i fluorescence microscope equipped with a Nikon Plan Apo VC100 (1.4 numerical aperture) oil immersion objective and a Cool SNAP HQ2 digital camera. Captured images were analyzed with NIS-Elements imaging software (Nikon).

**Figure S5** The structure of GspDN0-N1 is virtually the same in the GspCHR–GspDN0-N1 and GspDN0-N1-N2-Nb7 structures. A stereo view of a superposition of GspDN0-N1 from the GspCHR–GspDN0-N1 complex (cyan) and the GspD N0-N1-N2-Nb7 complex (purple, PDB 3EZJ) [24]. The superposition is based on the N0 domain only (r.m.s.d. 0.49 Å for 72 Ca atoms). The mutual orientation of the N0 and N1 domains is very similar indeed. (TIF)

**Figure S6** Analysis of the *Dickeya dadantii* GspCHR–GspDN0-N1 complex. (A) A homology model of the D. dadantii (previously Erwinia chrysanthemi) GspC–GspD complex. The structures of the HR domain of *DdGspC* (light green) and the N0-N1 domains of *DdGspD* (light blue) were obtained by homology modeling based on our new structure of the ETEC GspCHR–GspDN0-N1 complex (grey) as template, using the SWISS-MODEL server (http://swissmodel.expasy.org/) [79]. (B) Mutations of *DdGspC*. The residues in the interface of *DdGspC*-GspD in the homology model are shown as sticks. A previously suggested interaction region SIP (secretin interacting peptide) that corresponds to residues 139–159 is highlighted in orange [40]. The residues which have been subjected to mutational analysis (R142, V143, V144, R150, E152 and Y157) are shown as sticks and labeled. The mutant *DdGspC* proteins R142I, V144A, R150L, E152A and Y157A fully supported secretion of pectinases in *D. dadantii*. Note that, in contrast to the other residues, V143 is completely buried in the model and the substitution V143S leads to decreased secretion [40]. For further discussion see main text. (C) Sequence alignments of GspCHR and GspDN0 from ETEC and *D. dadantii*. Secondary structure elements are shown above alignment according to ETEC GspCHR–GspDN0 crystal structure. A previously suggested interaction region SIP that corresponds to residues 139–159 is indicated by an orange bar. The residues which have been subjected to mutational analysis (R142, V143, V144, R150, E152 and Y157) are highlighted by circles. (TIF)
Acknowledgments

We thank Jonathan Kay for maintenance of equipment and computers, and the support staff of beamline 9-2 of SSRF, for assistance during data collection. We thank Michael Bagdasarian for the assistance of a nanobody. Structure determination of the EpsI:EpsJ pseudopilin heterodimer from Vibrio cholerae. We wrote the paper: KVK TLJ MS WGJH. All authors commented on the manuscript.

References

1. Sandkvist M (2001) Type II secretion and pathogenesis. Infect Immun 69: 3523–3535.

2. Giancotti NP (2003) Type II secretion: a protein secretion system for all seasons. Trends Microbiol 11: 581–588.

3. DebRoy S, Dao J, Soderberg M, Rossier O, Giancotti NP (2006) Legionella pneumophila type II secretion reveals unique exoproteins and a chitinase that promotes bacterial persistence in the lung. Proc Natl Acad Sci U S A 103: 19146–19151.

4. Sikora AE, Zielke RA, Lawrence DA, Andrews PC, Sandkvist M (2011) Proteomic analysis of the Vibrio cholerae type II secretion system reveals new proteins including three related serine proteases. J Biol Chem 286: 16353–16366.

5. Sandkvist M, Michel LO, Hough LP, Morales VM, Bagdasarian M, et al. (1997) General secretion pathway (gsp) genes required for toxin secretion and outer membrane biogenesis in Vibrio cholerae. J Bacteriol 179: 6994–7003.

6. Tauschek M, Gorrell RJ, Strugnell RA, Robins-Browne RM (2002) Identification of the protein secretory pathway for the secretion of heat-stable enterotoxins by an enterotoxigenic strain of Escherichia coli. Proc Natl Acad Sci U S A 99: 7066–7071.

7. Hiroi TR, Holmgren J (1987) Conformation of secreted protein across bacterial outer membranes: a study of enterotoxin translocation from Vibrio cholerae. Proc Natl Acad Sci U S A 84: 7418–7422.

