Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallographic characterization of a multidomain histidine protein kinase from an essential two-component regulatory system

Haiyan Zhao and Liang Tang*

Department of Molecular Biosciences, University of Kansas, 1200 Sunnyside Avenue, Lawrence, KS 66045, USA

Correspondence e-mail: tangl@ku.edu

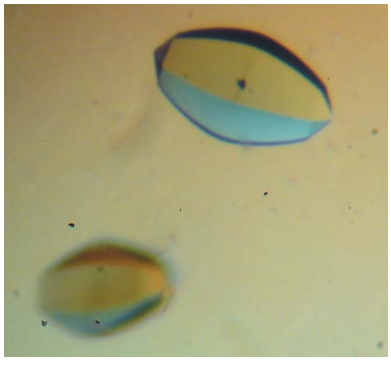
Received 16 January 2009
Accepted 23 February 2009

YycGF is a highly conserved two-component signal transduction system that is specific to low-G+C Gram-positive bacteria, including many important human pathogens. It has been recognized as a crucial regulatory system for cell-wall metabolism. YycG, the histidine protein kinase of this system, is a multidomain transmembrane protein. The truncated cytoplasmic portion of YycG from *Bacillus subtilis* encompassing the PAS domain, the dimerization domain and the catalytic domain was expressed, purified and crystallized. X-ray data were collected to 2.8 Å resolution with a completeness of 98.2% and an overall $R_{\rm merge}$ of 5.6%. The crystals belonged to space group $P6_1$ or $P6_5$, with unit-cell parameters a = 135.0, c = 133.0 Å. The selenomethionine-substituted version of the protein was crystallized and X-ray data were collected to 3.6 Å resolution for subsequent MAD phasing.

1. Introduction

YycGF is a two-component regulatory system (TCS) that is specific to low-G+C Gram-positive bacteria, which include important human pathogens such as *Staphylococcus*, *Streptococcus* and *Enterococcus* as well as potential bioterror agents such as *Bacillus anthracis* and *Clostridium botulinum* (Winkler & Hoch, 2008; Fabret & Hoch, 1998; Dubrac *et al.*, 2008). The histidine protein kinase (HK) YycG senses extracellular or intracellular signals and phosphorylates its cognate response regulator (RR) YycF, which in turn recognizes sequence-specific regions on the bacterial chromosome and regulates the expression of certain genes. The YycGF TCS has been demonstrated to be a crucial signal transduction system in regulation of bacterial cell-wall biosynthesis and homeostasis, cell-membrane composition, virulence, cell division and biofilm formation (Dubrac & Msadek, 2004; Fukuchi *et al.*, 2000; Howell *et al.*, 2003; Mohedano *et al.*, 2005; Ng *et al.*, 2005; Dubrac *et al.*, 2007; Fukushima *et al.*, 2008). It is among the very few two-component regulatory systems that are essential to cell viability (Fabret & Hoch, 1998; Clausen *et al.*, 2003; Hancock & Perego, 2004; Martin *et al.*, 1999), making it an attractive antibacterial target (Qin *et al.*, 2006; Furuta *et al.*, 2005; Winkler & Hoch, 2008; Watanabe *et al.*, 2008).

The histidine protein kinases YycG are highly conserved among low-G+C Gram-positive bacteria, showing sequence identity ranging from 37% to 65% and a similar domain organization. YycG from *B. subtilis* contains an amino-proximal extracellular signal-sensing domain flanked by two transmembrane regions. The cytoplasmic portion of YycG contains a HAMP (histidine kinase, adenylyl cyclase, methyl-accepting chemotaxis proteins and phosphatase; also called P-type linker) domain, followed by a PAS (Per–Arnt–Sim) domain, a dimerization domain that contains the active histidine residue and a C-terminal catalytic/ATP-binding domain. This multi-domain transmembrane architecture is shared by most HKs in two-component regulatory systems (West & Stock, 2001; Stock *et al.*, 2000). The molecular mechanisms of signal transduction across the plasma membrane through the cooperation of multiple domains in these HKs remain to be understood. Previous structural and mechanistic studies on HKs have been limited to isolated individual domains (Varughese, 2002; Varughese *et al.*, 2007; West & Stock, 2001;



Stock *et al.*, 2000) or portions containing two HK domains (Bilwes *et al.*, 1999; Marina *et al.*, 2005) and no structure is available for any full-length HK. Assembly of an RR with its truncated HK consisting of the dimerization domain, the catalytic domain and the PAS module from *Thermotoga maritima* has been studied using small-angle X-ray scattering and low-resolution X-ray crystallography (Yamada *et al.*, 2006). Nevertheless, it provided little detailed structural information owing to the limited resolution.

