The crystal structure of the periplasmic domain of *Vibrio parahaemolyticus* CpxA

Eunju Kwon,[1,2] Dong Young Kim,[2] Tri Duc Ngo,[1] Carol A. Gross,[3] John D. Gross,[2]* and Kyeong Kyu Kim[1]*

[1]Department of Molecular Cell Biology, Samsung Biomedical Research Institute, Sungkyunkwan University School of Medicine, Suwon 440-746, Korea

[2]Department of Pharmaceutical Chemistry, University of California, San Francisco, 600 16th Street, San Francisco, California 94107

[3]Department of Microbiology & Immunology, University of California, San Francisco, 600 16th Street, San Francisco, California 94143

Received 6 April 2012; Revised 21 June 2012; Accepted 25 June 2012
DOI: 10.1002/pro.2120
Published online 3 July 2012 proteinscience.org

Abstract: The Cpx two-component system of Gram-negative bacteria senses extracytoplasmic stresses using the histidine kinase CpxA, a membrane-bound sensor, and controls the transcription of the genes involved in stress response by the cytosolic response regulator CpxR, which is activated by the phosphorelay from CpxA. CpxP, a CpxA-associated protein, also plays an important role in the regulation of the Cpx system by inhibiting the autophosphorylation of CpxA. Although the stress signals and physiological roles of the Cpx system have been extensively studied, the lack of structural information has limited the understanding of the detailed mechanism of ligand binding and regulation of CpxA. In this study, we solved the crystal structure of the periplasmic domain of *Vibrio parahaemolyticus* CpxA (VpCpxA-peri) to a resolution of 2.1 Å and investigated its interaction with CpxP. VpCpxA-peri has a globular Per-ARNT-SIM (PAS) domain and a protruded C-terminal tail, which may be required for ligand sensing and CpxP binding, respectively. The direct interaction of the PAS core of VpCpxA-peri with VpCpxP was not detected by NMR, suggesting that the C-terminal tail or other factors, such as the membrane environment, are necessary for the binding of CpxA to CpxP.

Keywords: Gram-negative bacterium; PAS domain; extracytoplasmic stress response; two-component system; periplasm

Introduction

Gram-negative bacteria have developed extracytoplasmic stress response (ESR) systems to deal with environmental stress.[1–7] The Cpx system is one of several bacterial ESR systems operated by a two-component signaling system that is abundantly used for bacteria to sense, respond, and adapt to environmental changes. The two-component system generally consists of a sensory histidine kinase (HK) and a response regulator (RR) as basal elements.[8] In the Cpx system, CpxA is the inner membrane-spanning, sensory HK.[9] Recognition of a specific envelope stress by the periplasmic domain of CpxA triggers signal transduction to the cytoplasmic domain, which results in the autophosphorylation of a conserved histidine residue (His248 in *Escherichia coli* CpxA). Then, the phosphate group is transferred to its cognate RR, CpxR, which controls the transcription of the genes involved in the ESR.[10]

Many factors seemingly unrelated to each other are responsible for the activation of the Cpx

The coordinate and structure factors have been deposited in the Research Collaboratory for Structural Bioinformatics PDB with accession number 3V67.

Grant sponsor: Next-Generation BioGreen 21 Program; Grant number: SSAC PJ008107; Grant sponsor: Korea Healthcare Technology R&D Project; Grant number: A092006; Grant sponsor: a National Research Foundation of Korea (NRF); Grant number: 2011-0028878.

*Correspondence to: Kyeong Kyu Kim, Ph.D., Department of Molecular Cell Biology, Samsung Biomedical Research Institute, Sungkyunkwan University School of Medicine, Suwon 440-746, Korea. E-mail: kyeongkyu@skku.edu or John D. Gross, Ph.D., Department of Pharmaceutical Chemistry, University of California San Francisco, 600 16th Street, San Francisco, CA 94107, USA. E-mail: jdgross@cgl.ucsf.edu

pathway in *E. coli*.[1–3,7] Adhesion to abiotic surfaces triggers the Cpx response via NlpE, an outer membrane lipoprotein that is overexpressed by LacZ accumulation in the periplasm and alleviates the toxicity of periplasmic LacZ in a CpxAR-dependent manner.[11,12] YafY, an inner membrane lipoprotein, is also reported to be responsible for the Cpx response.[13] The Cpx pathway can also be activated in an NlpE-independent manner by various factors, such as bacterial growth,[14] misfolded or aggregated pili subunits,[15] alkaline pH,[16] and the presence of heavy metals.[17,18] However, to date, the ligand responsible for the activation of CpxA by direct binding has not been identified. The Cpx signaling system and the Bae two-component system, another ESR system, share spheroplast formation as an envelope stress signal.[19] In addition, both the Cpx and Arc two-component systems are commonly activated by antibiotic-mediated oxidative stress.[20]

The Cpx pathway regulates a variety of genes responsible for the envelope stress response and bacterial virulence. Approximately 100 genes were predicted as members of the *cpx* regulon based on bioinformatic predictions of CpxR-binding sites.[21] More recently, the systemic screening of known CpxR-binding promoters revealed that more than 40 genes are under the control of CpxR.[22] Alternatively, the Cpx signaling pathway is regulated by the negative feedback loop in the cells as the Cpx-regulated protein, CpxP, inhibits the Cpx signal via the periplasmic domain of CpxA.[23] Tethering CpxP to the inner membrane using an MBP tag constitutively blocks the Cpx signaling, suggesting that inhibition occurs at or near the inner membrane surface where CpxA is localized.[16] An *in vitro* reconstituted system containing recombinant CpxA, CpxP, and CpxR showed that CpxP inhibits the autophosphorylation activity of CpxA in the absence of other periplasmic proteins.[24] Recent structural and functional studies demonstrated the direct interaction between CpxP and the periplasmic domain of CpxA.[25,26] Therefore, it is proposed that CpxP represses the Cpx pathway, especially the autophosphorylation of CpxA, by directly interacting with the periplasmic domain of CpxA. The feedback loop of Cpx signaling is completed by the DegP-mediated degradation of CpxP and the activation of CpxA.[27] CpxP also plays a role in protein quality control by delivering the misfolded PapE or PapG to DegP.[28]

Structural studies of CpxA and CpxP were conducted to gain the structural implication of signal recognition of CpxA and its interaction with CpxP. In this study, we elucidated the crystal structure of the periplasmic domain of *Vibrio parahaemolyticus* CpxA (VpCpxA-peri) and tested its interaction with VpCpxP. *V. parahaemolyticus*, a Gram-negative halophilic bacterium causing gastroenteritis,[29] which belongs to the same class as

