



Published in final edited form as:

*J Struct Biol.* 2014 September ; 187(3): 223–235. doi:10.1016/j.jsb.2014.07.006.

## Crystal Structure of the Full-Length ATPase GspE from the *Vibrio vulnificus* Type II Secretion System in Complex with the Cytoplasmic Domain of GspL

Connie Lu<sup>1</sup>, Konstantin V. Korotkov<sup>1,#</sup>, and Wim G. J. Hol<sup>\*</sup>

Department of Biochemistry and Biomolecular Structure Center, University of Washington, Seattle, Washington 98195

### Abstract

The type II secretion system (T2SS) is present in many Gram-negative bacteria and is responsible for secreting a large number of folded proteins, including major virulence factors, across the outer membrane. The T2SS consists of 11-15 different proteins most of which are present in multiple copies in the assembled secretion machinery. The ATPase GspE, essential for the functioning of the T2SS, contains three domains (N1E, N2E and CTE) of which the N1E domain is associated with the cytoplasmic domain of the inner membrane protein GspL

Here we describe and analyze the structure of the GspE•cyto-GspL complex from *Vibrio vulnificus* in the presence of an ATP analog, AMPPNP. There are three such ~83 kDa complexes per asymmetric unit with essentially the same structure. The N2E and CTE domains of a single *V. vulnificus* GspE subunit adopt a mutual orientation that has not been seen before in any of the previous GspE structures, neither in structures of related ATPases from other secretion systems. This underlines the tremendous conformational flexibility of the T2SS secretion ATPase.

Cyto-GspL interacts not only with the N1E domain, but also with the CTE domain and is even in contact with AMPPNP. Moreover, the cyto-GspL domains engage in two types of mutual interactions, resulting in two essentially identical, but crystallographically independent, “cyto-GspL rods” that run throughout the crystal. Very similar rods are present in previous crystals of cyto-GspL and of the N1E•cyto-GspL complex. This arrangement, now seen four times in three entirely different crystal forms, involves contacts between highly conserved residues suggesting a role in the biogenesis or the secretion mechanism or both of the T2SS.

© 2014 Elsevier Inc. All rights reserved.

\*Correspondence to: Wim G. J. Hol, Department of Biochemistry and Biomolecular Structure Center, University of Washington, Seattle, Washington 98195. Phone: (206) 685-7044, Fax: (206) 685-7002, wghol@u.washington.edu.

<sup>1</sup>These authors contributed equally to this study.

<sup>#</sup>Present address: Department of Molecular & Cellular Biochemistry, University of Kentucky, Lexington, Kentucky 40536.

**Accession numbers:** The atomic coordinates and structure factors (code 4PHT) have been deposited in the Protein Data Bank (<http://www.pdb.org/>).

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Keywords

Assembly ATPase; Protein Secretion; EpsE; EpsL; T2SS; T4PS

---

## 1. Introduction

Secretion of proteins into the extra-cellular milieu is important for many pathogenic and non-pathogenic bacteria, which have developed an impressive variety of often complex multi-protein assemblies to perform this task. One of these protein secretion machineries is the sophisticated type II secretion system (T2SS) that spans the inner and outer membranes from many Gram-negative bacteria (Korotkov et al., 2012). The T2SS is highly relevant for the pathogenicity of several major pathogens since it translocates major virulence factors in a folded form from the periplasm to the extracellular milieu. Examples of bacterial human pathogens where the T2SS plays an important role include:

- i. *Vibrio cholerae*, where the T2SS secretes the heterohexameric AB<sub>5</sub> cholera toxin (CT) and ~20 other proteins (Hirst et al., 1984; Sikora et al., 2011). *V. cholerae* is responsible for estimated 100,000-120,000 deaths per year, mainly in low-income countries and disaster areas (<http://www.who.int/mediacentre/factsheets/fs107/en/>);
- ii. Enterotoxigenic *E. coli* (ETEC), where the T2SS translocates heat-labile enterotoxin (LT) (Hirst and Holmgren, 1987), a close structural and functional homolog of CT (Merritt and Hol, 1995). ETEC are an extremely important cause of diarrhea in the developing world (Qadri et al., 2005; Wenneras and Erling, 2004), and also are the most common cause of travelers' diarrhea (Steffen et al., 2005);
- iii. Enterohaemorrhagic *E. coli* (EHEC), which can cause severe foodborne disease, and even life-threatening renal failure in children and elderly. The T2SS deletion mutant of EHEC shows defects in colonization. In addition, the zinc metalloprotease StcE and the metal binding protein YodA, which are crucial for EHEC adherence to host cells, are secreted by the T2SS encoded on plasmid pO157 (Goldwater and Bettelheim, 2012; Ho et al., 2008; Toshima et al., 2007);
- iv. Enteropathogenic *E. coli* (EPEC), is one of the most important pathogens affecting children worldwide with the infection resulting in persistent diarrhea (Ochoa and Contreras, 2011). The T2SS is required for EPEC virulence (Baldi et al., 2012);
- v. *Pseudomonas aeruginosa*, an opportunistic pathogen of major importance in cystic fibrosis patients, contains two distinct T2SS machineries (Ball et al., 2002; Jyot et al., 2011) and in certain strains even three (Cadoret et al., 2014);
- vi. The intracellular pathogen *Legionella pneumophila*, the causative agent of Legionnaire's disease, is dependent on the T2SS (DebRoy et al., 2006; Rossier et al., 2008).

The T2SS is made up from ~11-15 proteins, most of these present in multiple copies in the assembled secretion complex. As a result of numerous biochemical and structural studies, and from analogies to related systems, a generally accepted picture has emerged with the T2SS thought to consist of three subassemblies: the Inner Membrane Platform, the dynamic

Pseudopilus, and the channel-forming Outer Membrane Complex (Cianciotto, 2013; Douzi et al., 2012; Filloux, 2004; Howard, 2013; Johnson et al., 2006; Korotkov et al., 2012; McLaughlin et al., 2012; Nivaskumar and Francetic, 2014). The Inner Membrane Platform (Py et al., 2001) is composed of the T2SS membrane proteins GspC, GspF, GspL, GspM and, in some species, GspN. The ATPase GspE resides in the cytoplasm interacting with the cytoplasmic domain of GspL (Abendroth et al., 2005; Sandkvist et al., 1995; Shiue et al., 2006) and with GspF (Arts et al., 2007; Py et al., 2001). The stoichiometry of the Inner Membrane Complex, the nexus of the T2SS since it interacts with all other subassemblies, is still a mystery. The Pseudopilus contains five different pseudopilins: GspK, GspI, GspJ, GspH and GspG. The tip is formed by a GspK•GspI•GspJ heterotrimer (Korotkov and Hol, 2008), most likely linked by one or a few GspH subunits (Douzi et al., 2011; Yanez et al., 2008a) to a helical filament made up of multiple copies of a calcium-requiring GspG (Campos et al., 2011; Kohler et al., 2004; Korotkov et al., 2009; Yanez et al., 2008a; Yanez et al., 2008b). The Outer Membrane Complex is composed of a dodecamer of GspD subunits which form a gated channel of ~880 kDa (Chami et al., 2005; Reichow et al., 2010). Intriguingly, the T2SS is possibly only fully assembled transiently, perhaps triggered by the presence of exoproteins in the periplasm (Chen and Hu, 2013; Howard, 2013).

Over the decades, an increasing number of bacterial multi-protein machineries spanning the inner and outer membrane of Gram-negative bacteria have been uncovered. The system closest related to the T2SS is the type IV pilus system (T4PS) (Ayers et al., 2010). At least two types of T4PS exist, with the best studied the Type 4a Pilus system (T4aPS) that differs from the Type 4b Pilus system (T4bPS) in several ways, including a different major pilin subunit (Craig and Li, 2008) and a different protein and domain organization of the homolog of the T2SS inner membrane protein GspL (Supplementary Fig. S1B). T4PS variants perform a diversity of functions in a wide range of species (Craig and Li, 2008; Giltner et al., 2012; Pelicic, 2008). More distantly related systems are the bacterial transformation system and the archaeum assembly system (AAS) (Korotkov et al., 2011; Korotkov et al., 2012; Lassak et al., 2012). The critical functions of these systems in bacterial survival and pathogenicity increase the importance of our understanding the T2SS.

Among the many protein components of the T2SS, the secretion ATPase GspE plays an essential role, and is likely responsible for providing energy for the protein translocation process (Camberg and Sandkvist, 2005; Camberg et al., 2007; Patrick et al., 2011; Sandkvist et al., 1995). GspE has several other, species-specific, names and is e.g. in *Vibrio* species called EpsE (Sandkvist et al., 1995). Here we will use the generic nomenclature, i.e. GspE. GspE is a protein of ~500 residues folding into three major domains, the N-terminal domains N1E and N2E, and the C-terminal domain CTE. The CTE can be subdivided into the subdomains C1E, CME and C2E (Supplementary Fig. S1A) (Lu et al., 2013; Robien et al., 2003), where the CME is the critical zinc-binding domain (Camberg and Sandkvist, 2005; Possot and Pugsley, 1997). In *Xanthomonas campestris*, an additional domain (N0E) occurs prior to N1E (Chen et al., 2005). However, this is an exceptional case in the T2SS GspE family. The amino acid sequences of *V. vulnificus* and *V. cholerae* GspE share 48, 94 and 90% identity for the N1E, N2E and CTE, respectively (a sequence alignment is provided in Supplementary Fig. S2).

Here we report the first crystal structure containing a full-length T2SS secretion ATPase, while previous structures of the T2SS GspE missed the N1E. The initial *V. cholerae* N1EGspE structure contained an arrangement of molecules with  $6_1$  helical symmetry (Robien et al., 2003). Solution studies have provided evidence that GspE tends to form multimers, most likely hexamers (Camberg et al., 2007; Shiue et al., 2006). Recently, crystal structures of two different hexamers of *V. cholerae* N1EGspE have been obtained by using an “assistant hexamer”, Hcp1 (Lu et al., 2013), which served to induce multimer formation of the fused N1EGspE chains. One of these *V. cholerae* N1EGspE hexamers adopts an arrangement with quite regular, quasi  $C_6$ , symmetry, another hexamer is elongated exhibiting  $C_2$  symmetry (Lu et al., 2013). These hexamers reveal considerable variability in the orientation of the N2E versus the CTE. In contrast, the association of a CTE and a N2E' (i.e. N2E from a neighboring subunit) is remarkably similar in both hexamers of N1EGspE-Hcp1 fusion proteins as well as in the helical *V. cholerae* N1EGspE structure. This CTE•N2E' “construction unit” has also been observed in ATPase hexamers from related systems such as in the retraction ATPase PilT from the *Pseudomonas aeruginosa* and *Aquifex aeolicus* T4aPS (Misic et al., 2010; Satyshur et al., 2007), and in the ATPases from the AAS, *Archaeoglobus fulgidus* GspE2 and *Sulfolobus acidocaldarius* FlaI (Reindl et al., 2013; Yamagata and Tainer, 2007). (Note: the symbol “•” is used throughout to indicate non-covalent complexes). These latter ATPases lack the CME, and contain either no N1E at all, or an N1E with a different fold from the T2SS N1Es (Supplementary Fig. S1A). Hence the T2SS, T4PS and AAS ATPases share two common core domains, the N2E and CTE. These domains display major variations in length and number of subdomains, and are often distantly related in sequence. While T4PS and AAS ATPases form hexamers readily, the T2SS ATPase has so far been captured only as a stable hexamer when fused to Hcp1 as assistant hexamer (Lu et al., 2013).

