PROTEIN STRUCTURE REPORT

The crystal structure *Escherichia coli* Spy

Eunju Kwon,^{1,2} Dong Young Kim,^{1,3} Carol A. Gross,² John D. Gross,^{3*} and Kyeong Kyu Kim^{1*}

¹Department of Molecular Cell Biology, Samsung Biomedical Research Institute, Sungkyunkwan University School of Medicine, Suwon, Korea

²Department of Microbiology and Immunology, University of California San Francisco, San Francisco, California

³Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California

Received 11 June 2010; Revised 4 August 2010; Accepted 9 August 2010

DOI: 10.1002/pro.489

Published online 26 August 2010 proteinscience.org

Abstract: *Escherichia coli* spheroplast protein y (EcSpy) is a small periplasmic protein that is homologous with CpxP, an inhibitor of the extracytoplasmic stress response. Stress conditions such as spheroplast formation induce the expression of Spy via the Cpx or the Bae two-component systems in *E. coli*, though the function of Spy is unknown. Here, we report the crystal structure of EcSpy, which reveals a long kinked hairpin-like structure of four α -helices that form an antiparallel dimer. The dimer contains a curved oval shape with a highly positively charged concave surface that may function as a ligand binding site. Sequence analysis reveals that Spy is highly conserved over the Enterobacteriaceae family. Notably, three conserved regions that contain identical residues and two LTxxQ motifs are placed at the horizontal end of the dimer structure, stabilizing the overall fold. CpxP also contains the conserved sequence motifs and has a predicted secondary structure similar to Spy, suggesting that Spy and CpxP likely share the same fold.

Keywords: Spy; spheroplast; extracytoplasmic stress response; Cpx; crystal structure

Abbreviations: Bae, bacterial adaptive response; Cpx, conjugation plasmid expression; EcSpy, the matured form of *Escherichia coli* Spy; ESR, extracytoplasmic stress response; MtCsoR, *Mycobacterium tuberculosis* copper-sensitive operon repressor; Rcs, regulator of colanic acid capsule synthesis; RMSD, root mean square deviation; SEC-MALLS, size exclusion chromatography with multiangle laser light scattering.

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: 21C Frontier Functional Proteomics Program; Grant number: FPR08B2-270; Grant sponsor: Ubiquitome Research Program; Grant number: M105 33010001-05N3301-00100; Grant sponsor: National Research Laboratory Program; Grant number: NRL-2006-02287.

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

Introduction

Extracytoplasmic stress response (ESR) systems play an important role in coping with envelope stress by regulating gene expression in response to stress signals. In *Escherichia coli*, three ESR systems are well known: the sigmaE proteolytic cascade, and the conjugation plasmid expression (Cpx) and the bacterial adaptive response (Bae) two-component systems.^{1–7} Each ESR system is activated by distinct triggers such as unfolded proteins and metal ions. ESR systems transduce signals to the cytoplasmic space across the inner membrane and regulate the expression of target genes. Although ESR systems independently regulate target genes, their regulons are partially overlapping.

The overexpression of Spy (Spheroplast protein y) is regulated by ESR systems. When a bacterial cell wall is stripped off, which is referred to as the

Table I. Data Collection and Refinement Statistics

Data collection statistics					
Data set	Native data (ALS 8.3.1)	MAD data (PAL BL 4A)			
		Peak	Edge	Remote 1	Remote 2
Space group	$P3_2$				$P6_222$
Unit cell (Å)	a = 69.20, c = 126.35				a = 69.50, c = 128.48
Resolution (Å)	20.0–2.70 (2.80–2.70) ^a				50.0–3.0 (3.11–3.00)
Wavelength (Å)	1.11587	0.97944	0.97976	0.98771	0.97220
Total/unique reflections	75029/18354	160659/4145	159089/4092	157808/4064	161819/4139
Completeness (%)	98.3 (95.3)	99.8 (100.0)	99.8 (100.0)	99.7 (100.0)	99.8 (100.0)
I/σ (I)	25.4 (2.1)	59.6 (8.0)	60.5 (9.8)	62.8 (13.4)	59.6 (8.6)
R_{merge} (%)	5.4 (38.5)	10.8 (41.1)	10.0 (35.8)	9.2 (29.0)	11.5 (40.1)
Refinement statistics					
Resolution (Å)	20.0–2.7 (2.77–2.70)				
No. of reflections	17348 (1252)				
$R_{\text{work}}/R_{\text{free}}$	0.252/0.300 (0.318/0.354)				
No. of atoms	3016/22/90				
protein/cadmium					
ion/water					
RMSD Bond lengths (Å)/Bond angles (°)	0.007/1.116				
Ramachandran statistics (%) ^b	85.3/14.4/0.3				