Here, we report the preliminary crystallographic characterization of the histidine protein kinase YycG from *B. subtilis*. A cytoplasmic portion of YycG encompassing the PAS domain, the dimerization domain and the catalytic domain was expressed, purified and crystallized. X-ray data were collected to 2.8 Å resolution. Selenomethionine-substituted (SeMet) protein was crystallized for MAD phasing. To our knowledge, this is the first HK containing the PAS domain, the dimerization domain and the catalytic domain that has been crystallized at a resolution sufficient for structural analysis at atomic detail.

2. Materials and methods

2.1. Protein expression and purification

We attempted the expression and purification of various truncated forms of YycG from *B. subtilis* in order to maximize the success rate of crystallization and also to allow comparative biochemical and functional analysis in order to dissect the respective roles of individual domains. One of these yielded soluble and well behaved protein; this was a truncated cytoplasmic portion of YycG that encompassed the PAS domain, the dimerization domain and the catalytic domain (residues 272–611; designated YycG-PDC). The gene of YycG-PDC was cloned from *B. subtilis* strain 168 genomic DNA into the pET30b vector (Novagen) using PCR primers that contained *Nde*I and *Xho*I restriction sites. The PCR primers were designed such that a stop codon was inserted before the *Xho*I restriction site and the resultant protein contained no His tag. The resultant plasmid was transformed in *Escherichia coli* strain BL21(DE3) pLysS (Novagen). The strain was grown at 288 K in LB media supplemented with 30 µg ml^{-1} kanamycin and 34 µg ml^{-1}chloramphenicol. Protein expression was induced at OD_{600} = 0.6 by the addition of 1 m*M* IPTG and shaken overnight. Cells were harvested, washed with buffer *A* (20 m*M* Tris pH 7.5, 50 m*M* NaCl, 1 m*M* EDTA), resuspended in buffer *A* and frozen at 253 K. The cells were dissolved in ice–water and lysed by sonication with 1 m*M* phenylmethylsulfonyl fluoride in the solution. Cellular debris was removed by centrifugation in a Sorvall RC5B centrifuge for 60 min at 15 000 rev min^{-1}. The supernatant was loaded onto a 5 ml HiTrap Q column, washed with buffer *A* and eluted with buffer *B* (buffer *A* + 500 m*M* NaCl). The fractions containing 80% pure YycG were pooled and loaded onto a Sephacryl S-300 size-exclusion chromatography column (GE Healthcare), concentrated to ∼20 mg ml^{-1} with an Ultrafree concentrator (Millipore; molecular-weight cutoff 30 000 Da) and stored at 193 K. The yield was ∼15–20 mg purified protein per litre of bacterial culture. The protein was >95% pure as shown by SDS–PAGE analysis. Size-exclusion chromatography showed the protein existed in solution as a dimer, which is consistent with observations of other TCS histidine protein kinases. Biochemical assays indicated that the purified protein had autophosphorylation activity (Fig. 1) and was capable of phosphorylating its cognate response regulator YycF (Zhao & Tang, unpublished results). SeMet YycG-PDC was expressed and purified in essentially the same way, except that the methionine-auxotrophic *E. coli* strain B834(DE3) pLysS and a minimal medium supplied with 40 µg ml^{-1} selenomethionine (Fisher Scientific) were used. MALDI–TOF mass spectrometry (Voyager, Applied Biosystems) showed a molecular mass of 39 494.3, which is consistent with the presence of nine Met residues in the protein (the N-terminal Met was absent).

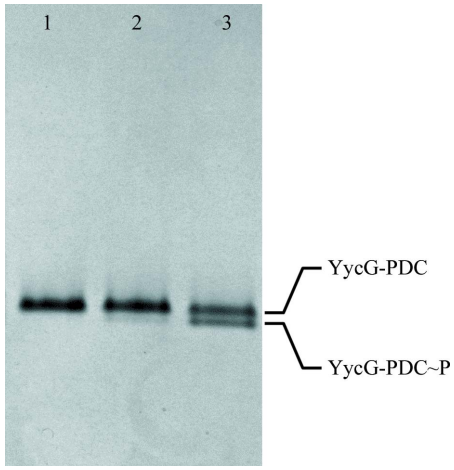


Figure 1
Autophosphorylation assay of YycG-PDC by native gel electrophoresis analysis. In lanes 2 and 3, 5 µg purified YycG-PDC protein was incubated with reaction buffer containing 10 m*M* ATP in 50 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl, 10 m*M* $MgCl_2$, 1 m*M* DTT and 0.1 m*M* EDTA at room temperature for 1 and 10 min, respectively, and the reaction was then stopped by the addition of EDTA to a final concentration of 50 m*M* and analyzed by native gel electrophoresis. In lane 1, the protein was incubated with reaction buffer without ATP for 30 min as a control. The band corresponding to the phosphorylated protein is indicated by YycG-PDC~P. In lane 2, a small fraction of the protein was autophosphorylated. In lane 3, about one third of the total protein was autophosphorylated.