Another important T2SS protein is GspL, which in *Vibrio* species is also called EpsL (Sandkvist et al., 1995), and has additional, species-specific, names (Supplementary Fig. S3). We use here the generic name GspL. GspL is a bitopic inner membrane protein that plays a central role in T2SS function since it interacts with several other T2SS proteins, including (1) the inner membrane platform protein GspM (Sandkvist et al., 1999); (2) the major pseudopilin GspG (Gray et al., 2011); and (3) GspE (Abendroth et al., 2005; Sandkvist et al., 1995). The cytoplasmic domain of GspL (cyto-GspL) is responsible for the interactions with the first domain of GspE (Sandkvist et al., 2000) and consists of three subdomains with similarities to proteins belonging to the actin-like ATPase superfamily (Abendroth et al., 2004a).

In the related T4aPS, the “homolog” of GspL is actually two separate proteins, PilM and PilN, corresponding to the cytoplasmic and periplasmic domains of GspL, respectively. PilM and PilN from *Thermus thermophilus* interact with each other in the cytoplasm via the N-terminus of PilN. PilM does not hydrolyze but binds ATP using a subdomain which is absent in *V. cholerae* cyto-GspL (Abendroth et al., 2004a; Abendroth et al., 2005; Karuppiah and Derrick, 2011). In the T4bPS responsible for the biogenesis of the bundle-forming pilus of EPEC, the homolog of GspL is a single protein, BfpC. The N-terminal domain of BfpC has only two of the three sub-domains in cyto-GspL (Yamagata et al.,

2012). Hence, there are distinct differences between the homologs of GspL in the T2SS, T4aPS and T4bPS, in particular regarding the cytoplasmic subdomains (Supplementary Fig. S1B).

Two structures of *V cholerae* cyto-GspL have been reported previously. The crystals containing *V cholerae* cyto-GspL (Abendroth et al., 2004a) revealed a three-subdomain architecture and also interactions between neighboring cyto-GspL subunits. The crystals containing *V cholerae* cyto-GspL in complex with the N1E of *V cholerae* GspE (Abendroth et al., 2005) showed the N1E•cyto-GspL heterodimer, revealing the interactions between the ATPase and inner membrane protein. Interestingly, in the crystals of this heterodimer, interactions were observed between neighboring cyto-GspL subunits, which are similar to interactions seen in the crystals of cyto-GspL by itself (Abendroth et al., 2004a).

The reluctance of T2SS ATPases by themselves to form hexamers in solution (Camberg and Sandkvist, 2005; Lu et al., 2013; Robien et al., 2003; Shiue et al., 2006), combined with evidence that cyto-GspL stimulates the ATPase activity of GspE (Camberg et al., 2007) and indirect evidence for hexamer formation in solution (Patrick et al., 2011; Shiue et al., 2006), encouraged us to undertake a study of a full-length T2SS GspE in complex with the cytoplasmic domain of GspL, to investigate whether or not the presence of cyto-GspL would induce hexamer formation of T2SS ATPases. We report here the 2.83 Å resolution crystal structure of the *V vulnificus* GspE•cyto-GspL complex with bound AMPPNP. The three copies of this complex per asymmetric unit are very similar, but no hexameric arrangement of GspE is present in the crystals. Entirely unexpected was the close interaction of cyto-GspL with the CTE of GspE and the bound AMPPNP. A novel N2E-*vs*-CTE orientation, which differs considerably from that in previously solved structures, was observed, expanding the range by which the N2E of T2SS ATPases can move and rotate with respect to the CTE. The current crystals of the *V vulnificus* GspE•cyto-GspL complex contain essentially the same two types of interfaces between cyto-GspL domains as present in the crystals of *V cholerae* cyto-GspL and of the *V cholerae* N1E•cyto-GspL heterodimer. In one of the interfaces a highly conserved hydrophobic cluster of residues is responsible for cyto-GspL•cyto-GspL contacts. The possible significance of these contacts for the T2SS is discussed.

## 2. Material and methods

### 2.1 Protein expression and purification

The DNA fragments corresponding to GspE and cyto-GspL were PCR amplified from genomic DNA of *V. vulnificus* YJ016 and cloned into a modified pET-21d vector (EMD Millipore) to create a bicistronic operon. The two *V. vulnificus* proteins, comprising residues 1-499 of GspE and 1-237 of GspL, with the latter carrying a C-terminal His<sub>6</sub>-tag, were expressed in BL21(DE3) *Escherichia coli* cells and purified via Ni<sup>2+</sup> immobilized affinity chromatography using standard methods. The final purification included size-exclusion chromatography using Superdex 200 column (GE Healthcare) in a 20 mM HEPES, pH7.5, 200 mM NaCl buffer.

## 2.2 Crystallization and data collection

The *V. vulnificus* GspE•cyto-GspL complex was crystallized in the presence of 1 mM AMPPNP and 2 mM MgCl<sub>2</sub>. Crystals were grown using the sitting drop vapor diffusion method by mixing 1 μL of protein solution and 1 μL of reservoir solution at 294 K. The reservoir solution was 0.2 M Na malonate pH 7.0, 18% PEG 3350. Crystals were cryoprotected using reservoir solution supplemented with 20% ethylene glycol. Data were collected on SSRL beam line BL12-2. Data were integrated and scaled using XDS and XSCALE (Kabsch, 2010). Data collection statistics are provided in Table 1.

## 2.3 Crystal structure determination

The 2.83 Å resolution structure was solved by molecular replacement with the program Phaser (McCoy et al., 2007) using the structures of *V. cholerae* N1EGspE (PDB 1P9R) and the *V. cholerae* N1E•cyto-GspL complex (PDB 2BH1) as search models (Abendroth et al., 2005; Robien et al., 2003). The crystals of the *V. vulnificus* GspE•cyto-GspL complex belong to space group C2 and contain three copies of the ~83 kDa complex per asymmetric unit. The model was iteratively built using Coot (Emsley and Cowtan, 2004) and refined with noncrystallographic symmetry restraints using REFMAC (Murshudov et al., 1997), resulting eventually in a R<sub>work</sub> of 24.8% and a R<sub>free</sub> of 28.0% with good geometry (Table 1). The electron density for almost all residues, including the 14-residue linker between the N2E and CTE in each of the three *V. vulnificus* GspE subunits, is well defined. Electron density is absent, however, for the linker residues (residues 79-95) between the N1E and N2E. Density is also missing for the first eleven and last two N- and C-terminal residues of GspE, and for the last two residues of cyto-GspL.

## 2.4 Structure analysis and figure preparation

Structures were superimposed and compared to other structures using SSM (Krissinel and Henrick, 2004) in Coot (Emsley and Cowtan, 2004). Interfaces between domains were analyzed with PISA (Krissinel and Henrick, 2007). Figures were prepared using Pymol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.) and ESPript (Gouet et al., 2003).

## 3. Results

### 3.1 Three *V. vulnificus* GspE•cyto-GspL complexes

The crystals of GspE•cyto-GspL contain three copies of full-length *V. vulnificus* GspE and three copies of cyto-GspL domains per asymmetric unit with a total molecular weight of ~250 kDa. The current 2.83 Å resolution structure is, to the best of our knowledge, the first structure containing a full-length ATPase from a T2SS. The overall B factors of the three complexes in the asymmetric unit are similar and range from 58 to 61 Å<sup>2</sup>. The average B-factor of the three cyto-GspL domains is ~53 Å<sup>2</sup>, of the N1E domains ~78 Å<sup>2</sup>, of the N2E domains ~63 Å<sup>2</sup>, and of the CTE ~59 Å<sup>2</sup>. The distances between the last visible residue of N1E and the first visible residue of the nearest three N2E domains in the crystal varies from 33 to 41 Å (Supplementary Fig. S4). This means that, given the ~17 residue-linker length, it is not possible to establish which N1E is actually connected to which N2E in the crystals. Since the three complexes are very similar, we describe below complex 1 (Fig. 1).

Key results are: (i) a new N2E-vs-CTE orientation angle, considerably larger than seen before; (ii) the cyto-GspL domain unexpectedly interacting with the CTE and with AMPPNP; and (iii) the presence of a linear arrangement of cyto-GspL domains observed now in three entirely different cyto-GspL-containing crystals. In these linear arrangements, two types of interfaces alternate. In one of these interfaces a conserved hydrophobic cluster of residues contacting each other across a twofold axis. The consistent occurrence of these linear arrangements in different crystal forms, leads to the suggestion of a pre-assembly complex of inner membrane and GspE proteins.

### 3.2 The *V. vulnificus* GspE subunit

Superposition of *V. vulnificus* N2E in the current structure onto N2E from the *V. cholerae* <sup>N1E</sup>GspE structure (Robien et al., 2003) yields an r.m.s.d. value of 0.49 Å, and the superposition of *V. vulnificus* CTE and *V. cholerae* CTE in these two structures an r.m.s.d. value of 0.61 Å.

These low r.m.s.d values reflect the 94 and 90 % amino acid sequence identities for these two domains from these two species. While the individual domains in these two structures superimpose very well, the orientation of the N2E with respect to the CTE differs, however, by not less than 171 degrees (Fig. 2). The two different orientations are observed in spite of the fact that in each instance the CTE binds Mg•AMPPNP. This difference in relative N2E-vs-CTE orientation in two structures is accompanied by a completely different conformation of the 14 residues in the linker connecting the N2E and CTE. The N2E-vs-CTE orientation seen in the current structure is also different from the four unique N2E-vs-CTE orientations observed in the recently published *V. cholerae* <sup>N1E</sup>GspE hexamers (Lu et al., 2013). Clearly, the flexible linker between the domains allows the N2E a large degree of freedom to move with respect to the CTE.

The most prominent interface in the current crystals occurs between the CTE of one *V. vulnificus* GspE chain and the N2E' of an adjacent GspE chain, with a buried solvent accessible surface area of ~1913 Å<sup>2</sup>. This interface is similar in size to that observed in helical and hexameric arrangements for *V. cholerae* <sup>N1E</sup>GspE (Lu et al., 2013; Robien et al., 2003) (PDB identifiers 1P9W, 4KSR, 4KSS), and have been called construction units of hexameric secretion ATPases (Lu et al., 2013). Structural comparisons of the *V. vulnificus* CTE•N2E' construction unit with that from the helical *V. cholerae* <sup>N1E</sup>GspE arrangement, and the four constructions units from the recently reported hexamers in *V. cholerae* <sup>N1E</sup>GspE-Hcp1 fusions, yield r.m.s.d. values between 0.50 and 0.73 Å for ~335 residues, showing that these construction units are essentially the same despite entirely different crystal environments.

This CTE•N2E' interface has also been consistently observed in the hexameric structures of the T2SS GspE homologs *Pseudomonas aeruginosa* PilT and *Aquifex aeolicus* PilT from the T4PS, and *S. acidocaldarius* FlaI and *A. fulgidus* GspE2 from the AAS (Misic et al., 2010; Reindl et al., 2013; Satyshur et al., 2007; Yamagata and Tainer, 2007). The mutual orientation of the N2E' and CTE in seven structures of these distantly related ATPases from a variety of secretion and assembly systems, with mutual amino acid sequence identities in the N2E domains as low as 14 %, differs by as little as 11 degrees.

### 3.3 The structure of the *V. vulnificus* cyto-GspL domain

The current *V. vulnificus* cyto-GspL structure superimposes onto that of *V. cholerae* cyto-GspL (Abendroth et al., 2004a) with an r.m.s.d. of 1.4 Å, and onto cyto-GspL from the *V. cholerae* N1E•cyto-GspL complex (Abendroth et al., 2005) with an r.m.s.d. of 1.2 Å with amino acid sequence identities of 61, 70 and 31 % for subdomains I, II and III, respectively. Clearly, the three-subdomain structure of cyto-GspL is essentially the same in these three crystal structures. While 1 mM AMPPNP was present in the crystallization solution, we observed no density representing this nucleotide analog bound to *V. vulnificus* cyto-GspL at a position equivalent to that of the distant homolog PilM of the *T. thermophilus* T4aPS, in agreement with the fact that subdomain 2B of PilM, involved in nucleotide binding by PilM, is absent in GspL proteins from the T2SS (Abendroth et al., 2004a; Abendroth et al., 2005; Karuppiah and Derrick, 2011) (Supplementary Fig. S1B).