^a The highest resolution shell is shown in parenthesis.

^b Ramachandran statistics indicate the fraction of residues in the favored, allowed, and disallowed regions of the Ramachandran diagram.

spheroplast formation, Spy is overexpressed by both the Cpx⁸ and the Bae two-component systems in *E. coli*.⁹ However, it is not detected in normal cells.⁸ In addition to spheroplast formation, other stimuli also activate the expression of Spy. Copper ions lead to Spy expression via the Cpx pathway,¹⁰ whereas zinc ions cause its expression via the Bae pathway.¹¹ Tannin¹² and indole¹³ enhance Spy transcription through the Bae pathway. The activation of the Rcs two-component system is also reported to enhance the expression of Spy.¹⁴ These data indicate various triggers induce the overexpression of Spy via either separation or simultaneous activation of ESR systems.

Though the triggers and signaling pathways leading to Spy expression have been well studied, the function of Spy is unknown. The *spy*-deletion mutant possesses no defect under both normal growth conditions and conditions of stress.^{8,9} However, a clue comes from the observation that Spy shares the identity of 25.5% with CpxP, which is a member of the Cpx two-component system. CpxP functions as a negative regulator of the Cpx two-component system and is a stress sensor of the Cpx pathway,^{3,15} implying that the function of Spy might be related to the ESR system.

To gain insights into the function of Spy, we have determined its crystal structure. We find that *E. coli* Spy (EcSpy) has a unique fold with a long kinked hairpin-like structure of four α -helices and forms a dimer. In the EcSpy dimer, the four helices are arranged in an antiparallel orientation that

results in a curved oval shape with a highly positively charged concave surface that may function as a ligand binding site. The residues on the dimeric interface are highly conserved among Spy proteins of enterobacteria as well as CpxP, suggesting that Spy and CpxP have similar structures.

Results and Discussion

Overall structure of EcSpy

The mature form of EcSpy, composed of residues 24–161, was expressed as a recombinant protein and purified by Ni-chelating resin and gel filtration. The crystal of selenomethionine-labeled EcSpy belongs to the hexagonal $P6_222$ space group and diffracted to 3.0 Å resolution. The phases were determined by the multiwavelength anomalous dispersion method (Table I). The native crystal of EcSpy diffracted to 2.7 Å resolution. Instead of the $P6_222$ space group, the native data were scaled in the $P3_2$ space group, containing four monomers in an asymmetric unit [Fig. 1(A)]. The final model was refined with the R and R_{free} of 25.2 and 30.0%, respectively (Table I). It contains residues 53–143 for each monomer, 22 cadmium ions, and 90 water molecules. Most of the cadmium ions are involved in crystal contacts except four, which interact with His119 of each monomer. Residues that were not observed in the electron density map due to disorder include N-terminal residues 24–52 and C-terminal residues 144–161 [Fig. 3(C)].

Figure 1. The overall structure of EcSpy. A: The ribbon diagram of four monomers in an asymmetric unit. Chain A, B, C, and D are shown as green, blue, yellow, and violet colors, respectively. B: The ribbon diagram of EcSpy monomer in the stereoview. The structure is colored with rainbow color: blue at N-terminus and red at C-terminus. The N-terminus, C-terminus, and helices (H1–H4) are labeled. The interhelical angle between H1 and H2 and the bend-angle of H3 are designated. An [interactive view](#) is available in the electronic version of the article.