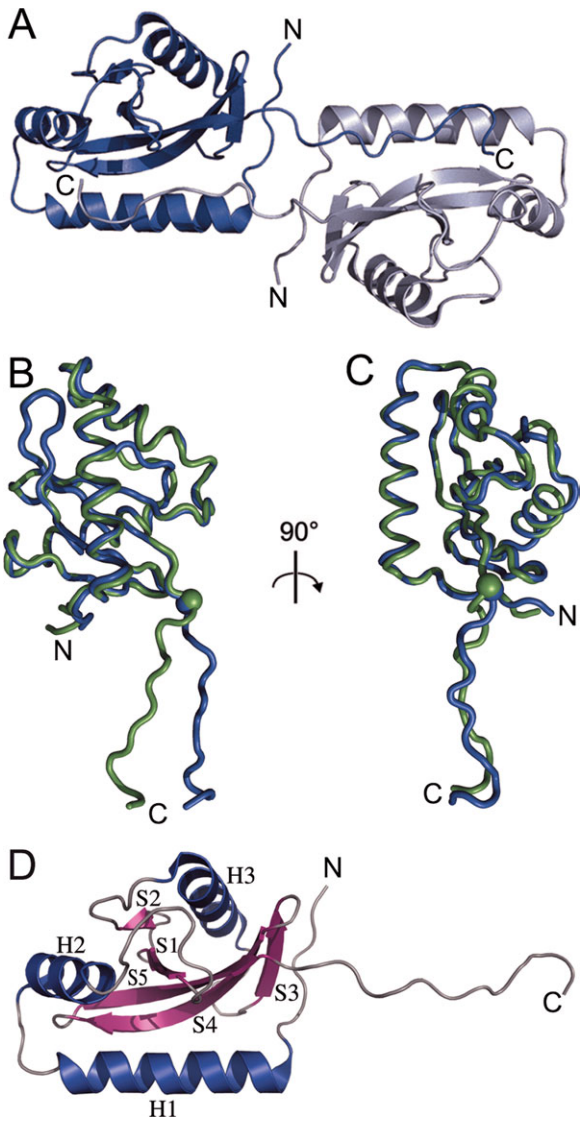

Figure 1. The overall structure of VpCpxA-peri. (A) Ribbon diagrams of Chain A and Chain B in an asymmetric unit are shown in blue and grey, respectively. (B) and (C) The Cα traces of Chain B of VpCpxA-peri were superimposed on Chain B in two different views. Chain A and Chain B are represented in blue and green, respectively. The Cα atom of Asn161 is drawn as a sphere. N and C represent the N- and C-termini of VpCpxA-peri, respectively. (D) The secondary structure elements of Chain A of VpCpxA-peri are indicated. α-helices (H1–H3), β-strands (S1–S5), and loops are colored blue, violet, and grey, respectively. An interactive view is available in the electronic version of the article.

E. coli, a *γ-proteobacteria*, harbors the *cpxa*, *cpxr*, and *cpxp* genes in the Cpx two-component system. Although the detailed Cpx pathways of bacteria belonging to the *Vibrio* genes have not been studied intensively, it is expected that the overall regulation mechanisms of the Cpx pathways must be similar to those in *E. coli*.

From the crystal structure analysis, we found that the sensory domain of CpxA contains a Per-ARNT-SIM (PAS) fold, which is not predicted from the sequence alignment. On the basis of NMR

Table I. *Data Collection and Refinement Statistics*

MAD data collection statistics (PAL BL 4A)					
	Native	Peak	Edge	Remote 1	Remote 2
Space group			*R3*		
Unit cell (Å, °)		$a = 148.58$, $b = 148.58$, $c = 32.31$, $\alpha = 90$, $\beta = 90$, $\gamma = 120$			
Resolution (Å)	30.0–2.1		30.0–2.3		
Wavelength (Å)	0.97943	0.97943	0.97970	0.98751	0.97176
Total/unique reflections	65658/29204	67638/11801	67729/11812	67852/11830	67810/11823
Completeness (%)	94.0 (80.5)[a]	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
$I/\acute{o}(I)$	18.1 (2.8)	23.0 (4.0)	22.2 (3.4)	22.9 (3.6)	20.4 (3.0)
R_{sym} (%)	10.0 (34.4)	11.5 (36.7)	10.7 (40.3)	10.0 (40.2)	11.4 (48.4)
Refinement statistics					
Resolution (Å)	28.9–2.1				
No. of reflections	15389				
R_{work}/R_{free}	22.0/26.5				
No. of atoms protein/water	2191/128				
r.m.s. deviations bond lengths (Å)/bond angles (°)	0.008/1.348				
Ramachandran statistics(%)[b]	85.0/15.0/0.0				

[a] The highest resolution shell is shown in parenthesis.
[b] The fraction of residues in the favored, allowed, and disallowed regions of the Ramachandran diagram calculated by PROCHECK.[30]

chemical shift perturbation analysis, we also confirmed that the PAS domain of CpxA does not directly interact with CpxP. Our results will provide a foundation for additional biochemical and genetic studies aimed at dissecting the mechanism of ligand sensing by CpxA and inhibition by CpxP.

Results

The overall structure of VpCpxA-peri

The crystal of the periplasmic domain of *V. parahaemolyticus* CpxA (VpCpxA-peri) belongs to the *R3* space group and diffracted to a 2.1 Å resolution. The structure was solved by the multi-wavelength anomalous dispersion (MAD) method using selenomethionine-derived VpCpxA-peri crystals. There were two molecules in the asymmetric unit, which are referred to as Chain A and Chain B (Fig. 1). The amino acid residues 41–172 were assigned to Chain A, whereas the amino acid residues 42–172 were assigned to Chain B. The R_{work} and R_{free} of the final model were 22.0% and 26.5%, respectively. Table I summarizes the data collection and refinement statistics. Chain A is used for the structural description in this study unless otherwise indicated.

The structure of VpCpxA-peri has a globular domain with a protruded C-terminal tail (Fig. 1). The five-stranded β-sheet (S1-S5) surrounded by three α-helices (H1-H3) composes the globular domain [Fig. 1(D)]. Two α-helices (H1 and H2) are placed at the N-terminus, which are followed by two β-strands (S1 and S2). S2 is at the end of the β-sheet and connected to S3 via an α-helix, H3. The C-terminal tail may be more flexible than the globular domain [Fig. 1(B and C)]; when the Cα atoms of Chain A are superimposed on those of Chain B, the root mean square deviation (RMSD) in the presence of C-terminal tail is 2.1 Å, whereas the RMSD in the absence of the C-terminal tail (amino acids 162–172) is 0.8 Å.

