Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of a periplasmic domain of the EpsAB fusion protein of the *Vibrio vulnificus* type II secretion system

Dariusz Martynowski,[a] Pawel Grochulski[b] and Peter S. Howard[a]*

[a]Department of Microbiology and Immunology, University of Saskatchewan, 107 Wiggins Road, Saskatoon, SK S7N 2W8, Canada, and [b]Canadian Light Source, 101 Perimeter Road, Saskatoon, SK S7N 0X4, Canada

Correspondence e-mail: peter.howard@usask.ca

Received 25 July 2012
Accepted 12 October 2012

PDB Reference: EpsAB, 4g54

Vibrio vulnificus utilizes the type II secretion system (T2SS), culminating in a megadalton outer membrane complex called the secretin, to translocate extracellular proteins from the periplasmic space across the outer membrane. In *Aeromonas hydrophila*, the general secretion pathway proteins ExeA and ExeB form an inner membrane complex which interacts with peptidoglycan and is required for the assembly of the secretin composed of ExeD. In *V. vulnificus*, these two proteins are fused into one protein, EpsAB. Here, the crystal structure of a periplasmic domain of EpsAB (amino acids 333–584) solved by SAD phasing is presented. The crystals belonged to space group *C*2 and diffracted to 1.55 Å resolution.

1. Introduction

The type II secretion system (T2SS), also known as the main terminal branch of the general secretory pathway (GSP), is utilized by many Gram-negative bacteria, including *Vibrio vulnificus*, to secrete toxins and enzymes from the periplasm into the extracellular environment (Johnson *et al.*, 2006). The number of genes involved in the T2SS is species-dependent and varies from 12 to 16 (Pugsley *et al.*, 1997). The secretion system forms a megadalton assembly that includes a component in the cytoplasm, an inner membrane subcomplex that protrudes into the periplasmic compartment and a secretion pore called the secretin in the outer membrane. The transport of proteins into the extracellular space requires the formation of the secretin complex composed of 12–14 monomers encoded by the *gspD* gene (Bitter *et al.*, 1998). A pseudopilus containing the GspG–GspK proteins may be required to push secreted proteins through the secretin (Hu *et al.*, 2002). This piston-like multimer is anchored to the inner membrane GspE, GspF, GspL and GspM platform (Py *et al.*, 2001; for a review, see Korotkov *et al.*, 2012). Gsp is the generic prefix for the T2SS proteins, but the proteins from different species have also been given specific prefixes such as Exe for *Aeromonas hydrophila* and Eps for *V. vulnificus*.

ExeA and ExeB, which form a complex in the inner membrane of *A. hydrophila*, are required for both localization in the outer membrane and multimerization of ExeD (Howard *et al.*, 2006). ExeA has also been shown to oligomerize in the presence of peptidoglycan, suggesting that an inner membrane heteromultimer of ExeA and ExeB forms a platform for pore formation by the secretin (Li & Howard, 2010; Ast *et al.*, 2002). Recent studies in *Aeromonas* and *Vibrio* species showed that in *V. vulnificus* EpsAB is a natural fusion of the GspA and GspB proteins that is also involved in the assembly of the secretin (Strozen *et al.*, 2011).

Table 1
Summary of data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	Peak	Native
Data-collection parameters		
Beamline	08ID-1, CLS	08ID-1, CLS
Wavelength (Å)	0.97898	0.9793
Temperature (K)	100	100
Oscillation range (°)	1.0	1.0
No. of images	180	180
Data-integration statistics		
Space group	*C*2	*C*2
Unit-cell parameters		
a (Å)	128.54	128.82
b (Å)	44.61	44.52
c (Å)	51.28	51.04
β (°)	110.6	110.6
Resolution limits (Å)	50.00–1.95	47.78–1.55
Total No. of reflections	66962	128758
No. of unique reflections	35882	30885
Multiplicity	1.9 (1.6)	3.5 (2.5)
Completeness (%)	92.2 (62.5)	93.2 (55.8)
$R_{\rm merge}$†	0.069 (0.340)	0.056 (0.386)
Mean $I/\sigma(I)$	12.8 (2.1)	37.0 (2.3)
Molecules per asymmetric unit	1	1
Solvent content (%)	49	49
Refinement		
Resolution (Å)		47.78–1.55
$R_{\rm work}/R_{\rm free}$ (%)		17.7/20.2
No. of reflections		36801
No. of atoms		2062
B factor (Å^2)		23.8
B factor from Wilson plot (Å^2)		22.8
Estimated coordinate error (Å)		0.078
R.m.s.d. bond lengths (Å)		0.028
R.m.s.d. angles (°)		2.45
Ramachandran plot		
Favoured (%)		96.4
Allowed (%)		3.6

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity of the observations $I_i(hkl)$ of reflection *hkl*.

V. vulnificus EpsAB is comprised of an N-terminal/cytoplasmic domain (amino acids 1–295), a short transmembrane domain and a C-terminal periplasmic domain (amino acids 305–718). In *A. hydrophila* the cytoplasmic domain of ExeA has been shown to have ATPase activity *in vitro* and is required for normal secretion (Schoenhofen *et al.*, 2005). Sequence analysis of the periplasmic domain of ExeA revealed the presence of a peptidoglycan-binding domain. Mutation within this region abolished protein secretion and prevented secretin formation, and biochemical analysis demonstrated direct interactions between the purified periplasmic domain of ExeA and peptidoglycan (Li *et al.*, 2011). These findings suggest the importance of interactions of the GspAB complex with peptidoglycan for pore formation and toxin release in *Aeromonas* and *Vibrio* species. Here, we present the structure of a periplasmic domain of *V. vulnificus* EpsAB. The fragment of EpsAB consisting of amino acids 333–584 was cloned and crystallized to study the interactions between key components of the T2SS in Gram-negative bacteria at the atomic level. The structure was solved using SAD phasing and refined to 1.55 Å resolution. The protein structure revealed two distinct subdomains: A1 of unknown function and a peptidoglycan-binding domain containing only one of the two proposed binding-site repeated sequences of bacterial wall peptidoglycan hydrolases (Ghuysen *et al.*, 1994).