The structures of the four monomers observed in the asymmetric unit are almost identical [Fig. 1(A)]. The root mean square deviation (RMSD) values of main chain atoms for pairs of monomers fall between 0.866 Å (RMSD between chain A and C) and 1.095 Å (RMSD between chain C and D). Chain A is used as the EcSpy monomer for the following analysis and figures.

The monomer structure consists of four α -helices exhibiting a long and kinked hairpin-like conformation [H1–H4, Fig. 1(B)]. H1 and H2 and the longest helix H3 form two prongs that are connected via a hairpin-loop between H2 and H3 [Fig. 1(B)]. H3 is kinked with a bend-angle of 135° at Lys107. The loop between H1 and H2 also introduces a bend that makes an interhelical angle of 105° between H2 and H3. The C-terminal helix, H4, and the N-terminal loop are curved inward [Fig. 1(B)].

To gain insight into the function of EcSpy, the structure was searched against the Protein Data Bank using the DaliLite server.¹⁶ However, no ho-

mologous structures were found. The top scoring structures were helical regions of larger proteins or a part of helical bundles that only marginally resembled Spy (Supporting Information Figure 1). Therefore, the structure of EcSpy is unique, and there are no homologous structures that can give information about the function of EcSpy.

Dimer structure of EcSpy

Because four monomers have a close contact in the asymmetric unit, the oligomeric state of EcSpy in solution was tested by crosslinking and by SEC-MALLS, which revealed that EcSpy exists as a dimer (Supporting Information Figure 2). We calculated the buried surface areas and the theoretical solvation free energies of the probable dimer interfaces to deduce the correct dimer structure (Supporting Information Table I). The dimer interfaces between chain A and C and chain B and D are strong enough to form the stable dimer: the buried surface area of larger than 1400 Å² and the

Figure 2. The proposed dimer structure of EcSpy. A: The ribbon diagram of the EcSpy dimer that includes chain A and C. The color scheme is same as A, and helices are labeled. The dimer interfaces indicated in panels B (left) and C (right) are depicted by two red circles. B and C: The close-up views of the dimer interfaces. The residues making a hydrogen bond or a salt bridge are shown in the stick diagram on the ribbon diagram. Hydrogen bonds and the salt bridge are depicted as blue and red dashed lines, respectively. Chain A and C are colored with pale green and light orange. For a clear view, the orientation is slightly modified from that of A. D: The surface charge distribution analysis of the EcSpy dimer. The charge distribution is shown on EcSpy dimer surface figure within a range of -5 kT (red) to $+5$ kT (blue) on left. The orientation of the surface model on the left is same as that of Figure 2(A). An [interactive view](#) is available in the electronic version of the article.

theoretical solvation free energy of less than -23 kcal/mol, an order of magnitude less than that predicted by the other interfaces (Supporting Information Table 1). Therefore, we suggest that the dimer of chain A and C (or equivalently chain B and D) is the stable dimer. Accordingly, the dimer structure of EcSpy chain A and C is referred to as the EcSpy dimer in the following discussion.

In the EcSpy dimer, chain A interacts with chain C in an antiparallel manner. The hairpin-loop and the nearby region (H2 and N-terminus of H3) in chain A is placed in proximity with the C-terminus of H3 and H4 in chain C and visa versa [Fig. 2(A)].

The EcSpy dimer has a rather flat and concave oval shape [Fig. 2(A,D)], reminiscent of a ligand receptor.