VpCpxA-peri possesses a PAS domain

The periplasmic sensory domain of VpCpxA has no significant sequence homology with known proteins, whereas the remaining region is comprised of three conserved domains found in other two-component HKs: the HAMP domain, the ATPase domain, and the phosphoacceptor domain. However, VpCpxA-peri matches many PAS domains in various sensory HKs when the structural homologues were searched against the structure in Protein Data Bank. Five structures with z scores greater than 6.0 were selected for structural comparison to VpCpxA-peri (Table II and Fig. 2): the AbfS arabinofuranosidase

Table II. *The Structural Homologues of VpCpxA-Peri*

PDB ID (chain)	Name	Z score	RMSD	Amino acids	Sequence identity (%)
2VA0 (B)	AbfS	7.0	4.3	97[a]/119[b]	10
2QHK (A)	MCP	6.6	2.8	99/148	6
3C38 (A)	LuxQ	6.3	4.3	109/217	10
3FOS (B)	KinD	6.3	3.7	108/202	8
1ZHH (B)	LuxQ	6.3	4.6	107/208	7

[a] Number of amino acid residues that are used for the RMSD calculation.
[b] Number of amino acid residues present in the homologous structures.

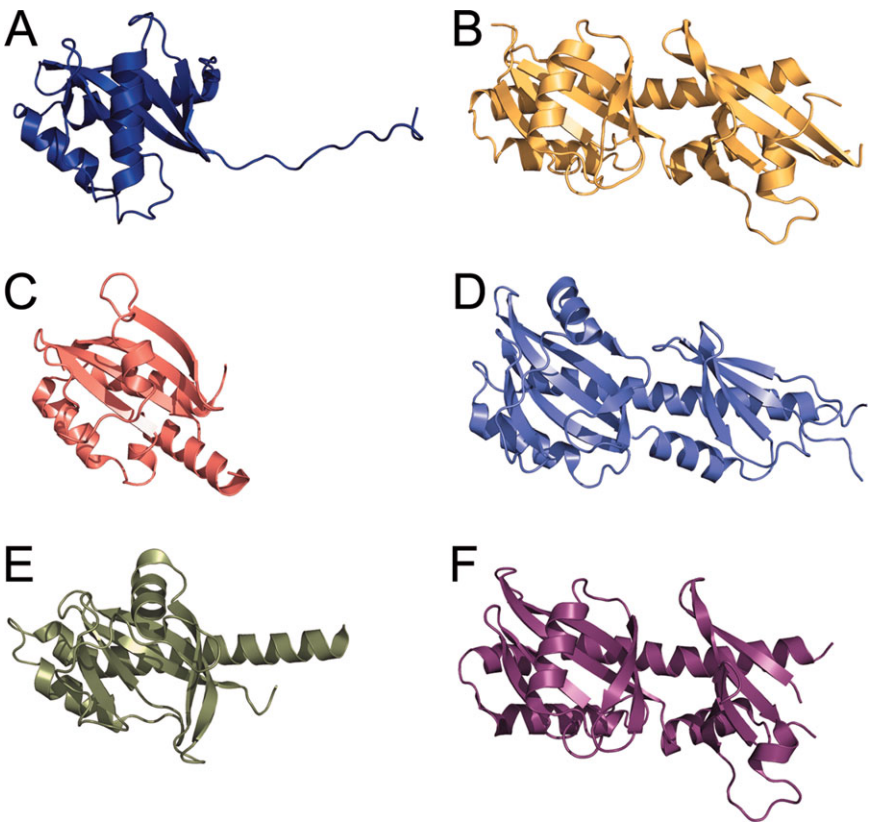

Figure 2. Structural similarity between the periplasmic domain of VpCpxA-peri and other HKs containing PAS domains in their periplasmic regions. VpCpxA-peri (A) is compared with the LuxQ autoinducer binding protein of *V. cholerae* (B) (3C38), AbfS arabinofuranosidase two-component sensor protein of *Cellvibrio japonicus* (C) (2VA0), KinD sporulation kinase D of *Bacillus subtilis* (D) (3FOS), methyl-accepting chemotaxis protein of *V. parahaemolyticus* (E) (2QHK), and LuxQ autoinducer 2-binding periplasmic protein of *Vibrio harveyi* (F) (1ZHH) in ribbon diagrams. The orientations of (B–F) were determined by overlapping (B–F) with (A). N- and C-termini were indicated by N and C, respectively.

two-component sensor protein of *Cellvibrio japonicus* (PDB ID: 2VA0),[31] the methyl-accepting chemotaxis protein of *V. parahaemolyticus* (PDB ID: 2QHK), the KinD sporulation kinase D of *Bacillus subtilis* (PDB ID: 3FOS), the LuxQ autoinducer binding proteins of *Vibrio cholerae* (PDB ID: 3C38) and *Vibrio harveyi* (PDB ID: 1ZHH).[32] Therefore, the fold of the globular domain of VpCpxA-peri can be classified into a classical PAS domain fold, which is found in many kinds of sensory proteins across the three kingdoms, such as serine/threonine kinases, HKs, photo receptors, and circadian clock proteins, and is involved in sensing of signals such as oxygen, redox molecules, light, or chemical compounds.[33] Every PAS domain shows a low sequence identity (less than 10%) to the sequence of the VpCpxA PAS domain (Table II), so this PAS domain was not predicted from sequence-based homology search.

The central β-pleated sheet consisting of five antiparallel β-strands which is the most conserved structural feature in the PAS domain is also found in VpCpxA-peri. In the classical PAS domain, the second and third β-strands are connected by a long loop or additional helices. In the crystal structure of VpCpxA-peri, the helix H3 is found between two strands, S2 and S3. Although the PAS domains share the overall fold, they must harbor structural differences to accommodate various ligands.[34] Accordingly, the conformational diversity is observed in the N-terminus, the C-terminus, and the loop or helix between the second and third strands (Fig. 2). In addition, the structural elements and their arrangement are also quite different in each PAS domain (Fig. 2). The structure of AbfS is most similar to the VpCpxA-peri structure, except for the additional small helix between H2 and S1 and the orientation of H1 [Fig. 2(C)]. The methyl-accepting chemotaxis protein has an additional α-helix and the long H1 [Fig. 2(E)]. KinD and LuxQ also have a long H1 helix that extends to the second β-sheet and an additional PAS domain [Fig. 2(B, D, F)].