### 3.4 The interaction between *V. vulnificus* N1E and cyto-GspL

The second largest interface in the current crystals occurs between *V. vulnificus* N1E and cyto-GspL domains, and buries ~1823 Å<sup>2</sup>. Each N1E in the crystal makes only contacts with a cyto-GspL domain, mainly by placing its α2 helix in the cleft between domains II and III of cyto-GspL (Fig. 3). The interface in the *V. cholerae* N1E•cyto-GspL interaction (Abendroth et al., 2005) buries ~1758 Å<sup>2</sup>, a number close to that in the current *V. vulnificus* N1E•cyto-GspL structure. Superposition of these two *Vibrio* N1E•cyto-GspL complexes yields an r.m.s.d. of 1.4 Å while many of the interface residues are hydrophobic and highly conserved, as can be seen from the amino acid sequence alignment of a diverse set of cyto-GspL sequences (Supplementary Fig. S3).

### 3.5 AMPPNP binding by *V. vulnificus* CTE and cyto-GspL

Density for AMPPNP is clearly present in the active site of all three *V. vulnificus* CTE domains in the current structure (Supplementary Fig. S5). The Mg<sup>2+</sup> ion interacts with oxygens from the β- and γ-phosphates of AMPPNP. AMPPNP is bound to *V. vulnificus* CTE by the side chains of residues Leu235 and Arg437 which interact with the adenine (Fig. 4A). Residues from the highly conserved P-loop Gly265-Lys266-Ser267-Thr268 make extensive contacts with the phosphate moiety of the nucleotide. The oxygen atoms of the α-phosphate form hydrogen bonds with the backbone amide groups of Ser267 and Thr268, and the side-chain oxygen of Thr268. The oxygen atoms of the β-phosphate form hydrogen bonds with the backbone amide groups of Gly265, Lys266 and Ser267, and with the side chains of Ser264, Lys266 and Ser267. The side chain amino group of Lys266 forms hydrogen bonds with the oxygen atoms of the γ-phosphate.

The interactions of CTE residues with AMPPNP are similar in the *V. vulnificus* and *V. cholerae* secretion ATPases (Fig. 4A). Whereas residues 239-497 of the CTE domains in the *V. vulnificus* GspE and *V. cholerae* N1E•GspE structures (Robien et al., 2003) adopt almost the same conformation, the loop formed by residues 225-238 connecting the N2E and CTE adopts a very different conformation in the current structure compared to the helical *V. cholerae* N1E•GspE structure, as mentioned before. As a result of this conformational change, residues Thr228 and Leu230 of the *V. vulnificus* GspE structure are not in contact



with the ribose and adenine moieties, respectively, while the equivalent Thr232 and Leu234 in the *V. cholerae* GspE structure are (Fig. 4A).

The cyto-GspL domain is also in contact with AMPPNP, specifically: (1) the N<sup>ε</sup>2 of Gln13 of *V. vulnificus* cyto-GspL makes hydrogen bonds with an oxygen of the  $\gamma$ -phosphoryl group; (2) the O<sup>ε</sup>1 of Gln13 interacts with the NH of AMPPNP bridging the  $\beta$ - and  $\gamma$ -phosphorous atoms; (3) the N<sup>ε</sup>2 of His120 forms an hydrogen bond with an oxygen of the  $\gamma$ -phosphoryl group; and (4) the side chain of Tyr14 contacts the ribose through hydrophobic interactions (Fig. 4B). This unanticipated result will be discussed below.

### 3.6 The interactions between *V. vulnificus* CTE and cyto-GspL

A most unexpected finding in our current structure is that *V. vulnificus* cyto-GspL interacts with the CTE domain while burying  $\sim 1349 \text{ \AA}^2$  solvent accessible surface in the interface. Three residues from cyto-GspL involved in the contacts are engaged in salt bridges: Asp37, Arg116 and Asp123 are within 2.9-3.6  $\text{\AA}$  from the side chains of residues Arg494, Asp289 and Arg332 of the CTE, respectively. Additional contacts occur between side chains of Gln39 and Arg116 of cyto-GspL form hydrogen bonds with main chain atoms of CTE residues Val303 and Leu388 (Fig. 4C). Residues Ser36, Trp38, Leu41 and Ala124 of cyto-GspL and residues Pro261, Thr262, Gly263, Ser264 and Val 495 of CTE contribute van der Waals interactions to the interface.

The position of cyto-GspL with respect to the CTE is quite remarkable. When comparing the cyto-GspLCTE complex in the current structure with that of the N2E-CTE arrangement in the helical *V. cholerae* N1EGspE structure (Robien et al., 2003), it appears that in the current structure the cyto-GspL chain occupies approximately the position of the N2E in the helical structure. The N2E has swung away to allow cyto-GspL to approach the AMPPNP and the active site of the CTE (Fig. 2).

### 3.7 Interactions between cyto-GspL domains: similar “rods” in multiple crystals

The *V. vulnificus* GspE•cyto-GspL structure provides an interesting opportunity to look at mutual interactions between cyto-GspL domains. In the current crystals, there are three independent Cyto-GspL domains per asymmetric unit. Two of these domains, #1 and #2, interact with each other mainly involving residues from the  $\beta$ C strand of subdomain I. This contact is hereafter called the “ $\beta$ -interface” (Fig. 5A). These two domains are related by a pseudo-twofold axis. Each of these domains also interacts, at the opposite side, with a domain from a neighboring asymmetric unit engaging primarily residues from helix  $\alpha$ 2 and the loop connecting strand  $\beta$ E and helix  $\alpha$ 2 of subdomain II. This interface between cyto-GspL domains #1 and #2' is hereafter called the “ $\alpha$ -interface” (Fig. 5B). Also these pairs of domains are related by a pseudo-twofold axis. The third crystallographically independent cyto-GspL domain, #3, uses crystallographic twofold axes to form contacts with neighboring cyto-GspL domains with a  $\alpha$ -interface on one side and a  $\beta$ -interface on the other side of the domain, essentially the same arrangement as between domains #1 and #2', and between #1 and #2, respectively (Fig. 5A).

The  $\alpha$ -interface (Fig. 5B) buries  $\sim 1008 \text{ \AA}^2$  solvent accessible surface area in the #1•#2 contact and  $\sim 835 \text{ \AA}^2$  solvent accessible surface area in the #3•#3' contact. The #1•#2 and #3•#3' pairs of cyto-GspL domains are very similar and superimpose with an r.m.s.d. of  $\sim 0.6 \text{ \AA}$ . The  $\alpha$ -interface has a hydrophobic center with Pro70 and Tyr83 engaged in hydrophobic interactions with Pro70, Ser79, Met80, Tyr83 and Leu84 of the neighboring subunit. Due to the twofold axis relating the two domains in this interface, this set of contacts occurs twice between this pair of cyto-GspL domains. Hydrogen bonds and electrostatic interactions involving Arg75, Gln76 and Asp87, occur on both sides of the hydrophobic center of the  $\alpha$ -interface.

The  $\beta$ -interface (Fig. 5A) buries  $\sim 530 \text{ \AA}^2$  solvent accessible surface area in the #1•#2 contact and  $\sim 529 \text{ \AA}^2$  in the #3•#3'' contact, where #33'' is related to #3 by a crystallographic twofold. The area buried in the  $\beta$ -interface is therefore almost half of that buried in the  $\alpha$ -interface. The #1•#2 and #3•#3'' pairs of cyto-GspL domains superimpose with an r.m.s.d. of  $\sim 0.6 \text{ \AA}$ . The domains forming the  $\beta$ -interface are related by a pseudo-twofold axis. The major contacts in this interface are the hydrophobic interactions between residues Val29 and Glu34 and the two main-chain hydrogen bonds from Ser32, involving two anti-parallel  $\beta$ C strands, one from each domain.

Both the  $\alpha$ -interface and  $\beta$ -interface, seen twice in the current crystal of *V. vulnificus* GspE•cyto-GspL, occur in two previous crystal structures. One of these crystals contains *V. cholerae* cyto-GspL (Abendroth et al., 2004a) and the other the *V. cholerae* N1E•cyto-GspL heterodimer (Abendroth et al., 2005). Hence, we have four crystallographically independent views of both the  $\alpha$ -interface and the  $\beta$ -interface, in crystals grown under different conditions, in different crystal forms, and with very different protein content (Table 2). The interdomain angles between the various pairs of *Vibrio* cyto-GspL domains forming  $\alpha$ - and  $\beta$ -interfaces in these crystals vary only by up to  $\sim 19$  degrees (Table 2). In all four instances of the  $\alpha$ -interface, a completely conserved Pro70 and two highly conserved hydrophobic residues Tyr83 and Leu84 (Supplementary Fig. S3) are forming the hydrophobic center of this interface. In addition, in three of the four structures (two *V. vulnificus* GspE•cyto-GspL and *V. cholerae* N1E•cyto-GspL) Arg75, Gln76 and Asp87 engage in hydrogen bonds or electrostatic interactions with twofold related side chains.

The similarity of the interactions between cyto-GspL domain contacts in the three crystals extends far beyond that of nearest neighbors. In the three cyto-GspL-containing crystals with known structure now available, it appears that the same arrangement of adjacent subunits of cyto-GspL domains with alternating  $\alpha$ - and  $\beta$ -interfaces form “cyto-GspL rods”, spanning the entire crystal. The angular deviations seen in a pairwise comparison of  $\alpha$  and  $\beta$  interfaces (Table 2) are such that essentially the same straight rods occur four times in these crystal forms (Fig. 5C). Possible implications of this tendency to form such linear arrangements are discussed below.

## 4. Discussion

### 4.1 The *V. vulnificus* GspE•cyto-GspL complex

The 2.83 Å resolution crystal structure of the complex of *V. vulnificus* GspE and cyto-GspL, reveals many features of and contacts between domains and subdomains of these two key proteins of the T2SS. In particular, the cyto-GspL domain is involved in several interactions. This leads to a hypothesis of a pre-assembly complex of this Inner Membrane Platform as described in the final section below.

### 4.2 The *V. vulnificus* GspE subunit

The GspE subunit appears to be remarkably flexible, even when in complex with its partner, the cytoplasmic domain of GspL. Not only is the linker between the N1E and N2E invisible in the crystals, the current N2E-*vs*-CTE orientation is very different from the orientations in helical *V. cholerae* N1E GspE (Robien et al., 2003) and in the two recently determined *V. cholerae* GspE hexamers (Lu et al., 2013). In the current structure, the N2E and CTE domains of the GspE subunit are oriented with a novel N2E-*vs*-CTE orientation without contacts of the N2E with AMPPNP (Fig. 1). The variability of conformations that the full-length GspE adopts during the biogenesis and action of the T2SS remains to be determined, but the currently available structures of the T2SS ATPase indicate that this key T2SS protein can adopt a tremendous range of N2E-*vs*-CTE orientations.

In contrast to the large variation in the N2E-*vs*-CTE orientations observed within the same subunit, the CTE•N2E' construction unit, bringing two domains from different GspE subunits in contact, is very similar to that observed in the helical (Robien et al., 2003) and hexameric *V. cholerae* N1E GspE structures, with r.m.s.d. values of 0.7 and 0.5 Å, respectively (Lu et al., 2013). As previously discussed (Lu et al., 2013; Reindl et al., 2013), this same construction unit also occurs in secretion ATPases from the T4aPS (Mistic et al., 2010; Satyshur et al., 2007) and in the more distantly related homologs from the AAS (Reindl et al., 2013; Yamagata and Tainer, 2007). This provides further support for a general significance of the CTE•N2E' interactions in a wide variety of secretion and related systems, probably an early invention during the evolution of these systems.