2. Materials and methods

2.1. Cloning, expression and initial crystallization trials

The periplasmic domain of *V. vulnificus* EpsAB (amino acids 333–718) was initially cloned with an N-terminal hexahistidine tag into vector pET30a; subsequently, a domain encoding amino acids 333–584 was cloned with an N-terminal hexahistidine tag into pET47 (see below). The clones were transformed into *Escherichia coli* BL21 (DE3) and 20 ml inoculation cultures were grown overnight at 310 K in rich medium containing 50 µg ml^{-1} kanamycin. Large-scale cultures of 0.4–0.8 l were grown at 303 K until an OD_{600} of about 0.8 was reached and were induced with 0.5 m*M* IPTG for 3 h at 303 K. The cells were harvested by centrifugation at 6000*g* for 10 min at 277 K and frozen at 253 K. After purification using Ni–NTA affinity and desalting columns (GE Healthcare Life Sciences), the protein was concentrated to ∼5 mg ml^{-1} in buffer consisting of 75 m*M* NaCl, 20 m*M* Tris pH 8.0. Crystallization trials of the EpsAB (333–718) protein were performed at 277 and 293 K using Hampton Research kits. The crystals obtained showed either no diffraction or salt-like diffraction patterns. The EpsAB (333–718) protein produced by the original clone was also very unstable during purification, resulting in protein loss and degradation (data not shown). Therefore, limited proteolysis was performed to identify fragments of the protein that were potentially more suitable for crystallization.

2.2. Proteolysis and identification of partial proteolysis fragments

A Proti-Ace Kit (Hampton Research) was used to determine fragments of intact protein suitable for crystallization. Initially, six different proteases were tried with different digestion times and variable protease concentrations. Finally, 10 µl of the purified periplasmic domain (10 mg ml^{-1}) was treated with 10 µl elastase (10 µg ml^{-1}) in a solution consisting of 10 m*M* HEPES pH 7.5, 500 m*M* NaCl. After 60 min of incubation at 310 K, the reaction was stopped by adding 1 µl glacial acetic acid. Proteolytic fragments were passed through a 1 ml Mono Q ion-exchange column (GE Healthcare Life Sciences) using a 0–100% gradient from buffer *A* (20 m*M* Tris pH 8.0) to buffer *B* (buffer *A* plus 500 m*M* NaCl). Selected fractions containing pure fragments of intact protein were analysed by mass spectrometry (MALDI-MS), which indicated that a fragment composed of EpsAB amino acids 333–584 was partially resistant to proteolysis; this was therefore cloned into pET47 and used in subsequent purifications.

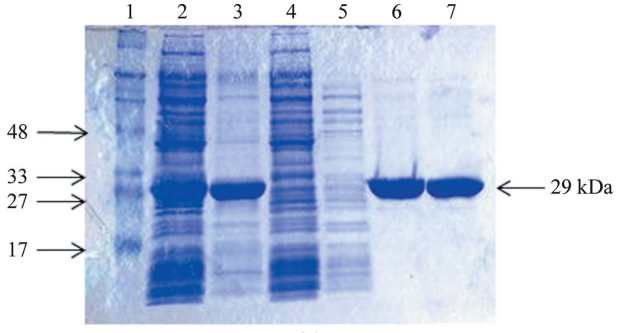
2.3. Purification of EpsAB (333–584)

The pET47 plasmid expressing the periplasmic domain of EpsAB (amino acids 333–584) from *V. vulnificus* was transformed into *E. coli* BL21 (DE3) cells and expressed as

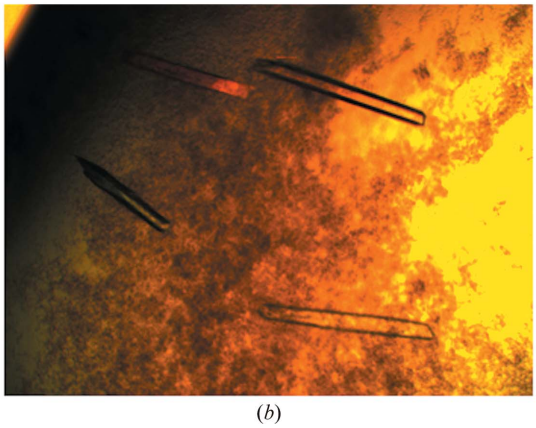
described above. The cell pellets were resuspended in 10 ml 20 m*M* Tris pH 8.0, 250 m*M* NaCl, 10% glycerol, 10 m*M* imidazole (buffer *A*) supplemented with 40 µl 1 mg ml^{-1} DNase, 100 µl 1 mg ml^{-1} RNase, 400 µl 25× EDTA-free protease-inhibitor cocktail, 100 µl 0.5 *M* $MgCl_2$ and passed twice through a French pressure cell at 6.9 MPa. The lysate was centrifuged at 12 000*g* for 30 min at 277 K to remove cell debris. EpsAB (330–584) was purified using nickel-affinity chromatography (GE Healthcare Life Sciences) with a gradient of 0–100% buffer *B* (buffer *A* + 500 m*M* imidazole) followed by desalting in 20 m*M* Tris pH 8.0 (buffer *C*; Fig. 1*a*). Initially, after removing imidazole, ion exchange was performed with a Resource Q column (GE Healthcare Life Sciences) using a gradient of 0–100% buffer *D* (buffer *C* + 500 m*M* NaCl). This process resulted in protein loss and degradation (data not shown) and was abandoned for further purification experiments. After adding NaCl to a final concentration of 75 m*M* (final protein buffer: 75 m*M* NaCl, 20 m*M* Tris pH 8.0), the Ni–NTA affinity-purified and desalted EpsAB (333–584) sample was concentrated to 5 mg ml^{-1} and used for crystallization as described below.