Surface charge distribution

The periplasmic domain of VpCpxA has 14 arginine and 8 lysine residues out of a total of 136 residues, and thus it can be categorized as a basic polypeptide because its Arg+Lys frequency (16.2%) is higher than the average value (11.4%). When the electrostatic surface potential of the VpCpxA-peri monomer was visualized (Fig. 3), most of the basic residues are

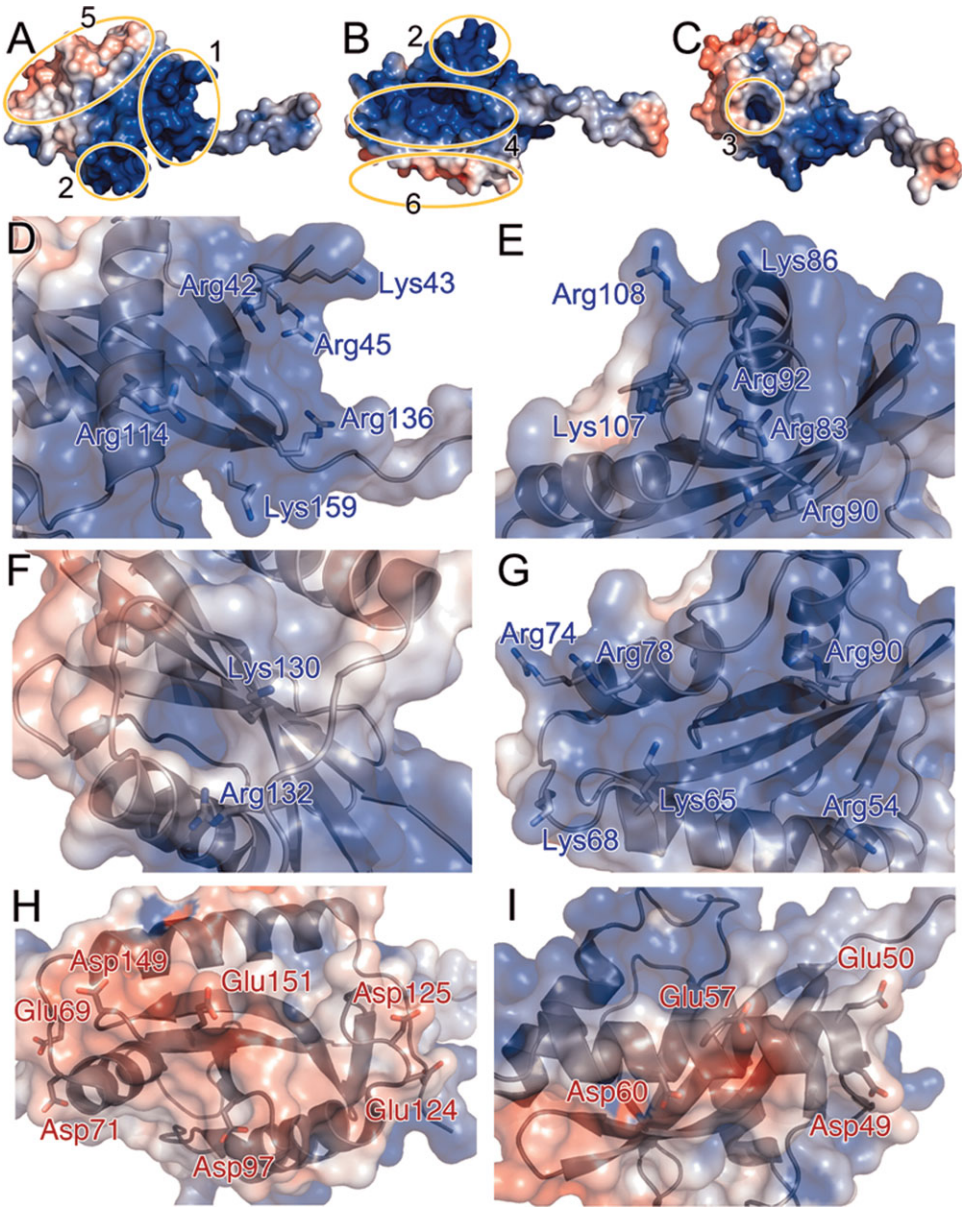

Figure 3. The surface charge distribution of VpCpxA-peri. (A-C) The charge of VpCpxA-peri is calculated within a range of −5 kT (red) to +5 kT (blue) and displayed on the surface diagram of VpCpxA-peri. The basic juxtamembrane region, the basic bulge, the basic hole, the basic side, the large acidic surface, and the small acidic surface are depicted with numbers 1–6, respectively. The side chains of residues constituting each basic or acidic region are drawn as a stick diagram: (D) the basic juxtamembrane region, (E) the basic bulge, (F) the basic hole, (G) the basic side, (H) the large acidic surface, and (I) the small acidic surface.

distributed over four regions: a basic juxtamembrane region, a basic bulge, a basic side, and a basic hole [Fig. 3(A–C)]. Arg42, Lys43, and Arg45 in the N-terminal loop, Arg114 on H3, Arg136 on S4, and Lys159 on S5 constitute the basic juxtamembrane region [Fig. 3(D)]. The basic bulge consists of residues on S1 and loops nearby: Arg83, Lys86, and Arg90 on $L_{H2\text{-}S1}$, Arg92 on S1, and Lys107 and Arg108 on $L_{S2\text{-}H3}$ [Fig. 3(E)]. The basic side is distributed over H1 (Arg54 and Lys65), $L_{H1\text{-}H2}$ (Lys68), H2 (Arg74 and Arg78), and $L_{H2\text{-}S1}$ (Arg90) [Fig. 3(G)]. In addition to these basic surfaces, Lys130 and Arg132 on S3 form a basic hole on a rather hydrophobic surface [Fig. 3(C and F)]. Arg59 and Lys128 do not belong to these basic regions, whereas Arg169 is placed at the C-terminus. VpCpxA-peri also has two acidic regions [Fig. 3(A and B)]. Seven acidic residues, Glu69, Asp71, Asp97, Asp125, Glu124, Asp149, and Glu151, form a large acidic region [Fig. 3(H)]. Four acidic residues on H1 constitute the other small acidic region: Asp49, Glu50, Glu57, and Asp60 [Fig. 3(I)].

The interaction between the periplasmic domains of VpCpxA and VpCpxP

It has been suggested that CpxP represses the Cpx signal by direct association with the periplasmic

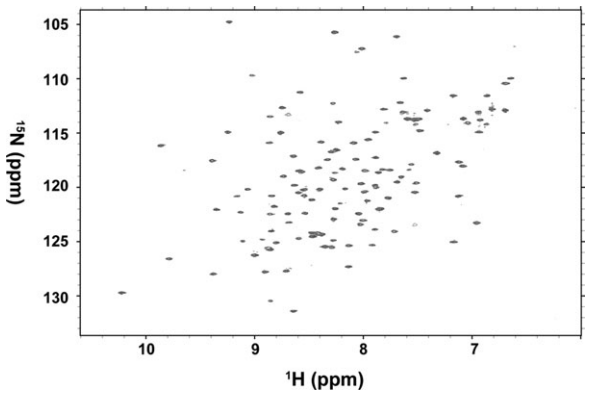

Figure 4. The ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled VpCpxA-ΔCT in the presence and absence of VpCpxP. The ^{1}H-15N HSQC spectrum of ^{15}N-labeled VpCpxA-ΔCT in the presence of a five-fold molar excess of VpCpxP is superimposed with the spectrum of VpCpxA-ΔCT alone, but they are indistinguishable.