Studies on purified GspE in solution indicates the presence of transient multimers, probably hexamers, as judged by e. g. ATPase activity in the size-exclusion chromatography fractions corresponding to hexamers (Camberg and Sandkvist, 2005). It is apparently not easy to capture hexamers of GspE in structural studies, even in the presence of cyto-GspL as evident from the results in the current paper. So far, only fusion to the “assistant hexamer” Hcp1 has resulted in GspE hexamers (Lu et al., 2013). In contrast, several GspE homologs do not display a similar reluctance to form hexamers (Mistic et al., 2010; Reindl et al., 2013; Satyshur et al., 2007; Yamagata and Tainer, 2007). Also, the assembly ATPase PilF from the *Thermus thermophilus* T4aPS forms, in the presence and absence of AMPPNP, hexamers as observed by negative-stained electron microscopy (Collins et al., 2013). However, this T4aPS assembly ATPase has two additional N-terminal domains compared to T2SS assembly ATPases (Supplementary Fig. S1A), and these domains could possibly promote hexamer formation. The reason that the T2SS assembly ATPase is so reluctant to

form hexamers remains a puzzle. Possibly contacts with additional inner membrane T2SS proteins, including GspF (Arts et al., 2007; Py et al., 2001), are required to obtain the putative hexameric arrangement of full-length GspE.

#### 4.3 The *V. vulnificus* CTE•cyto-GspL contacts and AMPPNP binding

The quite extensive interface and direct contact of cyto-GspL with the CTE and with AMPPNP is an unexpected result of the current structure. Therefore, we looked into the conservation of the residues engaged in the CTE•cyto-GspL interface. Out of 14 cyto-GspL and 16 CTE interface residues, 2 and 11 residues are conserved, respectively (Supplementary Fig. S2 and S3). Seven of the conserved CTE residues are located in the highly conserved Walker A and Walker B boxes, which are functionally important for the ATP binding and catalysis in the secretion ATPases (Possot and Pugsley, 1994; Rivas et al., 1997; Sandkvist et al., 1995). None of the residues of cyto-GspL engaged in the three Arg-Asp salt bridges across the *V. vulnificus* CTE•cyto-GspL interface (Fig. 4C) are conserved in any other species, neither are the two *V. vulnificus* cyto-GspL residues, Gln13 and His120, that contact oxygens of the bound AMPPNP. In fact, Gln13 and His120 in other species are in several instances an Ala, a residue that is unable to make the favorable interactions with, respectively, the oxygen linking the  $\beta$ - and  $\gamma$ -phosphates in ATP, or with a terminal  $\gamma$  phosphate oxygen. In view of this lack of conservation, it is rather unlikely that these cyto-GspL•CTE interactions and the cyto-GspL•AMPPNP interactions are of physiological importance.

#### 4.4 The interaction between *V. vulnificus* N1E and cyto-GspL

The interactions between N1E and cyto-GspL are very similar in the T2SS from two *Vibrio* species, *V. vulnificus* and *V. cholerae*, with the mutual orientation of the two domains differing by less than 11 degrees (Fig. 3) (Abendroth et al., 2004a; Abendroth et al., 2005). Several N1E residues in the N1E•cyto-GspL interface are highly conserved. In fact, four (Leu14, Pro15, Glu50 and Arg53) out of the six completely conserved residues of N1E are located in this interface (Supplementary Fig. S2). The N1E•cyto-GspL interaction is therefore most likely of major importance in the T2SS of many species in order to link the ATPase GspE to the Inner Membrane Platform. It is interesting that in *X. campestris* an additional N-terminal N0E domain occurs (called XpsL) (Chen et al., 2005), but we do not consider this N0E domain further since it is absent in the species of our current investigation and also occurs rarely in the T2SS GspE family.

#### 4.5 Linear arrangements of cyto-GspL domains – possible implications for T2SS assembly

Most intriguingly, crystals of *V. cholerae* cyto-GspL, *V. cholerae* N1E•cyto-GspL and the current *V. vulnificus* GspE•cyto-GspL all contain essentially the same linear arrangement of cyto-GspL domains consisting of straight rods with alternating  $\alpha$ - and  $\beta$ -interfaces between adjacent cyto-GspL domains (Table 2; Fig. 5C and D). While the  $\beta$ -interface is not very extensive, the  $\alpha$ -interface buries almost twice as much surface area and has a hydrophobic center. Moreover, this hydrophobic interface center contains a completely conserved Pro70 residue, which is in contact with a highly conserved Tyr and Leu of helix  $\alpha 2$  from an adjacent chain (Supplementary Fig. S3). These hydrophobic interactions occur around a

twofold axis in the center of the  $\alpha$ -interface (Fig. 6A). Hydrophilic and electrostatic interactions between Arg75, Gln76 and Asp87, which are conserved hydrophilic residues across GspL family members (Supplementary Fig. S3), surround the hydrophobic interactions in this interface.

This linear arrangement of cyto-GspL domains has now been observed four times in three different crystals in the absence and presence of other protein chains (Table 2). This suggests that it is a favorable arrangement of cyto-GspL domains, which may also occur *in vivo*. This linear array might play a role during the biogenesis of the T2SS along the following lines:

- In the cell, rods of cyto-GspL domains resembling those observed in the crystals (Fig. 6C), might be a transient arrangement in order to obtain a high local concentration of GspL subunits for rapid subsequent assembly of the Inner Membrane Platform. In this regard, it is of interest that periplasmic domains of *V. parahaemolyticus* GspL (peri-GspL) form dimers in solution as well as in crystals (Abendroth et al., 2009). In the crystal lattice, this dimer interface buries 1487 Å<sup>2</sup> solvent accessible surface and involves a number of conserved residues. For convenience, we will call this the “p-interface” of GspL. Purified full-length GspL forms dimers (Sandkvist et al., 1999). It is conceivable that full-length GspL dimerizes with simultaneous inter-subunit interactions between the periplasmic and cytoplasmic domains. Full-length GspL dimers with an  $\alpha$ -interface in the cytoplasm and the p-interface in the periplasm might be quite stable due to a total buried surface area of  $\sim 2400$  Å<sup>2</sup> (Fig. 6A). Multiple GspL dimers with  $\alpha$ - and p-interfaces on opposite sides of the inner membrane can then form linear arrangements using the  $\beta$ -interface as points of contact between cytoplasmic domains (Fig. 6B).
- Since one cyto-GspL domain interacts with one N1E domain (Abendroth et al., 2005; Karupiah and Derrick, 2011; Shiue et al., 2006) (Fig. 3), it is likely that (GspE•GspL)<sub>2</sub> hetero-tetramers form with each N1E domain tightly bound to a cyto-GspL domain. In this way, multiple GspE domains form a linear pre-assembly complex of (GspE•GspL)<sub>2</sub> hetero-tetramers (Fig. 6C). Given the flexibility of the N1E-N2E linker, as evidenced by the lack of density for these linker residues in our current crystals, the N2E and CTE domains are at this stage highly flexible with respect to the N1E.
- Specific interaction of Inner Membrane Platform proteins with either the Outer Membrane Complex (Howard, 2013), or with exoproteins to be secreted (Chen and Hu, 2013), might be a trigger in assembling a T2SS (Fig. 6D). For rapid assembly of the T2SS Inner Membrane Platform, rods consisting of (GspE•GspL)<sub>2</sub> hetero-tetramers in the inner membrane might provide a ready pool of building blocks for assembly.
- The C<sub>2</sub> symmetry of (GspL•N1E)<sub>2</sub> hetero-tetramers needs to be reconciled with the approximate C<sub>6</sub> symmetry of hexamers formed by the N2E and CTE domains. A key role might be played by the N1E-N2E linker that consists of  $\sim 19$  residues. The

“top” part of the (GspL•GspE)<sub>6</sub> multimer comprising GspL•N1E domains may adopt approximate C<sub>3</sub> symmetry, relating three (GspL•N1E)<sub>2</sub> heterodimers with each C<sub>2</sub> symmetry. Six flexible N1E-N2E linkers would connect this upper part to six “lower” N2E-CTE domains with approximate cyclic C<sub>6</sub> symmetry. This generates a GspL<sub>6</sub>•GspE<sub>6</sub> assembly with approximate overall C<sub>3</sub> symmetry. This point has been discussed in the paper reporting the *V. parahaemolyticus* peri-GspL structure (Abendroth et al., 2009). (An example of flexible linkers connecting “upper” and “lower” parts of an assembly which do not obey the same symmetry axes has been reported for the glutamate receptor tetramer (Sobolevsky et al., 2009).) Here we propose that the α-interface is the predominant GspL interface on the cytoplasmic side of the (GspL•GspE)<sub>6</sub> assembly (Fig. 6).

- Other inner membrane proteins, like GspM, are known to interact with GspL (Sandkvist et al., 1999; Sandkvist et al., 2000) and may be incorporated into a linear pre-assembly complex comprising multiple (GspM•GspL•GspE)<sub>6</sub> assemblies.
- Given the fact that the hexamer composed of N2E-CTE domains can adopt different arrangements (Lu et al., 2013), alterations in the mutual positions of these domains during ATP hydrolysis by GspE likely conveys motions of the N2E-CTE domains via the N1E-N2E linker to the N1E•cyto-GspL heterodimer. Changes in positions of cyto-GspL domains will alter positions of periplasmic GspL domains, and of other inner membrane platform proteins, thereby promoting assembly of the pseudopilus in the periplasm.
- The arrangement of full-length-GspL dimers as sketched above might only exist transiently in the fully assembled T2SS since it has recently been proposed, on the basis of disulfide cross-linking studies, that the periplasmic domains of full-length peri-GspL are part of the time forming homodimers, and part of the time peri-GspM•peri-GspL heterodimers (Lallemand et al., 2013). Whether or not the α-interface of the GspL dimer would be maintained in the cytoplasm during such proposed rearrangements of GspL domains in the periplasm remains to be determined.

Recent electron microscopy studies on a *Thermus thermophilus* T4aPS sub-assembly have proposed the presence of hetero-multimers formed by two copies each of the inner membrane proteins PilM, PilN and PilO, which are the homologs of the T2SS GspL and GspM Inner Membrane Platform proteins (Karupiah et al., 2013; Sampaleanu et al., 2009). This is in line with the dimers of the cytoplasmic and periplasmic domains of GspL and of GspM observed in crystal structures of these T2SS proteins (Abendroth et al., 2009; Abendroth et al., 2004a; Abendroth et al., 2004b; Abendroth et al., 2005). However, there are also important differences between inner membrane proteins of the T2SS and T4aPS. Particularly relevant for the current discussion is the fact that the conserved T2SS residues in the α-interface of cyto-GspL are not conserved in PilM (Supplementary Fig. S6). In view of such key differences, it seems best to refrain from extending the ideas presented here regarding the biogenesis, composition and symmetry of the T2SS Inner Membrane Complex to that of the T4aPS.