2.4. Selenomethionine-labelled protein expression and purification

For selenomethionine-labelled protein, an M9 SeMet high-yield medium kit (Medicilion) was used to overexpress EpsAB (333–584) protein. One pouch of Kit A (M9 salts and amino acids) was resuspended in 950 ml water and 50 µg ml^{-1}

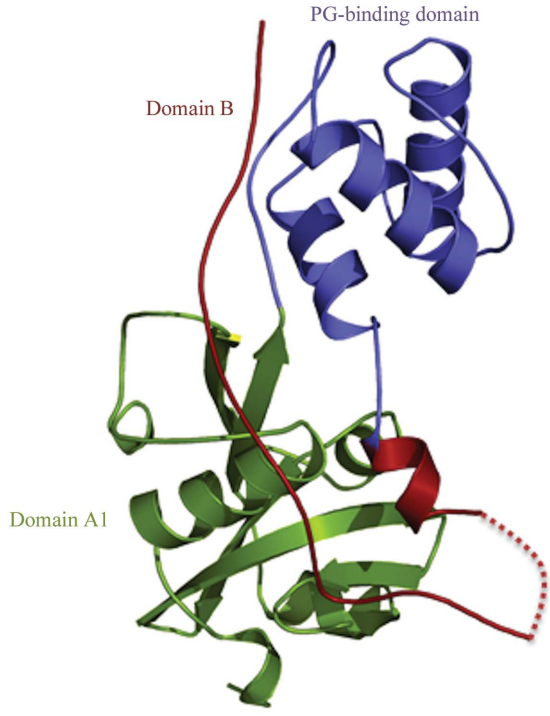


(*a*)

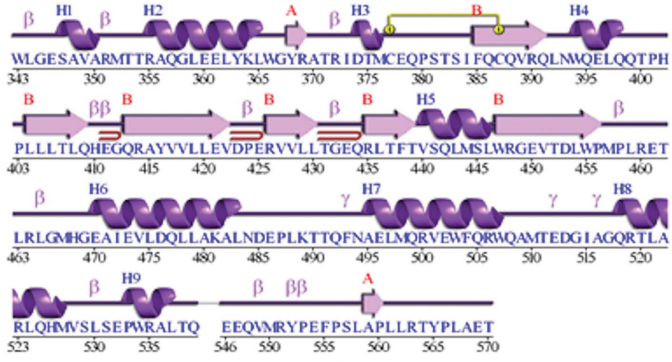


(*b*)

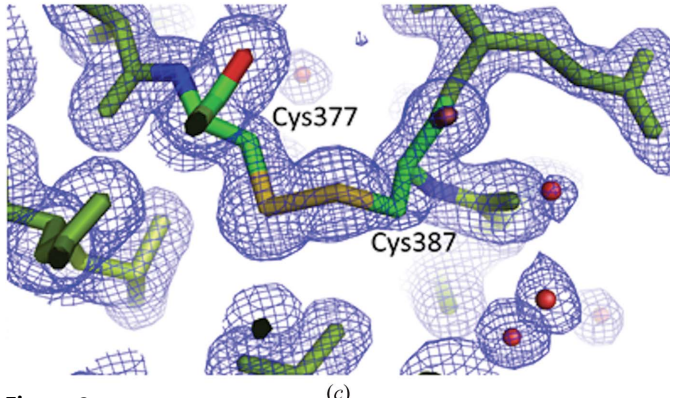
Figure 1
(*a*) SDS–PAGE analysis of EpsAB (333–584). Lane 1, markers (labelled in kDa); lane 2, supernatant; lane 3, pellet; lane 4, wash fraction; lane 5, wash fraction; lane 6, Ni–NTA elution peak; lane 7, desalted peak. (*b*) Crystals of EpsAB (333–584) from *V. vulnificus* grown by hanging-drop vapour diffusion at room temperature.



(*a*)



(*b*)



(*c*)

Figure 2
(*a*) Overall structure of *V. vulnificus* EpsAB (333–584). Green, N-terminal A1 domain; violet, peptidoglycan-binding domain; red, N-terminal fragment of B domain; dashed red, residues not observed in the electron-density map. (*b*) Sequence of *V. vulnificus* EpsAB. (*c*) A representative example of the quality of the final $2F_o - F_c$ electron-density map (blue, 1.55 Å resolution, contoured at 1.0σ) and the disulfide bridge between Cys377 and Cys387.

kanamycin. 10 ml Kit B (mineral supplements), 10 ml 50% glycerol and 10 ml inoculated culture were added. Cells were grown in the dark at 310 K until an OD_{600} of about 1.3 was reached and were induced with 0.5 m*M* IPTG overnight at 295 K after adding 20 ml Kit D (inhibitory amino-acid cocktail and SeMet). The purification of this derivative was the same as described above for the native protein using Ni–NTA affinity and desalting columns. The sample was concentrated to a final concentration of ∼5 mg ml^{-1} in 75 m*M* NaCl, 20 m*M* Tris pH 8.0 buffer.