8. Filloux A (2004) The underlying mechanisms of type II protein secretion. Mol Microbiol 53: 1710–1719.

9. Johnson TL, Abendroth J, Hol WGJ, Sandkvist M (2006) Type II secretion: from structure to function. FEBS Lett 585: 173–186.

10. Korokov KO, Gonzén T, Hol WGJ (2011) Secretion: dynamic channels for protein transport across membranes. Trends Biochem Sci 36: 433–443.

11. Opalka N, Beckmann R, Bousset N, Simon MN, Russell M, et al. (2003) Structure of the filamentous phage β IV middletail by cryo-electron microscopy. J Mol Biol 325: 461–470.

12. Collins RF, Ford RG, Kimmitt AO, Rosen LO, Tonjum T, et al. (2003) Three-dimensional structure of the Neisseria meningitidis secretin PSIq determined from negative-stain transmission electron microscopy. J Bacteriol 185: 2611–2617.

13. Jain S, Moscicka KB, Bes MP, Pachalie E, Stuart MC, et al. (2011) Structural characterization of outer membrane components of the type IV pili system in pathogenic Neisseria. PLoS One 6: e16624.

14. Burkhardt J, Vonck J, Averhoff B (2011) Structure and function of PSIq, a secretin of the DNA transporter from the thermophilic bacterium Thermus thermophilus HB8. J Biol Chem 286: 9977–9984.

15. Hodgkinson JL, Horley A, Sahlad D, Simon M, Johnson S, et al. (2009) Three-dimensional reconstruction of the Shigella T3SS transmembrane regions reveals 12-64α symmetry and novel features through structure. Nat Struct Mol Biol 16: 415–425.

16. Schratt O, Marlovits TC (2011) Three-dimensional model of Nele Buys for the selection, expression and purification of the nanobodies. PLoS One 6: e16624.

17. Korokov KO, Knüppel B, Sandkvist M (2009) Docking and assembly of the type II secretion complex of Vibrio cholerae. J Bacteriol 191: 3149–3161.

18. Ayers M, Howell PL, Burrows LL (2010) Architecture of the type II secretion and type IV pilus machineries. Future Microbiol 5: 1203–1218.

19. Chami M, Guilvout I, Gregorini M, Remigy HW, Muller SA, et al. (2005) Structure and functional studies of XecP, a component involved in general secretory pathway-dependent protein secretion in Pseudomonas aeruginosa. FEBS Lett 545: 175–180.

20. Gerard-Vincent M, Robert V, Ball G, Bleves S, Michel GP, et al. (2002) Identification of XecP domains that confer functionality and specificity to the Pseudomonas aeruginosa type II secretion apparatus. Mol Microbiol 44: 1651–1665.

21. Logue TF, Fries M, Wang X, Pickersgill RW, Shchelkoch VE, et al. (2010) A 20-residue peptide of the inner membrane protein OutC mediates interaction with two distinct sites of the outer membrane secretin OutD and is essential for the functional type II secretion system in Erwinia chrysanthemi. Mol Microbiol 74: 145–157.

22. Korokov KO, Knüppel B, Sandkvist M, Hol WGJ (2006) Structural and functional studies of EpsC, a crucial component of the type 2 secretion system of Escherichia coli. Mol Microbiol 58: 1226–1237.

23. Lybarger SR, Johnson TL, Gray MD, Sikora AE, Sandkvist M (2009) Identification of the T2SS secretin GspD with the assistance of a nanobody. Structure 17: 255–265.

24. Rasmussen SG, Choi HH, Fung JJ, Pardon E, Casarosa P, et al. (2011) Structure of a nanobody-stabilized active state of the βl adrenoceptor. Nature 469: 175–180.

25. Garcia-Herreros A, Vogel HJ (2005) Nuclear magnetic resonance solution structure of the periplasmic signalling domain of the TonB-dependent outer membrane transporter FecA from Escherichia coli. Mol Microbiol 58: 1226–1237.

26. Leiman PG, Basler M, Ramagopala UA, Bonanno JB, Sauder JM, et al. (2009) Type VI secretion apparatus and phage tail-associated complex proteins share a common evolutionary origin. Proc Natl Acad Sci U S A 106: 4145–4159.

27. Nagano N, Kuhori T, Kinoshita M, Ikeda K, Nagai H (2010) Crystal structure of Legionella DosD: insights into the relationship between type IVB and type II/III secretion systems. J Mol Biol 399: 317–329.