The dimer interface contains residues that form hydrogen bonds and a salt bridge. The NE2 atom of His88 and the O atom of Thr92 on H2 in chain A make hydrogen bonds with the O atom of Asn139 and the OH atom of Tyr127 in chain C, respectively [Fig. 2(B)]. The O atom of Ser93 on the hairpin-loop in chain A forms a hydrogen bond with the ND2 atom of Asn128 in chain C, whereas the OD1 atom of Asp94 on the hairpin-loop in chain A forms a salt bridge with the NZ atom of Lys135 in chain C [Fig. 2(B)]. The ND2 atom of Asn124 and the ND2 atom of

distributed over the N-terminus and H1. Phe52 is not modeled in the structure. Region 2 (His88, Phe96, and Asp97) is located in H2 and within the hairpin-loop. Region 3 comprises eight identical residues (Asn124, Tyr127, Asn128, Leu130, Thr131, Gln134, Lys135, and Asn139), which are placed over H3 to H4 [Fig. 3(B,C)]. On right of the dimer structure [Fig. 3(A)], Region 2 of chain C and Regions 1 and 3 of chain A are placed in parallel forming a cluster of residues that are close in space [Fig. 3(B)]. Four of ten residues making a hydrogen bond or a salt bridge in the dimer belong to these conserved residues: His88, Asn124, Tyr127, and Asn128 [Figs. 2(B,C) and 3(C)].

Both Regions 1 and 3 of EcSpy contain the LTxxQ motif, which is also found in CpxP [Fig. 3(C); where x can be any amino acid]. This motif (repeated twice) is found in the Spy and CpxP proteins of many bacterial species, as well as in multiple unannotated CpxP-homologs having less than 200 amino acids. The variable residues of the LTxxQ motif in Regions 1 and 3 form a pattern that is conserved in aligned homologs. The LTxxQ motif in Region 1 has an acidic amino acid (Asp or Glu) and Ala for the third and the fourth residues, whereas Region 3 has Pro and Glu (Supporting Information Figure 3). It is intriguing that both LTxxQ motifs are placed together at the very end of the dimer [Fig. 3(B)]. Though the role of LTxxQ motif is not understood, it is known that the mutations in the fifth Gln residue of LTxxQ motif in EcCpxP reduce the protein stability and induce the DegP-dependent degradation of EcCpxP.² These observations suggested that the LTxxQ motif seems to be important for maintaining the folding of CpxP and Spy. Taken together, the amino acid similarity suggests that the structure and the dimeric interface of EcSpy are very well conserved in the Spy and CpxP proteins of enterobacteria.

CpxP is a well-known homolog of Spy, and its sequence also exhibits characteristics that are shown in EcSpy. The predicted α -helical region in EcCpxP matches the secondary structure of EcSpy [Fig. 3(C)]. In addition, EcCpxP retains 13 of the 15 residues that are conserved in the multiple sequence alignment of Spy homologs, and three of them are involved in the dimerization: His82 (His88), Asn118 (Asn124), and Tyr121 (Tyr127), where the number in parenthesis indicates the residue number of EcSpy [Fig. 3(C)]. The LTxxQ motif, which is likely to be critical for maintaining the structure, is also nearly identical in EcCpxP and EcSpy [Fig. 3(C)]. Accordingly, EcCpxP is predicted to have helical structure similar to the EcSpy structure and to form a dimer in the same antiparallel manner.

In summary, we have determined the crystal structure of EcSpy. It has a long and kinked hairpin-like conformation, which is not represented in

the structural database. Spy forms a stable dimer of interlocked antiparallel helical hairpins that generate a concave oval shape. The LTxxQ motif and interfacial residues are conserved in Spy homologs of enterobacteria and help maintain structural integrity. These same residues are conserved in CpxP, a negative regulator of Cpx signaling. Additionally, we suggest that the positively charged concave surface of Spy may serve as a ligand binding site, possibly engaging lipopolysaccharides or phospholipids. Therefore, the Spy fold may be used for ligand recognition and modulation of ESR in multiple contexts across enterobacteria. Our structural studies will help to guide further biochemical and genetic experiments designed to identify the ligands of Spy and CpxP and their role in stress responses.