2.5. Crystallization

Crystallization of both native and SeMet-labelled proteins was performed by the hanging-drop vapour-diffusion method at 295 K in 24-well plates. Each drop consisted of 1 µl protein solution and 1 µl reservoir solution, with 500 µl reservoir solution in the well. Initial screening was carried out using PEG/Ion kits (Hampton Research) and positive hits were then optimized. The optimized reservoir solution consisted of 0.16–0.26 *M* calcium chloride, 20%(*w*/*v*) polyethylene glycol (PEG) 3350. Rod-shaped crystals with maximum dimensions of 0.3 × 0.1 × 0.1 mm were obtained within 5 d (Fig. 1*b*). For selenomethionine samples, 1.5 µl each of the protein solution and the reservoir solution were used.

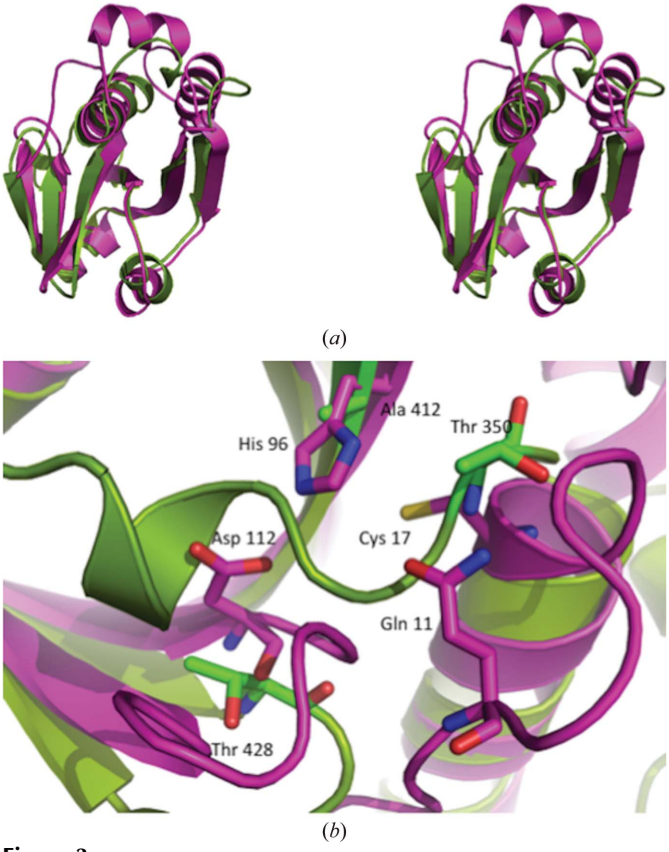


Figure 3
(*a*) Stereoview of the superposition of the A1 domain of *V. vulnificus* EpsAB (green) and 3k8u (pink). (*b*) Superposition of the active site of the peptidase domain of ComA (pink) and EpsAB A1 (green). The residues involved in peptidase activity are Gln11, Cys17, His96 and Asp112 (pink) and the corresponding residues in EpsAB A1 are Thr350, Ala412 and Thr428 (green). Gln11 of ComA does not have a structurally equivalent residue in EpsAB A1.

2.6. Structure solution and refinement

Selected crystals were transferred into a solution consisting of 0.2 *M* $CaCl_2$, 40% PEG 3350 and were flash-cooled in liquid nitrogen. Diffraction data were collected at 100 K using a Rayonix MX300 CCD detector on beamline 08ID-1 at the Canadian Light Source (CLS; Grochulski *et al.*, 2011). A MAD scan with *MxDC* (Fodje *et al.*, 2012) was performed to determine the energies for peak, inflection and remote edges. A total of 180 frames per edge of data were collected with an oscillation angle of 1.0°. The data were indexed, integrated and scaled with the *HKL*-2000 software package (Otwinowski & Minor, 1997). X-ray data-collection and model-refinement statistics are summarized in Table 1.

The structure was solved by SAD phasing with *AutoSol* from *PHENIX* (Adams *et al.*, 2010) using the peak data set only. Eight out of the nine potential sites of selenomethionine incorporation were found to provide initial phases for structure solution and refinement. Experimentally determined maps at 1.95 Å resolution allowed automated building of the SeMet-containing protein with *ARP*/*wARP* (Langer *et al.*, 2008), resulting in the docking of 204 out of 251 residues into electron density. The native EpsAB crystals diffracted to a resolution of 1.55 Å in space group *C*2, with unit-cell parameters a = 28.82, b = 44.52, c = 51.04 Å, β = 110.6°, and contained one molecule per asymmetric unit. The SAD-derived model was used as the search structure for molecular replacement with *MOLREP* (Vagin & Teplyakov, 2010) using the native data set. Several rounds of iterative model building and refinement were performed using *Coot* (Emsley *et al.*, 2010) and *REFMAC*5 (Murshudov *et al.*, 2011); 5% of the data were excluded from refinement and used for cross-validation. The final model contained 213 of the 251 residues of the *V. vulnificus* EpsAB (333–584) fragment, with both the N-terminal and the C-terminal ends of the protein not being visible in the 1.55 Å resolution electron-density map.

Subsequent rounds of refinement included TLS refinement with seven groups of peptides in the monomer (Painter & Merritt, 2006), resulting in an R value of 17.8% and an R_{free} of 20.2%. Geometry was analysed using *PROCHECK* (Laskowski *et al.*, 1993). The final structure was deposited in the Protein Data Bank as entry 4g54.

3. Results and discussion

3.1. Structural overview

The originally expressed EpsAB (305–718) protein was very unstable during purification, resulting in protein loss and degradation (data not shown). Limited proteolysis was performed to determine fragments of the intact protein that were suitable for purification and crystallization. EpsAB (333–584) was then purified and crystallized in space group *C*2 with one molecule per asymmetric unit.