It is clear from the above that the intriguing but fascinating T2SS is revealing many of its secrets in a most reluctant manner. Further structural and biochemical studies are obviously required to confirm current hypotheses and to add new information.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We gratefully acknowledge contributions from Jan Abendroth to the early stages of these studies. We thank Stewart Turley and Jonathan Kay for support regarding equipment and computing, and the staff of BL12-2 beam line at the Stanford Synchrotron Radiation Lightsource (SSRL) for invaluable assistance with data collection. This study was funded by National Institute of Health grant R01 AI34501 to WGJH. Use of the Stanford Synchrotron Radiation Lightsource is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## References

- Abendroth J, Kreger AC, Hol WG. The dimer formed by the periplasmic domain of EpsL from the Type 2 Secretion System of *Vibrio parahaemolyticus*. *Journal of structural biology*. 2009; 168:313–322. [PubMed: 19646531]
- Abendroth J, Bagdasarian M, Sandkvist M, Hol WG. The structure of the cytoplasmic domain of EpsL, an inner membrane component of the type II secretion system of *Vibrio cholerae*: an unusual member of the actin-like ATPase superfamily. *J Mol Biol*. 2004a; 344:619–633. [PubMed: 15533433]
- Abendroth J, Rice AE, McLuskey K, Bagdasarian M, Hol WG. 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*. 2004b; 338:585–596. [PubMed: 15081815]
- Abendroth J, Murphy P, Sandkvist M, Bagdasarian M, Hol WG. The X-ray structure of the type II secretion system complex formed by the N-terminal domain of EpsE and the cytoplasmic domain of EpsL of *Vibrio cholerae*. *J Mol Biol*. 2005; 348:845–855. [PubMed: 15843017]
- Arts J, de Groot A, Ball G, Durand E, El Khattabi M, Filloux A, Tommassen J, Koster M. Interaction domains in the *Pseudomonas aeruginosa* type II secretory apparatus component XcpS (GspF). *Microbiology*. 2007; 153:1582–1592. [PubMed: 17464073]
- Ayers M, Howell PL, Burrows LL. Architecture of the type II secretion and type IV pilus machineries. *Future microbiology*. 2010; 5:1203–1218. [PubMed: 20722599]
- Baldi DL, Higginson EE, Hocking DM, Praszquier J, Cavaliere R, James CE, Bennett-Wood V, Azzopardi KI, Turnbull L, Lithgow T, Robins-Browne RM, Whitchurch CB, Tauschek M. The type II secretion system and its ubiquitous lipoprotein substrate, SsIE, are required for biofilm formation and virulence of enteropathogenic *Escherichia coli*. *Infect Immun*. 2012; 80:2042–2052. [PubMed: 22451516]
- Ball G, Durand E, Lazdunski A, Filloux A. A novel type II secretion system in *Pseudomonas aeruginosa*. *Mol Microbiol*. 2002; 43:475–485. [PubMed: 11985723]
- Cadoret F, Ball G, Douzi B, Voulhoux R. Txc, a New Type II Secretion System of *Pseudomonas aeruginosa* Strain PA7, Is Regulated by the TtsS/TtsR Two-Component System and Directs Specific Secretion of the CbpE Chitin-Binding Protein. *J Bacteriol*. 2014; 196:2376–2386. [PubMed: 24748613]
- Camberg JL, Sandkvist M. Molecular analysis of the *Vibrio cholerae* type II secretion ATPase EpsE. *J Bacteriol*. 2005; 187:249–256. [PubMed: 15601709]
- Camberg JL, Johnson TL, Patrick M, Abendroth J, Hol WG, Sandkvist M. Synergistic stimulation of EpsE ATP hydrolysis by EpsL and acidic phospholipids. *Embo J*. 2007; 26:19–27. [PubMed: 17159897]

- Campos M, Francetic O, Nilges M. Modeling pilus structures from sparse data. *Journal of Structural Biology*. 2011; 173:436–444. [PubMed: 21115127]
- Chami M, Guilvout I, Gregorini M, Remigy HW, Muller SA, Valerio M, Engel A, Pugsley AP, Bayan N. Structural insights into the secretin PulD and its trypsin-resistant core. *J Biol Chem*. 2005; 280:37732–37741. [PubMed: 16129681]
- Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. 2010; 66:12–21. [PubMed: 20057044]
- Chen YL, Hu NT. Function-Related Positioning of the Type II Secretion ATPase of *Xanthomonas campestris* pv. *campestris*. *PLoS ONE*. 2013; 8
- Chen Y, Shiue SJ, Huang CW, Chang JL, Chien YL, Hu NT, Chan NL. Structure and function of the XpsE N-terminal domain, an essential component of the *Xanthomonas campestris* type II secretion system. *J Biol Chem*. 2005; 280:42356–42363. [PubMed: 16162504]
- Cianciotto NP. Type II Secretion and Legionella Virulence. *Curr Top Microbiol Immunol*. 2013; 376:81–102. [PubMed: 23900831]
- Collins RF, Hassan D, Karuppiiah V, Thistlethwaite A, Derrick JP. Structure and mechanism of the PilF DNA transformation ATPase from *Thermus thermophilus*. *The Biochemical journal*. 2013; 450:417–425. [PubMed: 23252471]
- Craig L, Li J. Type IV pilli: paradoxes in form and function. *Current Opinion in Structural Biology*. 2008; 18:267–277. [PubMed: 18249533]
- DebRoy S, Dao J, Soderberg M, Rossier O, Cianciotto NP. Legionella pneumophila type II secretome reveals unique exoproteins and a chitinase that promotes bacterial persistence in the lung. *Proc Natl Acad Sci U S A*. 2006; 103:19146–19151. [PubMed: 17148602]
- Douzi B, Filloux A, Voulhoux R. On the path to uncover the bacterial type II secretion system. *Philosophical Transactions of the Royal Society B-Biological Sciences*. 2012; 367:1059–1072.
- Douzi B, Ball G, Cambillau C, Tegoni M, Voulhoux R. Deciphering the Xcp Pseudomonas aeruginosa type II secretion machinery through multiple interactions with substrates. *J Biol Chem*. 2011; 286:40792–40801. [PubMed: 21949187]
- Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Biological Crystallography*. 2004; 60:2126–2132.
- Filloux A. The underlying mechanisms of type II protein secretion. *Biochim Biophys Acta*. 2004; 1694:163–179. [PubMed: 15546665]
- Giltner CL, Nguyen Y, Burrows LL. Type IV pilin proteins: versatile molecular modules. *Microbiology and molecular biology reviews : MMBR*. 2012; 76:740–772. [PubMed: 23204365]
- Goldwater PN, Bettelheim KA. Treatment of enterohemorrhagic *Escherichia coli* (EHEC) infection and hemolytic uremic syndrome (HUS). *BMC medicine*. 2012; 10:12. [PubMed: 22300510]
- Gouet P, Robert X, Courcelle E. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res*. 2003; 31:3320–3323. [PubMed: 12824317]
- Gray MD, Bagdasarian M, Hol WG, Sandkvist M. In vivo cross-linking of EpsG to EpsL suggests a role for EpsL as an ATPase-pseudopilin coupling protein in the Type II secretion system of *Vibrio cholerae*. *Mol Microbiol*. 2011; 79:786–798. [PubMed: 21255118]
- Hirst TR, Holmgren J. Transient entry of enterotoxin subunits into the periplasm occurs during their secretion from *Vibrio cholerae*. *J Bacteriol*. 1987; 169:1037–1045. [PubMed: 3493239]
- Hirst TR, Sanchez J, Kaper JB, Hardy SJ, Holmgren J. Mechanism of toxin secretion by *Vibrio cholerae* investigated in strains harboring plasmids that encode heat-labile enterotoxins of *Escherichia coli*. *Proc Natl Acad Sci U S A*. 1984; 81:7752–7756. [PubMed: 6393126]
- Ho TD, Davis BM, Ritchie JM, Waldor MK. Type 2 secretion promotes enterohemorrhagic *Escherichia coli* adherence and intestinal colonization. *Infect Immun*. 2008; 76:1858–1865. [PubMed: 18316380]
- Howard SP. Assembly of the type II secretion system. *Res Microbiol*. 2013; 164:535–544. [PubMed: 23541475]
- Johnson TL, Abendroth J, Hol WG, Sandkvist M. Type II secretion: from structure to function. *FEMS Microbiol Lett*. 2006; 255:175–186. [PubMed: 16448494]



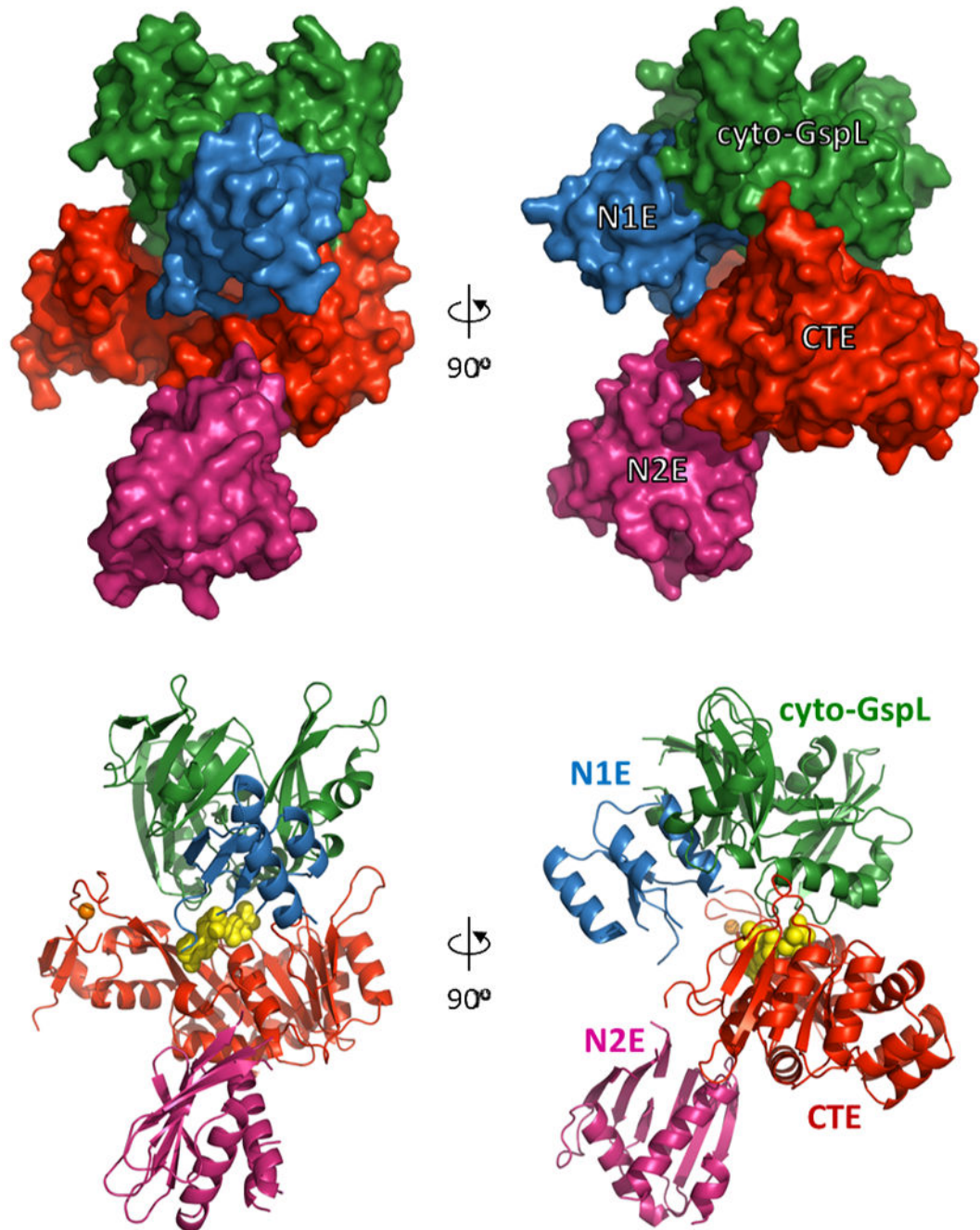
- Jyot J, Balloy V, Jouvion G, Verma A, Touqui L, Huerre M, Chignard M, Ramphal R. Type II secretion system of *Pseudomonas aeruginosa*: in vivo evidence of a significant role in death due to lung infection. *The Journal of infectious diseases*. 2011; 203:1369–1377. [PubMed: 21502078]
- Kabsch W. Xds. *Acta Crystallogr D Biol Crystallogr*. 2010; 66:125–132. [PubMed: 20124692]
- Karuppiah V, Derrick JP. Structure of the PilM-PilN inner membrane type IV pilus biogenesis complex from *Thermus thermophilus*. *J Biol Chem*. 2011; 286:24434–24442. [PubMed: 21596754]
- Karuppiah V, Collins RF, Thistlethwaite A, Gao Y, Derrick JP. Structure and assembly of an inner membrane platform for initiation of type IV pilus biogenesis. *Proc Natl Acad Sci U S A*. 2013; 110:E4638–4647. [PubMed: 24218553]
- Kohler R, Schafer K, Muller S, Vignon G, Diederichs K, Philippsen A, Ringler P, Pugsley AP, Engel A, Welte W. Structure and assembly of the pseudopilin PulG. *Mol Microbiol*. 2004; 54:647–664. [PubMed: 15491357]
- Korotkov KV, Hol WG. Structure of the GspK-GspI-GspJ complex from the enterotoxigenic *Escherichia coli* type 2 secretion system. *Nature structural & molecular biology*. 2008; 15:462–468.
- Korotkov KV, Gonen T, Hol WG. Secretins: dynamic channels for protein transport across membranes. *Trends in biochemical sciences*. 2011; 36:433–443. [PubMed: 21565514]
- Korotkov KV, Sandkvist M, Hol WG. The type II secretion system: biogenesis, molecular architecture and mechanism. *Nature reviews. Microbiology*. 2012; 10:336–351. [PubMed: 22466878]
- Korotkov KV, Pardon E, Steyaert J, Hol WG. Crystal structure of the N-terminal domain of the secretin GspD from ETEC determined with the assistance of a nanobody. *Structure*. 2009; 17:255–265. [PubMed: 19217396]
- Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallographica Section D Biological Crystallography*. 2004; 60:2256–2268.
- Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. 2007; 372:774–797. [PubMed: 17681537]
- Lallemant M, Login FH, Guschinskaya N, Pineau C, Effantin G, Robert X, Shevchik VE. Dynamic interplay between the periplasmic and transmembrane domains of GspL and GspM in the type II secretion system. *PLoS One*. 2013; 8:e79562. [PubMed: 24223969]
- Lassak K, Ghosh A, Albers SV. Diversity, assembly and regulation of archaeal type IV pili-like and non-type-IV pili-like surface structures. *Res Microbiol*. 2012; 163:630–644. [PubMed: 23146836]
- Lu C, Turley S, Marionni ST, Park YJ, Lee KK, Patrick M, Shah R, Sandkvist M, Bush MF, Hol WG. Hexamers of the type II secretion ATPase GspE from *Vibrio cholerae* with increased ATPase activity. *Structure*. 2013; 21:1707–1717. [PubMed: 23954505]
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007; 40:658–674. [PubMed: 19461840]
- McLaughlin LS, Haft RJF, Forest KT. Structural insights into the Type II secretion nanomachine. *Current Opinion in Structural Biology*. 2012; 22:208–216. [PubMed: 22425326]
- Merritt EA, Hol WG. AB5 toxins. *Current Opinion in Structural Biology*. 1995; 5:165–171. [PubMed: 7648317]
- Misic AM, Satyshur KA, Forest KT. *P. aeruginosa* PilT structures with and without nucleotide reveal a dynamic type IV pilus retraction motor. *J Mol Biol*. 2010; 400:1011–1021. [PubMed: 20595000]
- Murshudov GN, Vagin AA, Dodson EJ. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr D*. 1997; 53:240–255. [PubMed: 15299926]
- Nivaskumar M, Francetic O. Type II secretion system: A magic beanstalk or a protein escalator. *Biochim Biophys Acta*. 2014
- Ochoa TJ, Contreras CA. Enteropathogenic *Escherichia coli* infection in children. *Current opinion in infectious diseases*. 2011; 24:478–483. [PubMed: 21857511]
- Patrick M, Korotkov KV, Hol WG, Sandkvist M. Oligomerization of EpsE coordinates residues from multiple subunits to facilitate ATPase activity. *J Biol Chem*. 2011; 286:10378–10386. [PubMed: 21209100]