Materials and Methods

Cloning and purification

The mature form of EcSpy (24–161) was cloned into the modified pET28a(+) vector to have His-Trx tag before the thrombin cutting site. The plasmid was transformed into BL21(DE3) cells (Novagen). The protein expression was inoculated in Luria Bertani medium and cultured at 37°C. When the optical density at 600 nm reached 0.5–0.6, cells were induced by 0.5 mM isopropyl β -D-thiogalactopyranoside. Induced cells were cultured at 30°C for 5 h and harvested by centrifugation. The cell pellet was resuspended with buffer A (20 mM Tris-HCl pH 7.5, 0.1M NaCl, and 5% glycerol), disrupted by sonication, and fractionated by centrifugation. Clarified lysate was loaded onto a HisTrap HP column (GE Healthcare), and the protein was eluted by an imidazole gradient. The fractions containing HisTrx-EcSpy were pooled and dialyzed against buffer A. The dialyzed sample was treated with thrombin (Roche), and the cleaved His-Trx tag was removed by flowing over a HisTrap HP column (GE Healthcare). The sample was finally purified using a HiLoad Superdex75 16/60 gel filtration column (GE Healthcare). The selenomethionine-substituted protein was expressed in B834(DE3) cell with M9 minimal medium in the presence of selenomethionine and purified by same procedure as native protein.

Crystallization and structure determination

The EcSpy was crystallized using the modified microbatch method at 22°C. One microliter of protein solution (60 mg/mL) was mixed with 1 μ L of crystallization reagent containing 0.1M acetate pH 4.6, 26–32% (w/v) PEG400, 0.2M CdCl₂, and 2 mM TCEP under a layer of Al's oil (Hampton research). The crystals were fully grown in a month and flash-frozen in a cold nitrogen stream. The X-ray diffraction data of native EcSpy were collected to a 2.7 Å

resolution at ALS BL8.3.1 (Advanced Light Source Beamline 8.3.1). Mutiwavelength anomalous dispersion (MAD) data were collected to a 3.0 Å resolution with ADSC Quantum 210 CCD detector at PLS BL4A (Pohang Light Source Beamline4A, South Korea). Native and four wavelength MAD (peak, inflection, two remotes) data were processed and integrated using HKL2000 and scaled using SCALEPACK.¹⁹ The native and MAD data were processed to the space group $P3_2$ and $P6_222$ with nearly same unit cell parameters, respectively.

Seven selenium sites were found out of a total of 11 selenomethionine residues and used for phase calculation in the SOLVE program. Electron density was generated by RESOLVE after density modification.^{20,21} Because the electron density was too poor to trace automatically, the initial model was manually built using COOT,²² based on seven selenium sites and the secondary structure prediction. There is a monomer in the asymmetric unit of the $P6_222$ space group. To improve the model, several cycles of rigid-body refinement, positional refinement, restrained refinement, simulated annealing, B-factor refinement, and model rebuilding were performed at 3.0 Å resolution using CCP4,²³ Refmac,²⁴ CNS,²⁵ and COOT²² programs.

There are four monomers in an asymmetric unit of the native data in the $P3_2$ space group. The orientations of four monomers were searched using the Phaser program.²⁶ The model was further built and refined using the diffraction data of native Spy. The final refinement was performed by Refmac²⁴ with translation-libration-screw restraint. NCS (noncrystallographic symmetry) restraint was released to reduce R factor during final round of refinement. It resulted in the R and R_{free} of 25.2 and 30.0%, respectively. The quality of the structure was analyzed with PROCHECK.²⁷ The data collection and refinement statistics are summarized in Table I. Figures were generated using PyMOL.²⁸

Sequence and structure analysis

The protein sequence alignment of EcSpy and EcCpxP was performed using ClustalW2.²⁹ The secondary structure predictions were performed by PSIPRED.³⁰ The domain search against the protein family database was executed with Pfam.³¹ Interhelical angles between α -helices were calculated using QHELIX.³² The electrostatic surface potential was calculated by PDB2PQR server³³ and visualized by the APBS plugin of PyMOL.^{28,34}

Acknowledgments

The authors thank the staff of the PAL BL 4A and the ALS beamline 8.3.1 for their assistance during data collection, Dr. Daniel Southworth for the help in SEC-MALLS, and Robin Aglietti for the proofreading of

manuscript. Database: The coordinates of the structure for *Escherichia coli* Spy have been deposited in the Protein Data Bank under the accession number 3OEO.