The structure of EpsAB was determined to a resolution of 1.55 Å with an R_{work} of 17.8% (R_{free} = 20.2%) and root-mean square deviations (r.m.s.d.s) from ideal values of bond lengths and bond angles of 0.028 Å and 2.45°, respectively (Table 1). The final structure included residues 343–570 and 253 water molecules.

The EpsAB fusion protein comprises two major domains, A (343–533) and B (534–570), with domain A containing an N-terminal subdomain A1 (343–456) of unknown function and a peptidoglycan-binding domain (457–533). The C-terminal part of the EpsAB fragment (part of the B domain) includes an α-helix (534–537) and a long loop (538–570), both of which interact with the A domain and stabilize the structure (Fig. 2*a*). Six internal residues (540–545) were not observed in the electron-density map. Analysis of the coordinates using the *PDBePISA* server did not reveal any specific interactions that could result in the formation of stable quaternary structures.

3.1.1. Domain A1 of unknown function. The structure of the A1 subdomain has an α/β fold with six β-strands and five α-helices (Fig. 2*a*). The structure is organized around a central six-stranded antiparallel β-sheet with α-helices packing on two sides of the sheet, with the following order of these components: α1, α2, β1, α3, β2, α4, β3, β4, β5, α5, β6.

A search for similar structures in the Protein Data Bank with *DALI* (Holm & Rosenström, 2010) found among its top hits several members of the peptidase-like family. These include the peptidase domain of *Streptococcus* ComA (PDB entry 3k8u; *Z*-score = 11.1; Ishii *et al.*, 2010), the peptidase C39-like domain of *V. parahaemolyticus* (PDB entry 3b79; *Z*-score = 10.8; Midwest Center for Structural Genomics, unpublished work), the putative C39-like domain of *Bacillus anthracis* (PDB entry 3erv; *Z*-score = 7.4; New York SGX Research Center for Structural Genomics, unpublished work) and papain (PDB entry 1stf; *Z*-score = 7.2; Stubbs *et al.*, 1990). The overall r.m.s.d. values of the superimposable C^{α} positions (98–101 out of 114 in EpsAB) between *V. vulnificus* EpsAB and these structures ranged from 2.3 to 3.1 Å and the level of sequence identity ranged from 10 to 13%. Comparison with these structural homologues showed strong topological similarity in the core β-sheet, as shown for ComA in Fig. 3(*a*), although a region of higher structural differences exists between residues 32–64 of ComA and residues 366–382 of EpsAB. In addition, although the overall structure is very similar, the residues involved in peptidase activity (Gln11, Cys17, His96 and Asp112 of ComA) are not conserved in the A1 domain of the EpsAB fusion and the reaction pocket is not accessible owing to the presence of helix α1 and the following loop in the position of the peptidase reaction site (Fig. 3*b*).

Two-codon insertion analysis of ExeA, the EpsAB homologue from *A. hydrophila*, revealed that mutations in A1 decreased the ability of the cells