- Pellic V. Type IV pili: e pluribus unum? *Mol Microbiol.* 2008; 68:827–837. [PubMed: 18399938]
- Possot O, Pugsley AP. Molecular characterization of PulE, a protein required for pullulanase secretion. *Mol Microbiol.* 1994; 12:287–299. [PubMed: 8057853]
- Possot OM, Pugsley AP. The conserved tetracysteine motif in the general secretory pathway component PulE is required for efficient pullulanase secretion. *Gene.* 1997; 192:45–50. [PubMed: 9224873]
- Py B, Loiseau L, Barras F. An inner membrane platform in the type II secretion machinery of Gram-negative bacteria. *EMBO Rep.* 2001; 2:244–248. [PubMed: 11266368]
- Qadri F, Svennerholm AM, Shamsuzzaman S, Bhuiyan TR, Harris JB, Ghosh AN, Nair GB, Weintraub A, Faruque SM, Ryan ET, Sack DA, Calderwood SB. Reduction in capsular content and enhanced bacterial susceptibility to serum killing of *Vibrio cholerae* O139 associated with the 2002 cholera epidemic in Bangladesh. *Infect Immun.* 2005; 73:6577–6583. [PubMed: 16177333]
- Reichow SL, Korotkov KV, Hol WG, Gonen T. Structure of the cholera toxin secretion channel in its closed state. *Nature structural & molecular biology.* 2010; 17:1226–1232.
- Reindl S, Ghosh A, Williams GJ, Lassak K, Neiner T, Henche AL, Albers SV, Tainer JA. Insights into FlaI Functions in Archaeal Motor Assembly and Motility from Structures, Conformations, and Genetics. *Mol Cell.* 2013; 49:1069–1082. [PubMed: 23416110]
- Rivas S, Bolland S, Cabezon E, Goni FM, de la Cruz F. TrwD, a protein encoded by the IncW plasmid R388, displays an ATP hydrolase activity essential for bacterial conjugation. *J Biol Chem.* 1997; 272:25583–25590. [PubMed: 9325277]
- Robien MA, Krumm BE, Sandkvist M, Hol WG. Crystal structure of the extracellular protein secretion NTPase EpsE of *Vibrio cholerae*. *J Mol Biol.* 2003; 333:657–674. [PubMed: 14556751]
- Rossier O, Dao J, Cianciotto NP. The type II secretion system of *Legionella pneumophila* elaborates two aminopeptidases, as well as a metalloprotease that contributes to differential infection among protozoan hosts. *Applied and environmental microbiology.* 2008; 74:753–761. [PubMed: 18083880]
- Sampaleanu LM, Bonanno JB, Ayers M, Koo J, Tammam S, Burley SK, Almo SC, Burrows LL, Howell PL. Periplasmic domains of *Pseudomonas aeruginosa* PilN and PilO form a stable heterodimeric complex. *J Mol Biol.* 2009; 394:143–159. [PubMed: 19857646]
- Sandkvist M, Bagdasarian M, Howard SP, DiRita VJ. Interaction between the autokinase EpsE and EpsL in the cytoplasmic membrane is required for extracellular secretion in *Vibrio cholerae*. *Embo J.* 1995; 14:1664–1673. [PubMed: 7737119]
- Sandkvist M, Hough LP, Bagdasarian MM, Bagdasarian M. Direct interaction of the EpsL and EpsM proteins of the general secretion apparatus in *Vibrio cholerae*. *J Bacteriol.* 1999; 181:3129–3135. [PubMed: 10322014]
- Sandkvist M, Keith JM, Bagdasarian M, Howard SP. Two regions of EpsL involved in species-specific protein-protein interactions with EpsE and EpsM of the general secretion pathway in *Vibrio cholerae*. *J Bacteriol.* 2000; 182:742–748. [PubMed: 10633109]
- Satyshur KA, Worzalla GA, Meyer LS, Heiniger EK, Aukema KG, Misic AM, Forest KT. Crystal structures of the pilus retraction motor PilT suggest large domain movements and subunit cooperation drive motility. *Structure.* 2007; 15:363–376. [PubMed: 17355871]
- Shiue SJ, Kao KM, Leu WM, Chen LY, Chan NL, Hu NT. XpsE oligomerization triggered by ATP binding, not hydrolysis, leads to its association with XpsL. *Embo J.* 2006; 25:1426–1435. [PubMed: 16525507]
- Sikora AE, Zielke RA, Lawrence DA, Andrews PC, Sandkvist M. Proteomic analysis of the *Vibrio cholerae* type II secretome reveals new proteins, including three related serine proteases. *J Biol Chem.* 2011; 286:16555–16566. [PubMed: 21385872]
- Sobolevsky AI, Rosconi MP, Gouaux E. X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature.* 2009; 462:745–756. [PubMed: 19946266]
- Steffen R, Castelli F, Dieter Nothdurft H, Rombo L, Jane Zuckerman N. Vaccination against enterotoxigenic *Escherichia coli*, a cause of travelers' diarrhea. *Journal of travel medicine.* 2005; 12:102–107. [PubMed: 15996455]

- Toshima H, Hachio M, Ikemoto Y, Ogasawara J, Hase A, Takahashi K, Masaki H, Nishikawa Y. Prevalence of enteric bacteria that inhibit growth of enterohaemorrhagic *Escherichia coli* O157 in humans. *Epidemiology and infection*. 2007; 135:110–117. [PubMed: 16740195]
- Wenneras C, Erling V. Prevalence of enterotoxigenic *Escherichia coli*-associated diarrhoea and carrier state in the developing world. *Journal of health, population, and nutrition*. 2004; 22:370–382.
- Yamagata A, Tainer JA. Hexameric structures of the archaeal secretion ATPase GspE and implications for a universal secretion mechanism. *Embo J*. 2007; 26:878–890. [PubMed: 17255937]
- Yamagata A, Milgotina E, Scanlon K, Craig L, Tainer JA, Donnenberg MS. Structure of an essential type IV pilus biogenesis protein provides insights into pilus and type II secretion systems. *J Mol Biol*. 2012; 419:110–124. [PubMed: 22387466]
- Yanez ME, Korotkov KV, Abendroth J, Hol WG. Structure of the minor pseudopilin EpsH from the Type 2 secretion system of *Vibrio cholerae*. *J Mol Biol*. 2008a; 377:91–103. [PubMed: 18241884]
- Yanez ME, Korotkov KV, Abendroth J, Hol WG. The crystal structure of a binary complex of two pseudopilins: EpsI and EpsJ from the type 2 secretion system of *Vibrio vulnificus*. *J Mol Biol*. 2008b; 375:471–486. [PubMed: 18022192]

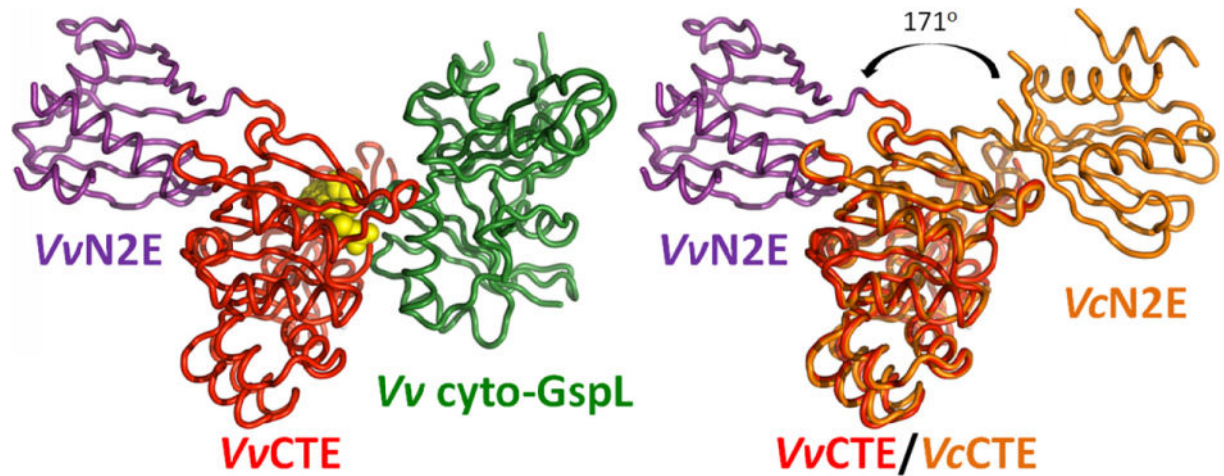
## Abbreviations

<b>AAS</b>	archaella assembly system
<b>BSA</b>	buried surface area
<b>Gsp</b>	general secretion pathway
<b>r.m.s.d</b>	root-mean-square deviation
<b>T2SS</b>	type II secretion system
<b>T4PS</b>	type IV pilus system.
<b>AMPPNP</b>	phosphoaminophosphonic acid-adenylate ester