domain of CpxA in *E. coli*.[25,26] To confirm the binding between VpCpxA and VpCPxP and investigate the residues involved in the CpxP binding, an NMR chemical shift perturbation assay was performed by titrating VpCpxA-peri with VpCpxP. Because VpCpxA-peri was not suitable for the NMR experiment due to poor spectral resolution of the 2D ^{15}N HSQC spectrum, VpCpxA-peri lacking the C-terminal 11 residues, VpCpxA-ΔCT (amino acid residues 38–162), was introduced. The ^{15}N HSQC spectra of VpCpxA-ΔCT were measured in the absence or presence of a five-fold molar excess of VpCpxP, and the two spectra were overlapped for comparison (Fig. 4). The peaks in both spectra are identical, which means VpCpxP does not interact with the VpCpxA periplasmic domain at least in the experimental conditions we used.

Discussion

The sequence identities of the full-length CpxA proteins belonging to γ-proteobacteria vary from 23% to 44%, whereas those of the periplasmic domains of CpxA proteins are in the range of 8%–17% except the sequence identity between two Vibrio CpxA proteins (Table III). Compared with the periplasmic domain, the remaining regions, two transmembrane helices and the cytoplasmic domains, show higher sequence homology, and particularly, all the essential motifs found in HK are well conserved in the cytoplasmic domain. Thereby, it is expected that the Cpx signal is transduced in the similar way as other known two-component system in spite of low sequence identities among the CpxA proteins belonging to the different genera. Although the periplasmic domains of the CpxA proteins are not homologous in their sequences, they seem to have similar three-dimensional topology because the predicted secondary structures are well matched with those of VpCpxA-peri (Fig. 5). This may reflect that the inducing signal recognized by the sensing domain of CpxA is diverse in each species, whereas the signaling mechanisms through the PAS, transmembrane, and cytoplasmic domain are similar.

Some stimuli of the *E. coli* Cpx system are shared with those of the *Vibrio* Cpx system such as copper,[35] but not all them. Based on the stimulation of the *Vibrio* Cpx pathway by chloride and magnesium ions, it was proposed that the *Vibrio* Cpx pathway have a role in responding to salinity.[35] *V. parahaemolyticus* and *V. cholerae* belong to the same genus and share a similar habitat. The high sequence homology between the CpxA periplasmic domains of these two species implies that stimuli for VpCpxA are similar to triggers for *V. cholerae* CpxA.

When the predicted secondary structure of VpCpxA-peri is compared with that of the crystal structure, the α-helices are in good agreement with each other, except the C-terminus that is connected to the second transmembrane helix (Fig. 5). This α-helix is also predicted to be present in the same position in other CpxA proteins despite the low sequence identities in the periplasmic domain (Fig. 5 and Table III). In addition, residues 169–173 of Chain A form a 3_{10} helix. Based on these results, it is tempting to propose that the C-terminal end of the periplasmic domain forms a helix that continues to the second

Table III. *The Sequence Identities Among CpxA Proteins*

		CpxA of γ-proteobacteria[a]				
		Vp	So	Lp	Ps	Vc
CpxA of γ-proteobacteria	Ec	44[b] (13[c])	35 (16)	27 (11)	25 (8)	44 (15)
	Vp	—	33 (17)	25 (9)	23 (7)	68 (48)
	So	—	—	32 (10)	27 (16)	33 (13)
	Lp	—	—	—	25 (8)	25 (10)
	Ps	—	—	—	—	24 (8)

[a] Ec, *Escherichia coli* str. K-12 sub. W3110 (geneID: 7438256); Vp, *V. parahaemolyticus* RIMD 2210633 (geneID: 1190422); So, *Shewanella oneidensis* MR-1 (geneID: 1172073); Lp, *Legionella pneumophila* str. Lens (geneID: 3114153); Ps, *Pseudomonas syringae* pv. tomato str. DC3000 (geneID: 1183444); Vc, *V. cholerae* O1 biovar El Tor str. N16961 (geneID: 2615521).
[b] Sequence identity among the CpxA proteins (%).
[c] Sequence identity of the periplasmic domain of CpxA proteins (%). Residues corresponding to 38–173 of VpCpxA-peri in Fig. 5 were used for the calculation of the sequence identity.