28. Souza DP, Andrade MO, Alvarez-Martinez CE, Arantes GM, Farah CS, et al. (2011) A Component of the Xanthomonadaceae Type IV Secretion System Combines a VexB Motif with a N0 Domain Found in Outer Membrane Transport Proteins. PLoS Pathog 7: e1002035.

29. Kanamaru S, Leiman PG, Kostyukenko VA, Chipman PR, Meyyazhagan VV, et al. (2002) Structure of the cell-puncturing device of bacteriophage T4. Nature 415: 533–537.

30. Valvère R, Edwards L, Rogan L (2000) Structure and function of KH domains. FEBS Lett 475: 2712–2726.

31. Souza DP, Andrade MO, Alvarez-Martinez CE, Arantes GM, Farah CS, et al. (2011) A Component of the Xanthomonadaceae Type IV Secretion System Combines a VexB Motif with a N0 Domain Found in Outer Membrane Transport Proteins. PLoS Pathog 7: e1002035.

32. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystallographic data: KVK TLJ MGJ RKH MS WGJH. Wrote the paper: KVK TLJ MS WGJH. Conceived and designed the experiments: KVK JS RKH MS WGJH.
Camberg JL, Johnson TL, Patrick M, Abendroth J, Hol WGJ, et al. (2007) Synergistic stimulation of EpsE ATP hydrolysis by EpsL and acidic phospholipids. EMBO J 26: 19–27.

Patrick M, Korolok KV, Hol WGJ, Sandkvist M (2011) Oligomerization of EpsE coordinates residues from multiple subunits to facilitate ATPase activity. J Biol Chem 286: 10378–10386.

Abendroth J, Murphy P, Sandkvist M, Bagdasarian M, Hol WGJ (2005) The X-ray structure of the type II secretion system complex formed by the N-terminal domain of EpsE and the cytoplasmic domain of EpsM of *Vibrio cholerae*. J Mol Biol 348: 845–855.

Ayers M, Sampaleanu LM, Tamamian S, Koo J, Harvey H, et al. (2009) PIM/ N/O/P proteins form an inner membrane complex that affects the stability of the *Pseudomonas aeruginosa* type IV pilus secretin. J Mol Biol 394: 129–142.

Sampaleanu LM, Bonanno JB, Ayers M, Koo J, Tamamian S, et al. (2009) Periplasmic domains of *Pseudomonas aeruginosa* PIN and PICO form a stable heterodimeric complex. J Mol Biol 384: 143–159.

Abendroth J, Rice AE, McLuskey K, Bagdasarian M, Hol WGJ (2004) The crystal structure of the periplasmic domain of the type II secretion system protein EpsM from *Vibrio cholerae*: the simplest version of the ferredoxin fold. J Mol Biol 339: 585–596.

O’Neal CJ, Amaya EI, Jolding MG, Holmes RK, Hol WGJ (2004) Crystal structures of an intrinsically active cholera toxin mutant yield insight into the toxin activation mechanism. Biochemistry 43: 3772–3782.

Reichow SL, Korolok KV, Gouen M, Sun J, dela Rosa J, et al. (2011) The binding of cholera toxin to the periplasmic vestibule of the type II secretion channel. Channels 5: 215–218.

Bouley J, Condemine G, Shevchik VE (2001) The PDZ domain of OutC and the cytoplasmic domain of EpsL of *Vibrio cholerae*. J Mol Biol 210: 129–142.

Guilvout I, Hardie KR, Sauvonnet N, Pugsley AP (1999) Genetic dissection of the pathway of *Escherichia coli* N-terminal region of OutD determine the secretion specificity of the type II out channel. Channels 5: 215–218.

Erwinia chrysanthemi. J Mol Biol 308: 205–219.

Guilvout I, Hardie KR, Sauvonnet N, Pugsley AP (1999) Genetic dissection of the pathway of *Escherichia coli* N-terminal region of OutD determine the secretion specificity of the type II out channel. Channels 5: 215–218.

Klock HE, Koersena EJ, Knuh MW, Lesley SA (2008) Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscanning to accelerate structural genomics efforts. Proteins 71: 902–994.

Barthelmes K, Reynolds AM, Peisach E, Jonker HR, DeNunzio NJ, et al. (2011) Engineering encodable lanthanide-binding tags into loop regions of proteins. J Am Chem Soc 133: 808–819.

Deng J, Davies DR, Wiedchaai G, Wu M, Hol WGJ, et al. (2004) An improved protocol for rapid freezing of protein samples for long-term storage. Acta Crystallogr D Biol Crystallogr 60: 203–204.