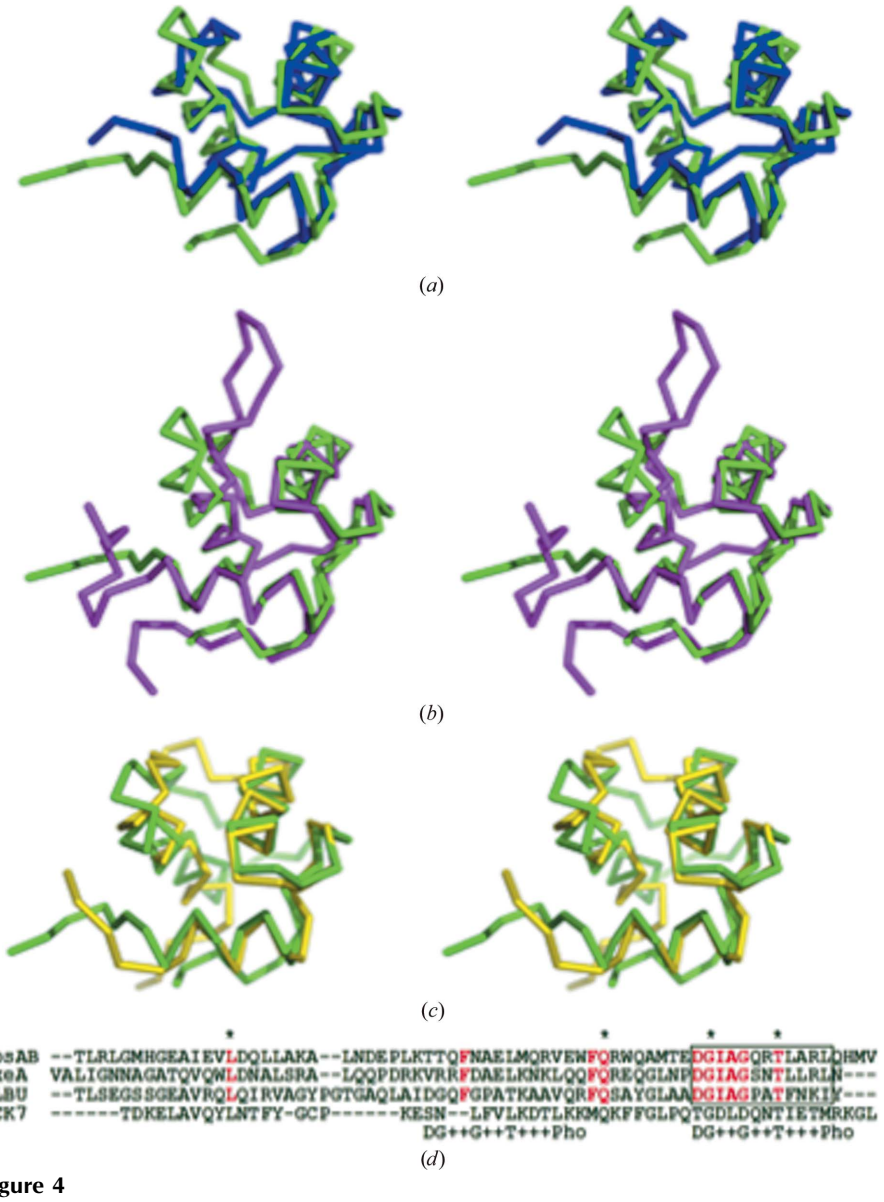


Figure 4
Structure of the peptidoglycan-binding domain of EpsAB. Comparison of the EpsAB peptidoglycan-binding domain (green) with (*a*) the bacteriophage φKZ lytic transglycosylate qp144 N-terminal domain (blue; PDB entry 3kbh), (*b*) the N-terminal region of D-Ala-D-Ala-cleaving carboxypeptidase from *S. albus* G (violet; PDB entry 1lbu) and (*c*) the prodomain of human matrix metalloproteinase (yellow; PDB entry 1ck7). (*d*) Sequence alignment of the peptidoglycan-binding domains of *V. vulnificus* EpsAB, *A. hydrophila* ExeA, *S. albus* carboxypeptidase (PDB entry 1lbu) and human MMP2 (PDB entry 1ck7). Red, residues conserved in the bacterial proteins; residues conserved in all four proteins are indicated by stars. The black box indicates the second repeat of the peptidoglycan-binding site.

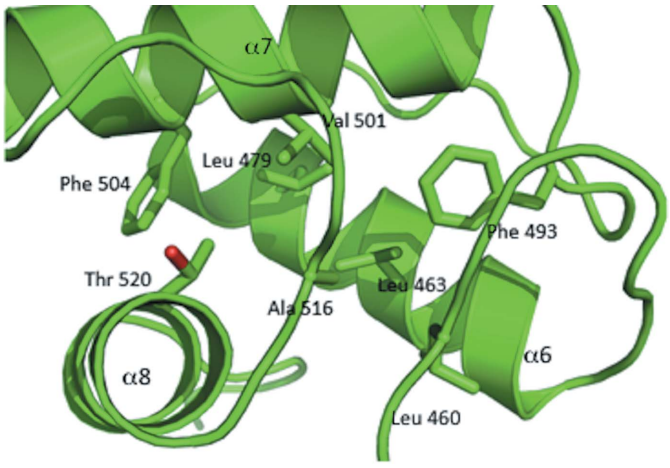


Figure 5
Conserved hydrophobic residues involved in the folding of the peptidoglycan-binding domain in bacterial proteins.

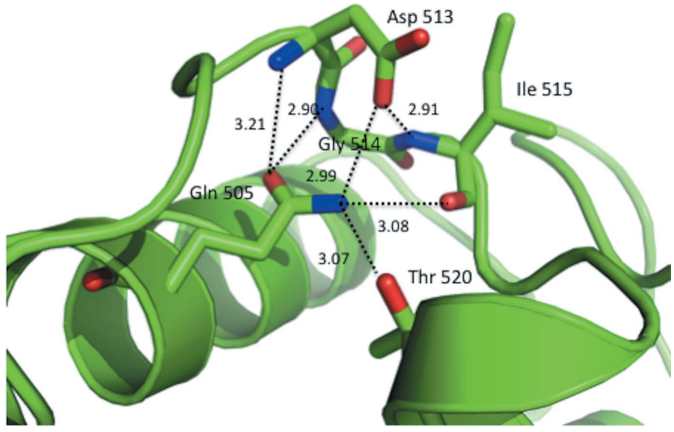


Figure 6
Hydrogen-bond interactions between the highly conserved Gln505 residue and Asp513, Gly514, Leu515 and Thr520.