2.2. Crystallization

Initial crystallization screening was performed using commercially available screening kits from Hampton Research and Emerald BioSystems. After optimization of crystallization conditions, crystals of YycG-PDC were grown at 293 K by the vapor-diffusion method in hanging drops containing 1 µl each of protein solution at 10 mg ml^{-1} and of reservoir solution consisting of 100 m*M* cacodylate pH 6.5, 0.2 *M* magnesium formate, 0.2 *M* sodium chloride and 10% PEG 200

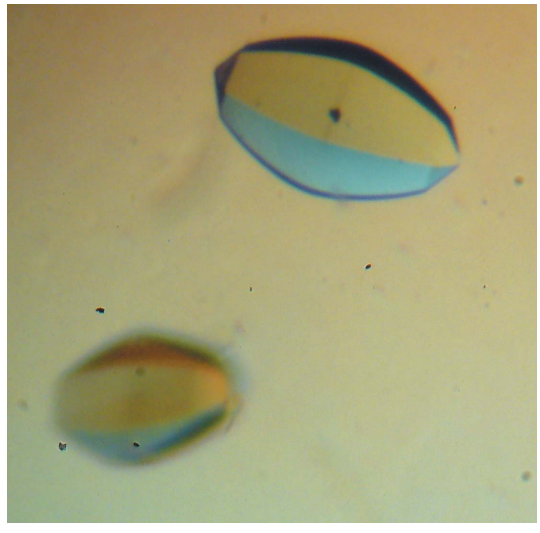


Figure 2
Native crystals of YycG-PDC. The larger crystal is approximately 0.4 × 0.2 × 0.2 mm in size.

Table 1
X-ray data-collection and processing statistics.

Values in parentheses are for the outermost resolution shell.

		SeMet YycG-PDC		
	Native YycG-PDC	Peak	Remote	Inflection
Beamline	SSRL 9-2	SSRL 9-2	SSRL 9-2	SSRL 9-2
Wavelength (Å)	0.9794	0.9792	0.9793	0.9116
Resolution (Å)	50–2.8	50–3.6	50–3.6	50–3.6
No. of measurements	146869	56293	56077	56258
No. of unique reflections	33486	16046	16038	16091
Completeness (%)	98.2 (95.7)	99.0 (99.0)	98.8 (99.9)	99.0 (99.9)
$I/\sigma(I)$	31.5 (2.1)	16.5 (2.8)	16.5 (3.0)	16.0 (2.7)
$R_{\rm merge}$† (%)	5.6 (35.6)	8.1 (32.2)	7.1 (30.1)	8.2 (33.3)
Space group	$P6_1$ or $P6_5$	$P6_1$ or $P6_5$	$P6_1$ or $P6_5$	$P6_1$ or $P6_5$
Unit-cell parameters (Å)	a = 135.0, c = 133.0	a = 135.5, c = 133.7		

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity of reflection hkl and $\langle I(hkl)\rangle$ is the averaged intensity of symmetry-equivalent measurements.

equilibrated against 0.5 ml reservoir solution. Crystals grew to maximal dimensions of approximately 0.4 × 0.2 × 0.2 mm within one to two weeks (Fig. 2). The crystallization conditions for SeMet YycG-PDC were slightly different from those of the native protein in that the optimal reservoir solution pH was 7.5. SeMet YycG-PDC crystals appeared to grow more slowly than native crystals and reached maximal dimensions of 0.2 × 0.1 × 0.1 mm within four weeks.

2.3. Data collection and processing

For cryogenic experiments, crystals were soaked in mother liquor plus 20% PEG 200 for 10–30 s and flash-frozen in liquid nitrogen. X-ray data were collected at 100 K on Stanford Synchrotron Radiation Laboratory (SSRL) beamline 9-2 using a MAR 325 detector equipped with Stanford Automated Mounting robotics. For the native crystal, a total of 60 frames were collected over 60° with an oscillation range of 1° and an exposure time of 15 s per frame (Fig. 3). For the SeMet crystal, a total of 55 frames were collected with an oscillation range of 1° and an exposure time of 2 s per frame for the peak, remote and inflection wavelengths. The diffraction data were processed with *HKL*-2000 (Otwinowski & Minor, 1997). Data-collection and processing statistics are summarized in Table 1.