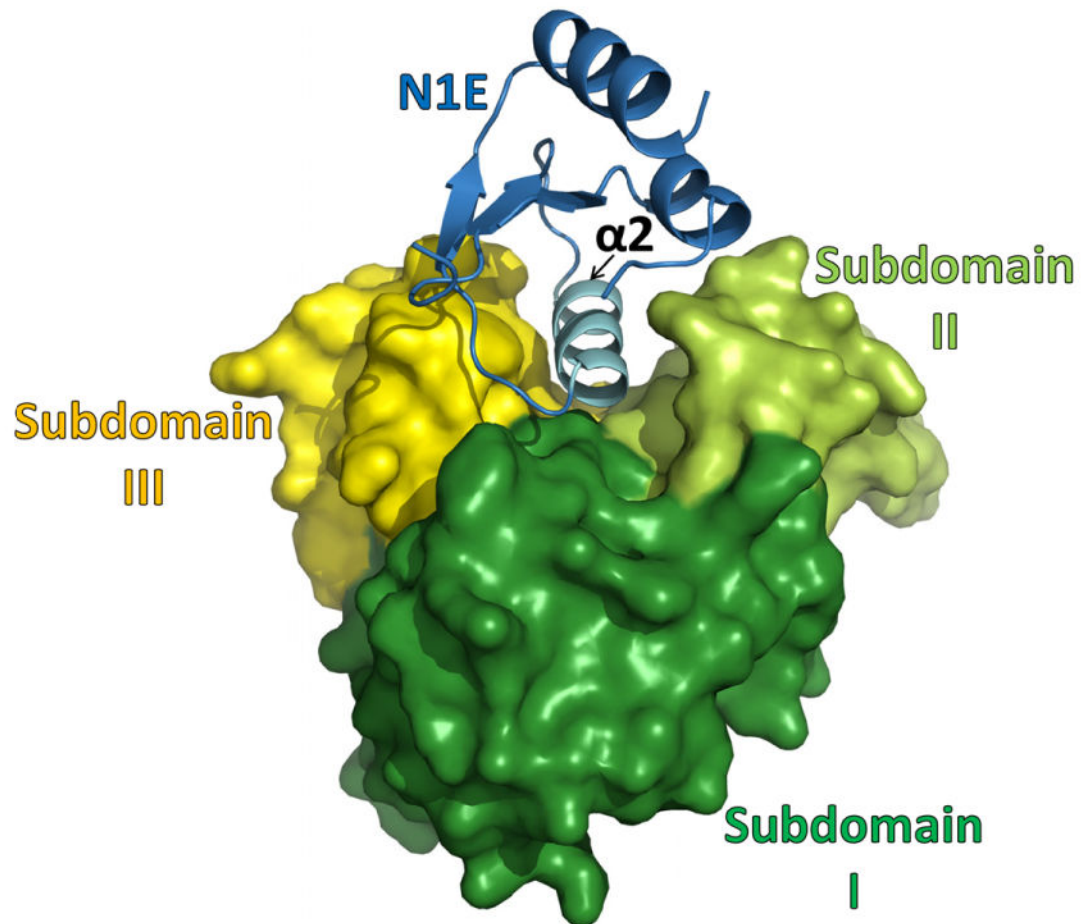
**Fig. 1. The *V. vulnificus* GspE•cyto-GspL complex**

Domains are shown with N1E in blue, cyto-GspL in green, CTE in red, and N2E in pink. AMPPNP atoms are shown as yellow spheres, the essential Zn ion (Robien et al., 2003) as an orange sphere. For nomenclature of the domains, see Supplementary Fig. S1.



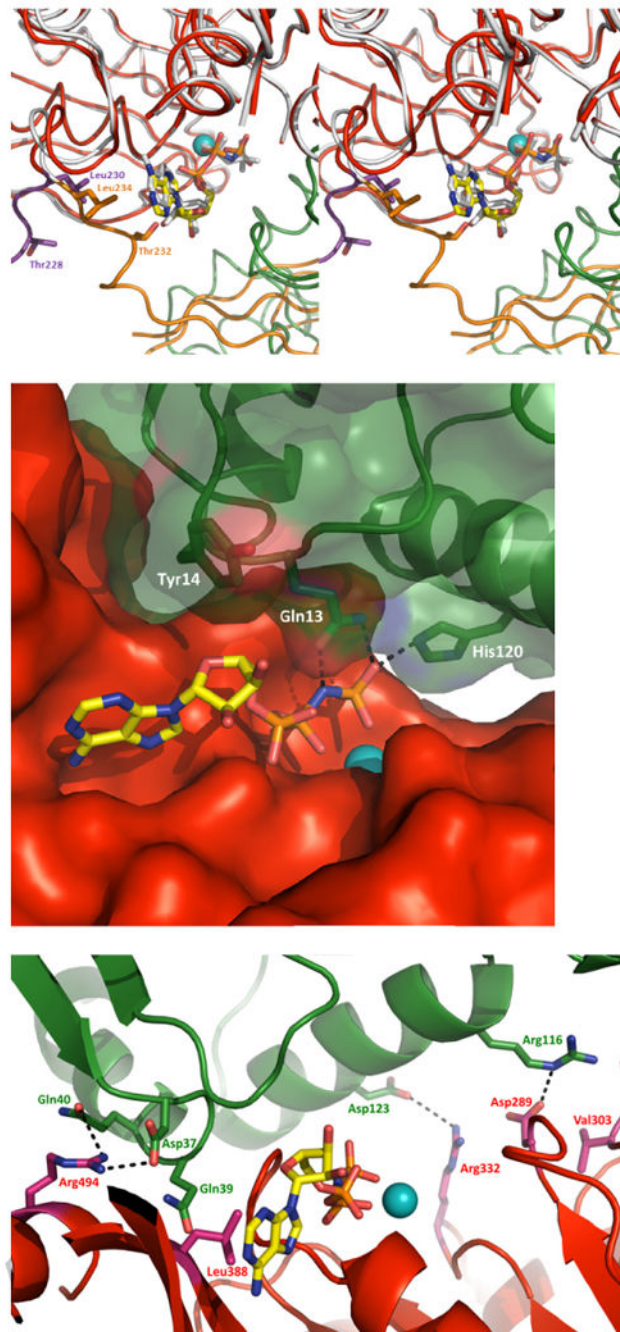
**Fig. 2. The variability of N2E-vs-CTE orientation in GspE**

Left: Overall structure of *V. vulnificus* N2E-CTE (N2E purple, CTE red) in complex with *V. vulnificus* cyto-GspL (green), shown as cartoon with bound AMPPNP (yellow spheres). Right: Comparison of *V. vulnificus* and *V. cholerae* N2E-CTE “di-domains” in the current *V. vulnificus* GspE•cyto-GspL structure and *V. cholerae* <sup>N1E</sup>GspE (Robien et al., 2003). Superimposed are the CTEs from *V. vulnificus* (red) and *V. cholerae* GspE (orange). There is a dramatic change of ~171 degrees in N2E-vs-CTE orientation in these two cases – compare the *V. vulnificus* N2E (purple) and the *V. cholerae* N2E (orange) domains. Comparison with the left panel also shows that *V. vulnificus* cyto-GspL in the current structure occupies approximately the same position as N2E in the *V. cholerae* N2E-CTE didomain structure (Robien et al., 2003).



**Fig. 3. Structure of the *V. vulnificus* N1E•cyto-GspL heterodimer**

The N1E of GspE shown in blue, subdomain I of cyto-GspL in green, subdomain II of cyto-GspL in lime, subdomain III of cyto-GspL in yellow. Note how all three GspL subdomains are involved in contacting N1E with helix  $\alpha 2$  of N1E, a major component of the interface. When compared with the *V. cholerae* N1E•cyto-GspL heterodimer (Abendroth et al., 2005) the r.m.s.d. is 1.4 Å with a difference of ~11 degrees in N1E-*vs*-cyto-GspL orientation, and ~48 and 53 % amino acid sequence identity for N1E and cyto-GspL, respectively (see also Supplementary Fig. S2 and S3).

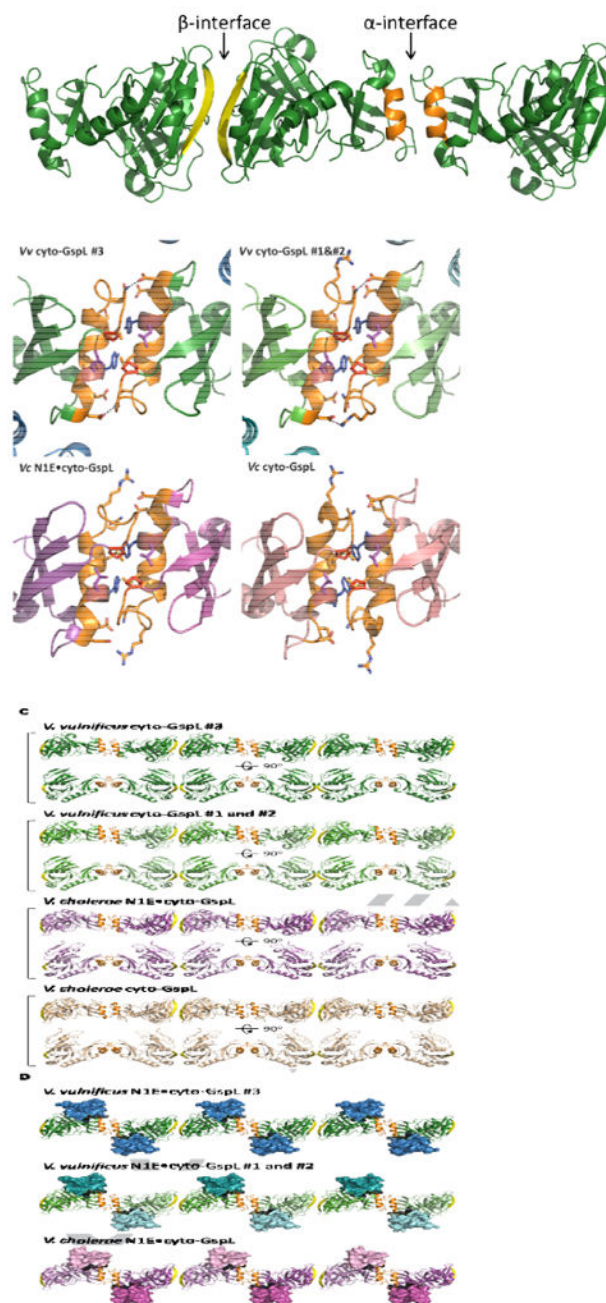


**Fig. 4. Nucleotide binding by *V. vulnificus* GspE•cyto-GspL and by *V. cholerae* GspE**  
**A. Superposition of *V. vulnificus* and *V. cholerae* GspE with AMPPNP bound (stereo figure).** Residues that interact with AMPPNP in the *V. cholerae* GspE•AMPPNP complex but do not interact with AMPPNP in the *V. vulnificus* GspE•cyto-GspL•AMPPNP, due to a major change in conformation of the N2E-CTE linker (225-238), are shown in sticks. *V. vulnificus*: CTE red, N2E-CTE linker purple, cyto-GspL green, AMPPNP as sticks with yellow carbons, and Mg in cyan. *V. cholerae*: N2E and N2E-CTE linker orange, CTE white, and AMPPNP as white sticks (PDB 1P9W). Some CTE residues are removed for clarity.