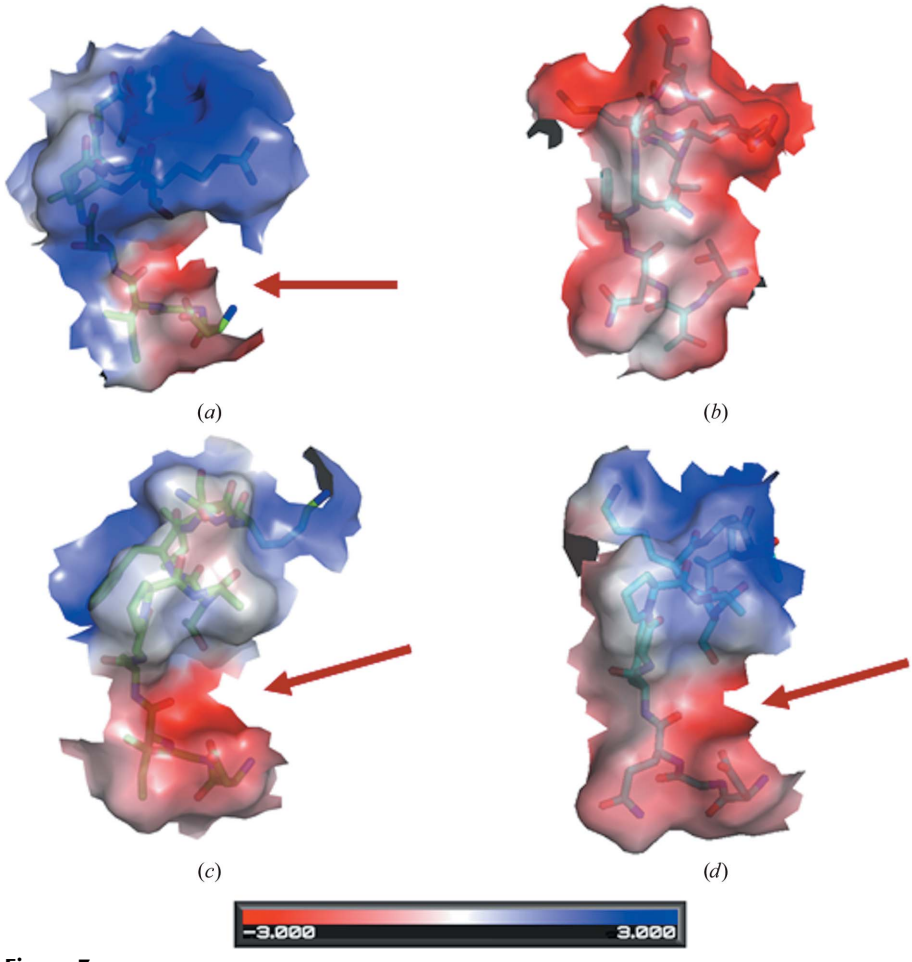


Figure 7
Electrostatic potential mapped onto the solvent-accessible surface of the peptidoglycan-binding sites. Blue indicates positive charge and red indicates negative charge. The red arrow indicates the pocket where the sugar moiety may interact with protein. (*a*) EpsAB binding site II, (*b*) EpsAB binding site I, (*c*) muramoyl-pentapeptide carboxypeptidase site II, (*d*) carboxypeptidase site I.

to secrete aerolysin and form secretin, but did not influence interaction with peptidoglycan (Howard *et al.*, 2006). Mapping those insertions onto the structure of EpsAB located them on the surface of the A1 subdomain, suggesting that A1 may be involved in ExeAB complex formation and is not required for peptidoglycan binding.

3.1.2. Peptidoglycan-binding domain. The *DALI* (Holm & Rosenström, 2010) search for structural homologues revealed several proteins deposited in the Protein Data Bank involved in peptidoglycan recognition and also the prodomain of human matrix metalloproteinase MMP2. The proteins interacting with peptidoglycan include bacteriophage φKZ lytic transglycosylase qp144 N-terminal domain (PDB entry 3kbh; *Z*-score = 11.6; Fokine *et al.*, 2008), with 28% identity and an overall r.m.s.d. value of superimposable C^{α} positions (72 out of 76) of 1.7 Å, and the N-terminal domain of the D-Ala-D-Ala-cleaving carboxypeptidase from *Streptomyces albus* (PDB entry 1lbu; *Z*-score = 11.0; Dideberg *et al.*, 1982), with 26% identity and an r.m.s.d. of 2.4 Å (Figs. 4*a* and 4*b*). However, regions of higher r.m.s.d. exist between the PG-binding domain of EpsAB and the transglycosylase at amino aicds 483–493 of EpsAB and amino acids 24–37 of 3kbh, and between EpsAB and the carboxypeptidase at amino acids 475–486 of EpsAB and amino acids 25–38 of 1lbu. A domain with a similar structure to the C-terminal region of the periplasmic domain of EpsAB with 16% sequence similarity occurs in the human matrix metalloproteinase (MMP2; PDB

entry 1ck7; *Z*-score = 6.0; Morgunova *et al.*, 1999). Superposition of the EpsAB peptidoglycan-binding domain onto the prodomain of human MMP2 gave an r.m.s.d. of 2.6 Å for 58 out of 76 C^{α} atoms (Fig. 4*c*). Greater differences in structure between the EpsAB PG-binding domain and MMP2 occur in two regions: amino acids 445–470, corresponding to amino acids 42–48 of MMP2, and amino acids 484–491, corresponding to amino acids 60–67 of MMP2. This eukaryotic protein also shares lower overall sequence similarity with the prokaryotic homologues identified by *DALI* (Fig. 4*d*). The MMP2s degrade extracellular matrix proteins and are considered to play an important role in tumour invasion and arthritis (Kohn & Liotta, 1995).

The C-terminal subdomain of EpsAB belongs to the peptidoglycan-binding domain family (Pfam PF01471) and features three α-helices: α6, α7 and α8. Several conserved residues among bacterial proteins containing this binding site (Li *et al.*, 2011) are Leu460, Leu463, Leu479, Phe493, Val501, Phe504, Ala516 and Thr520, and these are involved in hydrophobic interactions and proper protein folding, as shown in Fig. 5.

Based on the crystal structure of the carboxypeptidase 1lbu, it was proposed that two DG++G++T+++Pho (where Pho indicates hydrophobic residues) repeats are involved in peptidoglycan binding (Ghuysen *et al.*, 1994). However, in a recent genetic study of the structure/function of the peptidoglycan-binding site of *A. hydrophila* ExeA (Li *et al.*, 2011), it was proposed that the proteins containing the site form three evolutionary families: those containing two binding repeats, those containing only the first and those (including ExeA) containing only the second. The ExeA studies revealed that mutations Q488A, D496A, T503A and L507A (corresponding to Gln505, Asp513, Thr520 and Leu524 in *V. vulnificus* EpsAB) within the second repeat dramatically reduced lipase secretion. Furthur analysis of these mutations showed defects in ExeD multimerization and decreased binding of PG. The EpsAB structure determined here shows that the highly conserved residue Gln505 forms a hydrogen-bond net involving residues Asp513, Gly514, Ile515 and Thr520 of the binding motif **DGI**+G++**T**+++Pho (Fig. 6) and thus explains the demonstrated importance of Gln505 in the formation of the biologically functional PG-binding pocket. Notably, the Thr520 residue plays an important role in both hydrophobic interactions and hydrogen-bond formation. The prodomain of human MMP2, although a member of Pfam PF01471, does not show the presence of the proposed peptidoglycan-binding motif (Fig. 4*d*).