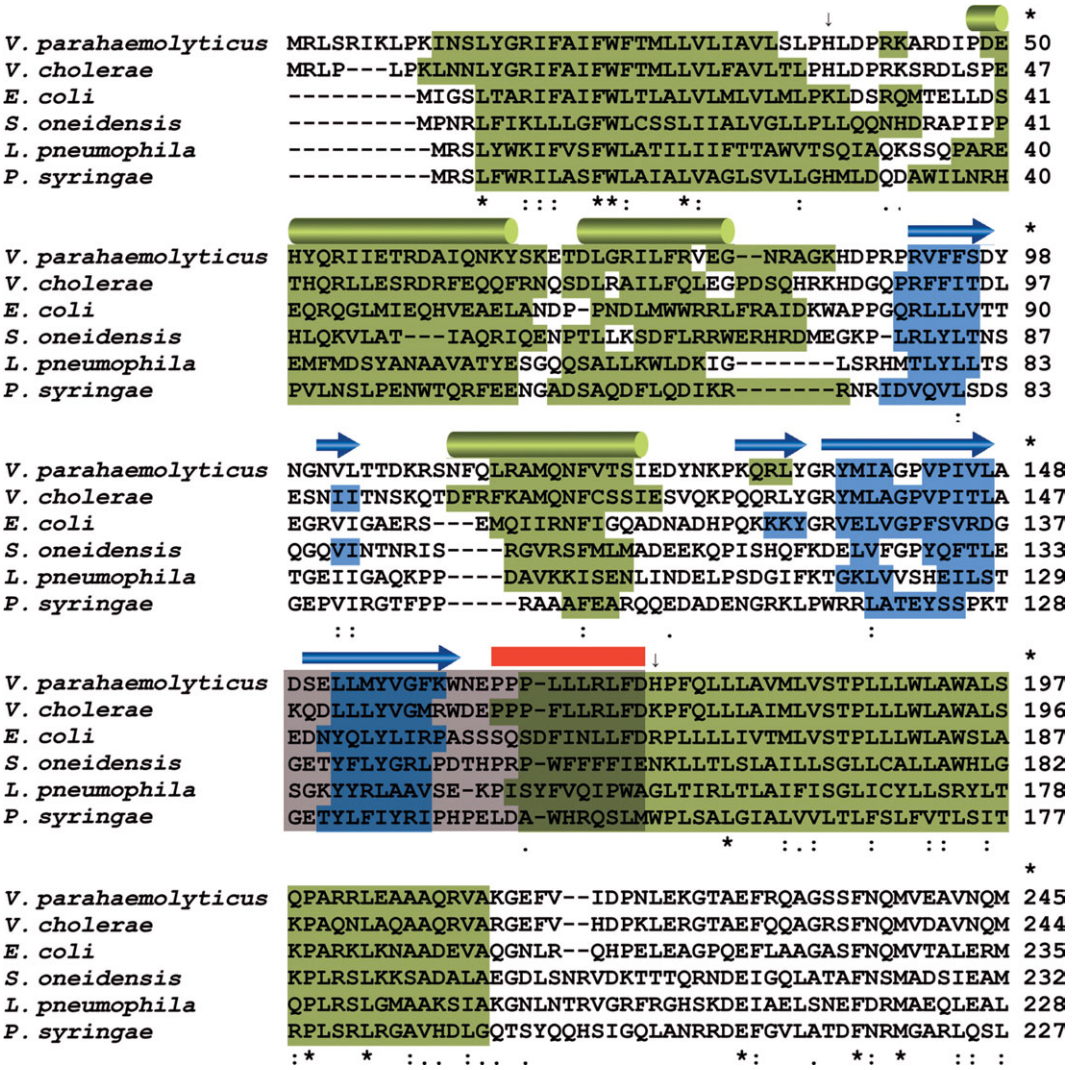

Figure 5. Multiple sequence alignment of CpxA proteins of γ-proteobacteria. Multiple sequence alignment of the CpxA proteins of *V. parahaemolyticus*, *V. cholerae*, *E. coli*, *Shevanella oneidensis*, *Legionella pneumophila*, and *Pseudomonas syringae*. Predicted α-helices and β-strands are highlighted on the sequence of each species in green and blue, respectively. α-helices (H1–H3) and β-strands (S1–S5) in Chain A of VpCpxA-peri are displayed in the first row as green cylinders and blue arrows, respectively. The C-terminal tail of VpCpxA-peri deleted for the NMR experiment is indicated by a red line, and the region predicted to play a role in *E. coli* CpxP binding is represented by a grey box.[25] The N- and the C-terminal residues of VpCpxA-peri are indicated by arrows.

transmembrane helix but is unfolded due to interruption of the helix in the crystal structure.

The PAS domain of VpCpxA-peri has low sequence homology to other PAS domains (Table II). In addition, each PAS domain uses a different region to recognize its cognate ligand. Neither the ligand nor a ligand-binding site was observed in the crystal structure of the periplasmic domain of VpCpxA, and thus it is difficult to identify the ligand-binding region in the structure of VpCpxA-peri. However, assuming that VpCpxA also recognizes Cl^- ions like *V. cholera* CpxA, one of the basic surfaces and especially the highly basic side between H1 and H2 may be a good candidate for the ligand recognition site [Fig. 3(B)]. If the basic side is accommodated by negatively charged Cl^- ions, then the flat surface can be used for dimerization, which is a requirement for the autophosphorylation of HK in the two-component system. This possibility should be addressed in a further study.

Recently, two groups independently reported the crystal structure of *E. coli* CpxP.[25,26] It shows the same fold as its homologue, *E. coli* Spy, which contains a basic concave surface.[36] Zhou et al. suggested that *E. coli* CpxA likely binds CpxP via its C-terminus (amino acid residues 138–162) that corresponds to S5 of the PAS domain (residues 149–160) and the protruded C-terminal tail (residues 161–172) of VpCpxA-peri (Fig. 5).[25] Assuming that the *E. coli* CpxA has the same structure as VpCpxA-peri, it is unlikely that the residues in S5 are involved in CpxP binding because they are buried in the PAS domain and form

a secondary structure with other strands (Figs. 1 and 5). Therefore, the reported interaction between the *E. coli* CpxP and CpxA are not likely biologically relevant because the synthesized peptide may not reflect the structural property of the native protein. Alternatively, current results can be interpreted as that the binding between CpxP and CpxA is mediated not by the S5 region but by the C-terminal end of the periplasmic domain of CpxA (residues 163–172) because the CpxP binding peptide identified by Zhou et al. harbors the C-terminal end[25] and the binding of VpCpxP to VpCpxA-ΔCT was not detected via NMR. Another possible scenario is that the lack of a helical structure on the C-terminal end of VpCpxP hinders binding to CpxP or the membrane lipids may mediate the binding between CpxP and CpxA because the proposed binding site is near the membrane. From the binding assay between *E. coli* CpxP and CpxA, it was proposed that their binding was mediated by the charge-charge interaction involving the positively charged patch of CpxA (Arg56, Arg60, and Arg67).[25] These residues are well conserved as Arg67, Lys73, and Arg80 in VpCpxP, respectively. Consistently, the negatively charged residues of *E. coli* CpxA in the proposed CpxP binding site are also found in VpCpxA (Fig. 5). Therefore, the CpxA-CpxP interaction may be similar in *E. coli* and *V. parahaemolyticus*.

VpCpxA-peri exists as a dimer in an asymmetric unit [Fig. 1(A)]; however, its biological relevance is not confirmed. In the crystal structure, the C-terminal tail of one subunit is protruded to and grabbed by the other subunit [Fig. 1(A)]. Both the N-terminus and C-terminus of VpCpxA-peri have to be connected to transmembrane helices, and hence the dimer structure seems to be improper as a membrane-spanning protein. These results suggest that the dimer formation is likely caused by the crystal packing or the intervention of the native structure in the recombinant protein. Many PAS domains of the sensory HKs, such as PhoQ,[37] LuxQ,[32] and FixL,[34] form a dimer with diverse dimeric interfaces. Furthermore, PAS domains can exist as several dimeric forms in the case of FixL from *Bradyrhizobium japonicum*[34] and *Sinorhizobium meliloti*.[38] Although there is no report that the CpxA periplasmic domain exists as a dimer, VpCpxA-peri is expected to form a dimer by ligand binding or the intact VpCpxA can exist as a dimer. However, predicting the dimeric interface of VpCpxA-peri from this study is not plausible because of its flexibility.

In summary, we solved the crystal structure of the periplasmic domain of *V. parahaemolyticus* CpxA. Although not predicted based on sequence, VpCpxA-peri contains a PAS domain that could play a role in sensing envelope stress signals. However, no direct clue necessary for identifying the ligand of CpxA was provided. We also examined the interaction between the periplasmic domains of CpxA and CpxP of *V. parahaemolyticus* using NMR spectroscopy, but their binding was not observed under the current experimental conditions. Our work that defined the structure of the sensor domain of CpxA may help future biochemical and genetic studies in the identification of natural ligands and the elucidation of its signaling and regulation mechanisms.