**B. *V. vulnificus* cyto-GspL•AMPPNP interactions.** Cyto-GspL residues Gln13 and His120 interact with the PNP of AMPPNP. Specifically, the side chains of Gln13 and His120 form hydrogen bonds with the oxygen atoms of  $\gamma$  phosphorous atom with distances of  $\sim 2.5$ - $2.8$  Å. The side-chain oxygen of Gln13 forms a hydrogen bond with a distance of  $\sim 2.7$  Å with the nitrogen linking the  $\beta$  and  $\gamma$  phosphor atoms. Tyr14 makes hydrophobic contacts with the ribose of AMPPNP.

**C. Interactions between *V. vulnificus* Cyto-GspL and CTE.** Key residues contributing to the cyto-GspL and CTE interactions are shown in sticks. The three Arg-Asp salt bridges are indicated with dashed lines. Residues making hydrogen bonds with main chain atoms of another subunit are also labeled.





**Fig. 5. Linear arrangements of *Vibrio* cyto-GspL domains in multiple crystals**  
**A. The  $\alpha$ - and  $\beta$ -interfaces among cyto-GspL domains in the current *V. vulnificus* GspE•cyto-GspL crystals.** The interfaces between cyto-GspL domain #3 and crystallographically related domains are depicted. The interfaces between cyto-GspL domains #1 and #2 are essentially the same (see text). The  $\alpha$ -interface is formed mainly by side chain contacts between residues of  $\alpha 2$  helices (orange) and the loops between strand  $\beta E$  and helix  $\alpha 2$  from two domains. The domains are related by a twofold axis approximately parallel to the direction of view. (See text and Fig. 5B for further description of the

contacts). In the  $\beta$ -interface, main chain hydrogen bonds between antiparallel  $\beta_c$  strands (yellow) are the main contacts between two subunits, which are also related by a twofold approximately parallel to the direction of view.

**B. Close-ups of four similar cyto-GspL  $\alpha$ -interfaces in three different crystal forms.**

Key residues are shown in sticks. Hydrogen bonds and electrostatic interactions are indicated with dashed lines. Note the completely conserved Pro70 (red) in all interfaces, twice in contact with the highly conserved Tyr83 (blue) and Leu84 (purple).

Left upper: The  $\alpha$ -interface between two neighboring cyto-GspL domains #3 in the current *V. vulnificus* GspE•cyto-GspL crystals.

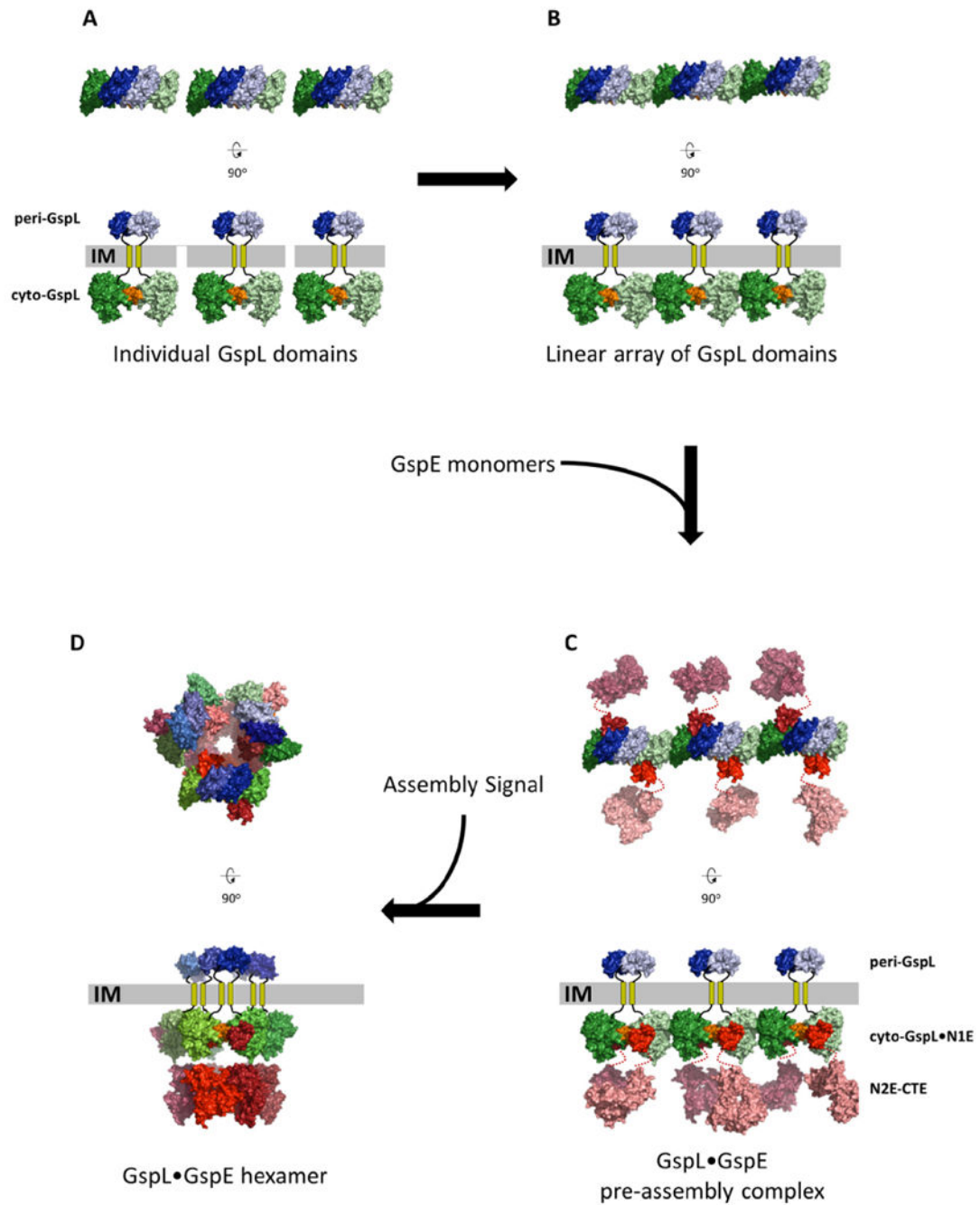
Right upper: The  $\alpha$ -interface between neighboring cyto-GspL domains #1 and #2 in the current *V. vulnificus* GspE•cyto-GspL crystals.

Left lower: The  $\alpha$ -interface between neighboring cyto-GspL domains in the *V. cholerae* N1E•cyto-GspL crystals (PDB 2BH1) (Abendroth et al., 2005).

Right lower: The  $\alpha$ -interface between neighboring cyto-GspL domains in the *V. cholerae* cyto-GspL crystals (PDB 1YF5) (Abendroth et al., 2004a)

C. Four similar linear arrangements of cyto-GspL domains in three different crystal forms. From top to bottom: cyto-GspL rods from: *V. vulnificus* GspE•cyto-GspL complex #3; *V. vulnificus* GspE•cyto-GspL complex #1 and #2; *V. cholerae* N1E•cyto-GspL; and *V. cholerae* cyto-GspL. The  $\alpha$ -interface and  $\beta$ -interfaces are colored orange and yellow, respectively.

**D. Three linear arrangements of cyto-GspL domains with associated N1E domains in two different crystal structures.** From top to bottom: *V. vulnificus* GspE•cyto-GspL complex #3; *V. vulnificus* GspE•cyto-GspL complex #1 and #2; and *V. cholerae* N1E•cyto-GspL. The  $\alpha$ -interface and  $\beta$ -interfaces are colored orange and yellow, respectively. N1E domains are shown in surface representation.



**Fig. 6. A possible pre-assembly complex of the T2SS Inner Membrane Platform**

The upper views are perpendicular to the membrane. The lower views are parallel to the membrane.

**A.** Three separate dimers of GspL with interactions across the  $\alpha$ -interface (orange) of the cytoplasmic domains (green colors) and,  $\epsilon$  p-interface in the periplasmic domains (blue colors).

**B.** Multiple GspL dimers form linear arrays by  $\beta$ -interfaces amongst cytoplasmic domains.

**C.** N1E domains of GspE interact with cyto-GspL domains of a linear array, yielding a pre-assembly complex. Additional proteins like GspM (not shown) might also be part of the pre-assembly complex (see text).

**D.** After an assembly signal, six GspL and six GspE subunits form an assembly with the lower N2E-CTE didomain as a hexamer with approximate  $C_6$  symmetry, connected by N1E-N2E linker residues to the N1E•GspL complex with approximate  $C_3$  symmetry. The latter  $C_3$  axis relates three N1E•GspL dimers with each of these dimers containing an approximate  $C_2$  axis. The aforementioned approximate  $C_6$ ,  $C_3$  and  $C_2$  axes run parallel to each other, perpendicular to the inner membrane plane, with the  $C_6$  and  $C_3$  axes coinciding. The  $C_2$  axes have a different position, approximately related by the  $C_3$  axis.

**Table 1**

Data collection and refinement statistics.

Space Group	C2
Unit cell dimensions	226.4, 133.9, 93.5, 90, 91.4, 90
Resolution range (Å)	43.5-2.83 (2.90-2.83) <sup>a</sup>
Unique reflections	124259 (8574)
Average redundancy	2.0
Completeness (%)	94.3 (87.7)
Rsym	6.1% (51.5%)
$\langle I/\sigma I \rangle$	12.6 (2.0)
Rwork/Rfree	24.8/28.0 (33.3/33.8)
Rms bond length (Å)	0.008
Rms bond angle (Deg)	0.978
Ramachandran outliers <sup>b</sup>	5 (0.26%)
Ramachandran favored	1897 (97.6%)
Number of residues per AU	2238 (three full E•cyto-GspL complexes)
Number of water molecules	16
Average B factors (Å <sup>2</sup> ):	
Protein	59.4
AMPPNP	38.1
Zn	56.3

<sup>a</sup>Values in parenthesis correspond to the highest-resolution shell.<sup>b</sup>Calculated using the Molprobrity server (Chen et al., 2010).

**Table 2**  
**Characteristics of structures containing linear cyto-GspL arrangements**

Structure	Vv #3	Vv #1 and #2	Vc NIE•cyto-GspL	Vc cyto-GspL
Crystallization Condition	0.2M Na Malonate pH7. 18% PEG335 1 mM AMPPNP 2 mM MgCl <sub>2</sub>		0.2M CaOAc <sub>2</sub> 0.1M Bis-Tris-HCl pH 5.75 15-20% PEG3350	0.15M CaOAc <sub>2</sub> 0.1M Tris-HCl pH 8 11-15% PEG6000
Space Group	C2	C2	P2	C2
α-interface	BSA (Å <sup>2</sup> )	960	1150	1102
	r.m.s.d.	-	1.23	1.43
	Kappa	4.1	19.1	16.0
β-interface	BSA (Å <sup>2</sup> )	528	679	766
	r.m.s.d.	-	1.23	1.43
	Kappa	4.1	19.1	16.0

Vv #3: cyto-GspL of *V. vulnificus* GspE•cyto-GspL #3 complex; Vv #1 and #2: cyto-GspL of *V. vulnificus* GspE•cyto-GspL #1 and #2 complex; Vc NIE•cyto-GspL: cyto-GspL of *V. cholerae* NIE•cyto-GspL (PDB 2BH1) (Abendroth et al., 2005); Vc cyto-GspL: *V. cholerae* cyto-GspL (PDB 1YF5) (Abendroth et al., 2004a).

R.m.s.d. and rotation angle (kappa) calculations are based on a superposition of cyto-GspL from three different structures to the cyto-GspL of *V. vulnificus* GspE•cyto-GspL #3 complex. The resultant superposition operation is applied to the entire cyto-GspL•cyto-GspL pairs. The second cyto-GspL in the pairs is subsequently superimposed on the cyto-GspL related to Vv #3 by a crystallographic twofold. The r.m.s.d. and interdomain angle of this second superposition are given in the Table.