Superposition of the peptidoglycan-binding site repeat 1 that is not conserved in EpsAB and its homologues onto the highly conserved second repeat shows structural differences between the peptidoglycan-binding regions of the proteins which contain both motifs and the same region in EpsAB. The electrostatic potential mapped onto the solvent-accessible surface shows differences in the shape and electrostatic character of the repeat 1 and repeat 2 regions of EpsAB (Figs. 7*a* and 7*b*). The upper area of the region containing repeat 2 is positively charged, while the lower area exhibits negatively charged properties and the presence of a distinctive pocket that may bind to the sugar moiety of peptidoglycan. This differs from site 1, which is mostly negatively charged. Comparison to the N-terminal region of the D-Ala-D-Ala carboxypeptidase from *S. albus* reveals similarity in electrostatic properties between the EpsAB peptidoglycan-binding motif and the carboxypeptidase repeats 1 and 2, with the upper half of both of the latter sites being positively charged and the potential binding pocket area being negatively charged (Fig. 7). Further studies of the interactions between peptidoglycan and these proteins are required in order to fully understand the mechanism of peptidoglycan binding and its role in secretin formation.

This research was supported by a Saskatchewan Health Research Foundation Group grant to the Molecular Design Research Group of the University of Saskatchewan and by a Natural Sciences and Engineering Research Council of Canada Discovery grant to SPH. The Canadian Light Source is supported by the Natural Sciences and Engineering Research Council of Canada, the National Research Council Canada, the Canadian Institutes of Health Research, the Province of Saskatchewan, Western Economic Diversification Canada and the University of Saskatchewan.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Ast, V. M., Schoenhofen, I. C., Langen, G. R., Stratilo, C. W., Chamberlain, M. D. & Howard, S. P. (2002). *Mol. Microbiol.* **44**, 217–231.
Bitter, W., Koster, M., Latijnhouwers, M., de Cock, H. & Tommassen, J. (1998). *Mol. Microbiol.* **27**, 209–219.
Dideberg, O., Charlier, P., Dive, G., Joris, B., Frère, J.-M. & Ghuysen, J.-M. (1982). *Nature (London)*, **299**, 469–470.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Fodje, M., Janzen, K., Berg, R., Black, G., Labiuk, S., Gorin, J. & Grochulski, P. (2012). *J. Synchrotron Rad.* **19**, 274–280.
Fokine, A., Miroshnikov, K. A., Shneider, M. M., Mesyanzhinov, V. V. & Rossmann, M. G. (2008). *J. Biol. Chem.* **283**, 7242–7250.
Ghuysen, J.-M., Lamotte-Brasseur, J., Joris, B. & Shockman, G. D. (1994). *FEBS Lett.* **342**, 23–28.
Grochulski, P., Fodje, M. N., Gorin, J., Labiuk, S. L. & Berg, R. (2011). *J. Synchrotron Rad.* **18**, 681–684.
Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
Howard, S. P., Gebhart, C., Langen, G. R., Li, G. & Strozen, T. G. (2006). *Mol. Microbiol.* **59**, 1062–1072.
Hu, N.-T., Leu, W.-M., Lee, M.-S., Chen, A., Chen, S.-C., Song, Y.-L. & Chen, L.-Y. (2002). *Biochem. J.* **365**, 205–211.
Ishii, S., Yano, T., Ebihara, A., Okamoto, A., Manzoku, M. & Hayashi, H. (2010). *J. Biol. Chem.* **285**, 10777–10785.
Johnson, T. L., Abendroth, J., Hol, W. G. J. & Sandkvist, M. (2006). *FEMS Microbiol. Lett.* **255**, 175–186.
Kohn, E. C. & Liotta, L. A. (1995). *Cancer Res.* **55**, 1856–1862.
Korotkov, K. V., Sandkvist, M. & Hol, W. G. J. (2012). *Nature Rev. Microbiol.* **10**, 336–351.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Li, G. & Howard, S. P. (2010). *Mol. Microbiol.* **76**, 772–781.

Li, G., Miller, A., Bull, H. & Howard, S. P. (2011). *J. Bacteriol.* **193**, 197–204.
Morgunova, E., Tuuttila, A., Bergmann, U., Isupov, M., Lindqvist, Y., Schneider, G. & Tryggvason, K. (1999). *Science*, **284**, 1667–1670.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Painter, J. & Merritt, E. A. (2006). *Acta Cryst.* D**62**, 439–450.
Pugsley, A. P., Francetic, O., Hardie, K., Possot, O. M., Sauvonnet, N. & Seydel, A. (1997). *Folia Microbiol. (Praha)*, **42**, 184–192.
Py, B., Loiseau, L. & Barras, F. (2001). *EMBO Rep.* **2**, 244–248.
Schoenhofen, I. C., Li, G., Strozen, T. G. & Howard, S. P. (2005). *J. Bacteriol.* **187**, 6370–6378.
Strozen, T. G., Stanley, H., Gu, Y., Boyd, J., Bagdasarian, M., Sandkvist, M. & Howard, S. P. (2011). *J. Bacteriol.* **193**, 2322–2331.
Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B. & Turk, V. (1990). *EMBO J.* **9**, 1939–1947.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.