Material and Methods

Cloning and purification

The N-terminal signal sequence of CpxP and the transmembrane regions of CpxA were predicted using SignalP[39] and TMHMM,[40] respectively. The mature form of CpxP and the periplasmic domain of CpxA of *V. parahaemolyticus* were designed based on the following domain boundaries: VpCpxP amino acid residues 15–170, VpCpxA-peri amino acid residues 38–173, and VpCpxAΔCT amino acid residues 38–162. Each gene was amplified by polymerase chain reaction (PCR) from genomic DNA and cloned into the modified pET-28a(+) or pDUET vector that has a 6xHis-tag and tobacco etch virus (TEV) protease cleavage site. Every protein was expressed using *Escherichia coli* BL21(DE3), except selenomethionine-labeled VpCpxA-peri, which was expressed using a *Escherichia coli* B834(DE3) cell. A single colony of transformed cells was inoculated into 50 ml lysogeny broth (LB) with 50 μg/ml ampicilin or kanamycin at 37°C overnight. Cultured cells were transferred into 500 ml LB for the unlabeled protein, 600 ml M9 minimal medium with 1 g/L of $^{15}NH_4Cl$ for ^{15}N-labeled protein or 600 ml M9 minimal medium with 50 mg/L of selenomethionine. Cells were cultured at 37°C to an optical density 0.5–0.6 at 600 nm. The recombinant proteins were expressed at 18°C in the presence 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 12 hours.

The cells were harvested by centrifugation at 3500 g at 4°C for 15 min, resuspended in buffer A (20 mM Tris-HCl pH 8.5) and disturbed by sonication. Remaining cells and aggregates were pelleted by centrifugation at 16,000 g at 4°C for 1 hour. Each protein was purified by the same method: the protein was loaded into a HisTrap column (GE Healthcare), washed with 40 mM imidazole, and subjected to a 40–500 mM imidazole gradient using FPLC (GE Healthcare). After the eluted protein was pooled, the sample was dialyzed to buffer A and treated with TEV protease to cut the N-terminal 6xHis-tag. After the cleaved tag was removed by the HisTrap column, the sample was purified by size-exclusion chromatography with a Superdex75 column (GE Healthcare). The purified VpCpxA-peri was concentrated to 10 mg/ml for crystallization purpose. The purity of VpCpxA-peri checked by SDS-gel electrophoresis was more than 95%.

Structure determination and analysis

The crystallization of selenomethinone-labeled VpCpxA-peri was performed using the microbatch method at 22°C. One microliter of protein solution at a concentration of 10 mg/ml was mixed with an equal volume of crystallization reagent containing 0.1 M Tris-HCl pH 7.0, 18–24% PEG 3000, 0.2 M $Ca(OAc)_2$, 3% (v/v) 1,6-hexanediol, and 2 mM Tris (2-carboxyethyl) phosphine under a layer of Al's oil (Hampton). The crystals were flash-frozen in a cold nitrogen stream. The X-ray diffraction data were collected at a resolution of 2.3 Å with an ADSC Quantum 210 CCD detector at beam line 4A, Pohang Light Source, South Korea. The data was processed and integrated using HKL2000 and scaled using SCALEPACK.[41] The VpCpxA-peri structure was determined by MAD using SOLVE/RESOLVE.[42,43] The model building was performed by Coot.[44] Several cycles of rigid-body refinement, positional refinement, restrained refinement, simulated annealing, B-factor refinement, and model rebuilding were performed at 2.3 Å resolution using Phenix,[45] Refmac,[46] CNS,[47] and Coot[44] programs.

The quality of the structure was analyzed by PROCHECK.[30] The RMSD values of the overlapped structure were calculated using LSQKAB.[48] A Dali-Lite v. 3 server[49] was used for searching the structural homologues in the Protein Data Bank.[50] Figures were drawn using PyMOL,[51] and the surface charge was calculated and visualized by PDB2PQR[52] and APBS,[53] respectively. The domain structure was searched using the Pfam database.[54] The multiple sequence alignment was implicated using ClustalW2.[55]

Interaction test

^{15}N HSQC spectra were recorded at 20°C on a Bruker Avance 800 MHz spectrometer outfitted with cryogenic probes, and the data were processed with nmrPipe.[56] Isotopically labeled VpCpxAΔCT was made from M9 minimal media containing ^{15}N NH_4Cl as the sole nitrogen source. The sample buffer for NMR was 50 mM HEPES pH 7.5, 0.2 M NaCl, 5% glycerol, 50 mM glutamate, and 50 mM arginine. HSQC spectra were recorded on VpCpxAΔCT at a final concentration of 60 μM in the absence or presence of VpCpxP.

Acknowledgment

The authors thank laboratory members John and Carol Gross for helpful discussions and the staff of the PAL BL 4A beamline for their assistance during data collection.