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. 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.
3. Raivio TL, Popkin DL, Silhavy TJ (1999) The Cpx envelope stress response is controlled by amplification and feedback inhibition. *J Bacteriol* 181:5263–5272.
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. Hagenmaier S, Stierhof YD, Henning U (1997) A new periplasmic protein of *Escherichia coli* which is synthesized in spheroplasts but not in intact cells. *J Bacteriol* 179:2073–2076.
9. Raivio TL, Laird MW, Joly JC, Silhavy TJ (2000) Tethering of CpxP to the inner membrane prevents spheroplast induction of the cpx envelope stress response. *Mol Microbiol* 37:1186–1197.
10. Yamamoto K, Ishihama A (2005) Transcriptional response of *Escherichia coli* to external copper. *Mol Microbiol* 56:215–227.
11. Yamamoto K, Ogasawara H, Ishihama A (2008) Involvement of multiple transcription factors for metal-induced spy gene expression in *Escherichia coli*. *J Biotechnol* 133:196–200.
12. Zoetendal EG, Smith AH, Sundset MA, Mackie RI (2008) The BaeSR two-component regulatory system mediates resistance to condensed tannins in *Escherichia coli*. *Appl Environ Microbiol* 74:535–539.
13. Garbe TR, Kobayashi M, Yukawa H (2000) Indole-inducible proteins in bacteria suggest membrane and oxidant toxicity. *Arch Microbiol* 173:78–82.
14. Bury-Mone S, Nomane Y, Reymond N, Barbet R, Jaquet E, Imbeaud S, Jacq A, Boulon P (2009) Global analysis of extracytoplasmic stress signaling in *Escherichia coli*. *PLoS Genet* 5:e1000651.
15. 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.
16. Holm L, Kaariainen S, Rosenstrom P, Schenkel A (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics* 24:2780–2781.
17. Tarricone C, Xiao B, Justin N, Walker PA, Rittinger K, Gamblin SJ, Smerdon SJ (2001) The structural basis of Arfaptin-mediated cross-talk between Rac and Arf signalling pathways. *Nature* 411:215–219.

18. Liu T, Ramesh A, Ma Z, Ward SK, Zhang L, George GN, Talaat AM, Sacchettini JC, Giedroc DP (2007) CsoR is a novel *Mycobacterium tuberculosis* copper-sensing transcriptional regulator. *Nat Chem Biol* 3:60–68.
19. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
20. Terwilliger TC (2003) Automated main-chain model building by template matching and iterative fragment extension. *Acta Cryst D* 59:38–44.
21. Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Cryst D* 55:849–861.
22. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Cryst D* 60:2126–2132.
23. Potterton E, Briggs P, Turkenburg M, Dodson E (2003) A graphical user interface to the CCP4 program suite. *Acta Cryst D* 59:1131–1137.
24. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst D* 53:240–255.
25. Brünger 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 D* 54:905–921.
26. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Cryst* 40:658–674.
27. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 26:283–291.
28. DeLano WL (2002) The PyMOL Molecular Graphics System. DeLano Scientific: San Carlos.
29. Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics Chapter 2:2.3.1–2.3.22*.
30. McGuffin LJ, Bryson K, Jones DT (2000) The PSIPRED protein structure prediction server. *Bioinformatics* 16:404–405.
31. Finn RD, Mistry J, Tate J, Coghill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A (2008) The Pfam protein families database. *Nucleic Acids Res* 36:D281–D288.
32. Lee HS, Choi J, Yoon S (2007) QHELIX: a computational tool for the improved measurement of inter-helical angles in proteins. *Protein J* 26: 556–561.
33. 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.
34. 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.