References

1. Dorel C, Lejeune P, Rodrigue A (2006) The Cpx system of *Escherichia coli*, a strategic signaling pathway for confronting adverse conditions and for settling biofilm communities? Res Microbiol 157:306–314.
2. Raivio TL (2005) Envelope stress responses and Gram-negative bacterial pathogenesis. Mol Microbiol 56: 1119–1128.
3. Raivio TL, Silhavy TJ (1999) The sigmaE and Cpx regulatory pathways: overlapping but distinct envelope stress responses. Curr Opin Microbiol 2:159–165.
4. Raivio TL, Silhavy TJ (2001) Periplasmic stress and ECF sigma factors. Annu Rev Microbiol 55:591–624.
5. Rowley G, Spector M, Kormanec J, Roberts M (2006) Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens. Nat Rev Microbiol 4: 383–394.
6. Ruiz N, Silhavy TJ (2005) Sensing external stress: watchdogs of the *Escherichia coli* cell envelope. Curr Opin Microbiol 8:122–126.
7. Wick LM, Egli T (2004) Molecular components of physiological stress responses in *Escherichia coli*. Adv Biochem Eng Biotechnol 89:1–45.
8. Chang C, Stewart RC (1998) The two-component system. Regulation of diverse signaling pathways in prokaryotes and eukaryotes. Plant Physiol 117:723–731.
9. Weber RF, Silverman PM (1988) The cpx proteins of *Escherichia coli* K12. Structure of the cpxA polypeptide as an inner membrane component. J Mol Biol 203: 467–478.
10. Dong J, Iuchi S, Kwan HS, Lu Z, Lin EC (1993) The deduced amino-acid sequence of the cloned cpxR gene suggests the protein is the cognate regulator for the membrane sensor, CpxA, in a two-component signal transduction system of *Escherichia coli*. Gene 136: 227–230.
11. Snyder WB, Davis LJ, Danese PN, Cosma CL, Silhavy TJ (1995) Overproduction of NlpE, a new outer membrane lipoprotein, suppresses the toxicity of periplasmic LacZ by activation of the Cpx signal transduction pathway. J Bacteriol 177:4216–4223.
12. Otto K, Silhavy TJ (2002) Surface sensing and adhesion of *Escherichia coli* controlled by the Cpx-signaling pathway. Proc Natl Acad Sci USA 99:2287–2292.
13. Miyadai H, Tanaka-Masuda K, Matsuyama S, Tokuda H (2004) Effects of lipoprotein overproduction on the induction of DegP (HtrA) involved in quality control in the *Escherichia coli* periplasm. J Biol Chem 279: 39807–39813.
14. DiGiuseppe PA, Silhavy TJ (2003) Signal detection and target gene induction by the CpxRA two-component system. J Bacteriol 185:2432–2440.
15. Lee YM, DiGiuseppe PA, Silhavy TJ, Hultgren SJ (2004) P pilus assembly motif necessary for activation of the CpxRA pathway by PapE in *Escherichia coli*. J Bacterial 186:4326–4337.
16. Danese PN, Silhavy TJ (1998) CpxP, a stress-combative member of the Cpx regulon. J Bacteriol 180:831–839.
17. Kershaw CJ, Brown NL, Constantinidou C, Patel MD, Hobman JL (2005) The expression profile of *Escherichia coli* K-12 in response to minimal, optimal and excess copper concentrations. Microbiology 151:1187–1198.
18. Yamamoto K, Ishihama A (2005) Transcriptional response of *Escherichia coli* to external copper. Mol Microbiol 56:215–227.
19. Raffa RG, Raivio TL (2002) A third envelope stress signal transduction pathway in *Escherichia coli*. Mol Microbiol 45:1599–1611.
20. Kohanski MA, Dwyer DJ, Wierzbowski J, Cottarel G, Collins JJ (2008) Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death. Cell 135:679–690.
21. De Wulf P, McGuire AM, Liu X, Lin EC (2002) Genome-wide profiling of promoter recognition by the

two-component response regulator CpxR-P in *Escherichia coli*. J Biol Chem 277:26652–26661.
22. Price NL, Raivio TL (2009) Characterization of the Cpx regulon in *Escherichia coli* strain MC4100. J Bacteriol 191:1798–1815.
23. Raivio TL, Popkin DL, Silhavy TJ (1999) The Cpx envelope stress response is controlled by amplification and feedback inhibition. J Bacteriol 181:5263–5272.
24. Fleischer R, Heermann R, Jung K, Hunke S (2007) Purification, reconstitution, and characterization of the CpxRAP envelope stress system of *Escherichia coli*. J Biol Chem 282:8583–8593.
25. Zhou X, Keller R, Volkmer R, Krauss N, Scheerer P, Hunke S (2011) Structural basis for two-component system inhibition and pilus sensing by the auxiliary CpxP protein. J Biol Chem 286:9805–9814.
26. Thede GL, Arthur DC, Edwards RA, Buelow DR, Wong JL, Raivio TL, Glover JN (2011) Structure of the periplasmic stress response protein CpxP. J Bacteriol 193: 2149–2157.
27. Buelow DR, Raivio TL (2005) Cpx signal transduction is influenced by a conserved N-terminal domain in the novel inhibitor CpxP and the periplasmic protease DegP. J Bacteriol 187:6622–6630.
28. Isaac DD, Pinkner JS, Hultgren SJ, Silhavy TJ (2005) The extracytoplasmic adaptor protein CpxP is degraded with substrate by DegP. Proc Natl Acad Sci USA 102: 17775–17779.
29. Sakazaki R, Iwanami S, Fukumi H (1963) Studies on the enteropathogenic, facultatively halophilic bacteria, *Vibrio parahaemolyticus*. I. Morphological, cultural and biochemical properties and its taxonomical position. Jpn J Med Sci Biol 16:161–188.
30. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26:9.
31. Emami K, Topakas E, Nagy T, Henshaw J, Jackson KA, Nelson KE, Mongodin EF, Murray JW, Lewis RJ, Gilbert HJ (2009) Regulation of the xylan-degrading apparatus of *Cellvibrio japonicus* by a novel two-component system. J Biol Chem 284:1086–1096.
32. Neiditch MB, Federle MJ, Miller ST, Bassler BL, Hughson FM (2005). Regulation of LuxPQ receptor activity by the quorum-sensing signal autoinducer-2. Mol Cell 18:507–518.
33. Ponting CP, Aravind L (1997) PAS: a multifunctional domain family comes to light. Curr Biol 7:R674–R677.
34. Ayers RA, Moffat K (2008) Changes in quaternary structure in the signaling mechanisms of PAS domains. Biochemistry 47:12078–12086.
35. Slamti L, Waldor MK (2009) Genetic analysis of activation of the *Vibrio cholerae* Cpx pathway. J Bacteriol 191:5044–5056.
36. Kwon E, Kim DY, Gross CA, Gross JD, Kim KK (2010) The crystal structure *Escherichia coli* Spy. Protein Sci 19:2252–2259.
37. Cheung J, Bingman CA, Reyngold M, Hendrickson WA, Waldburger CD (2008) Crystal structure of a functional dimer of the PhoQ sensor domain. J Biol Chem 283: 13762–13770.
38. Miyatake H, Mukai M, Park SY, Adachi S, Tamura K, Nakamura H, Nakamura K, Tsuchiya T, Iizuka T, Shiro Y (2000) Sensory mechanism of oxygen sensor FixL from *Rhizobium meliloti*: crystallographic, mutagenesis and resonance Raman spectroscopic studies. J Mol Biol 301:415–431.
39. Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2:953–971.
40. Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567–580.
41. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 276:20.
42. Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. Acta Cryst D55:849–861.
43. Terwilliger TC (2003) Automated main-chain model building by template matching and iterative fragment extension. Acta Cryst D59:38–44.
44. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Cryst D60:2126–2132.
45. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. J Appl Cryst 40:658–674.
46. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Cryst D53:240–255.
47. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Cryst D54:905–921.
48. Kabsch W (1976) A solution for the best rotation to relate two sets of vectors. Acta Cryst A32:2.
49. Holm L, Kaariainen S, Rosenstrom P, Schenkel A (2008) Searching protein structure databases with DaliLite v.3. Bioinformatics 24:2780–2781.
50. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF, Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures. J Mol Biol 112:535–542.
51. DeLano WL (2002) The PyMOL Molecular Graphics System. San Carlos, CA: DeLano Scientific.
52. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res 32:W665–W667.
53. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci USA 98:10037–10041.
54. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A (2008) The Pfam protein families database. Nucleic Acids Res 36:D281–D288.
55. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and clustal X version 2.0. Bioinformatics 23:2947–2948.
56. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293.