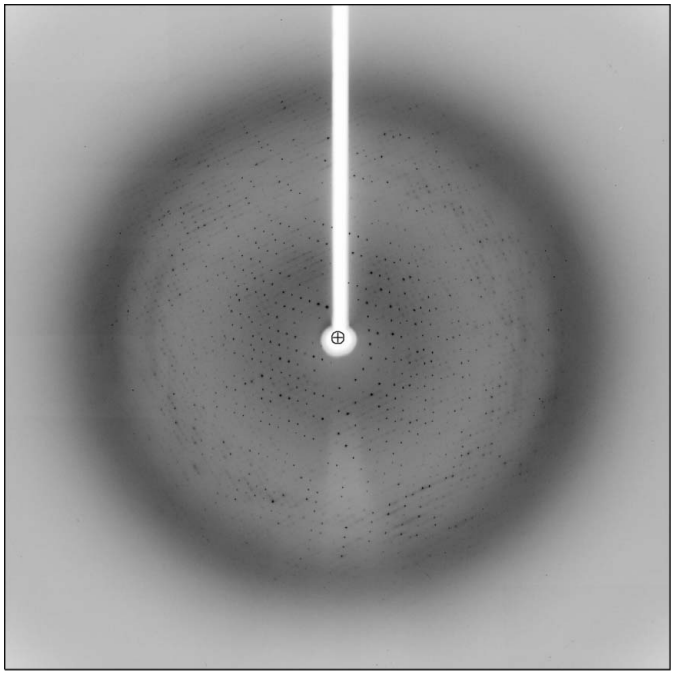


Figure 3
A diffraction pattern of a native YycG-PDC crystal.

3. Results and discussion

The crystals of native YycG-PDC belonged to space group $P6_1$ or $P6_5$, with unit-cell parameters a = 135.0, c = 133.0 Å. The SeMet crystals had a slightly different unit cell (Table 1). The native crystal data set contained a total of 33 486 reflections in the resolution range 50.0–2.8 Å, with a completeness of 98.2% and an $R_{\rm merge}$ of 5.6%. The crystal may contain two kinase homodimers in the crystallographic asymmetric unit, which corresponds to 47.2% solvent content and a $V_{\rm M}$ value of 2.3 Å^3 Da^{-1} (Matthews, 1968). However, considering the current medium resolution of 2.8 Å and the fragility of the crystals, it cannot be ruled out that the asymmetric unit may contain only one kinase homodimer, corresponding to 73.6% solvent content and a $V_{\rm M}$ value of 4.6 Å^3 Da^{-1}.

The cytoplasmic portion of a hypothetical HK from *T. maritima* (PDB code 2c2a; Marina *et al.*, 2005) contains the dimerization domain and the catalytic domain and is the nearest homologous structure to YycG, with a sequence identity of 34%. However, molecular replacement has not been successful using the structure of the *T. maritima* HK as the initial search model. This suggests a different domain arrangement in YycG-PDC compared with that in the *T. maritima* HK. Molecular-replacement attempts failed using search models built from available homologous structures of either the catalytic domain of other histidine protein kinases or the PAS domain of PAS-like sensing proteins such as the heme-binding protein (PDB code 1lsw; Hao *et al.*, 2002), which showed 19% sequence identity. Structure determination using MAD phasing is in progress.

We thank the staff at the Stanford Synchrotron Radiation Laboratory for assistance in X-ray data collection and Dr Nadezhda Gelava at the University of Kansas Analytical Proteomics Laboratory for assistance in mass spectrometry. This work was supported by the NIH/NCRR COBRE program P20RR17708.

References

Bilwes, A. M., Alex, L. A., Crane, B. R. & Simon, M. I. (1999). *Cell*, **96**, 131–141.
Clausen, V. A., Bae, W., Throup, J., Burnham, M. K., Rosenberg, M. & Wallis, N. G. (2003). *J. Mol. Microbiol. Biotechnol.* **5**, 252–260.
Dubrac, S., Bisicchia, P., Devine, K. M. & Msadek, T. (2008). *Mol. Microbiol.* **70**, 1307–1322.
Dubrac, S., Boneca, I. G., Poupel, O. & Msadek, T. (2007). *J. Bacteriol.* **189**, 8257–8269.
Dubrac, S. & Msadek, T. (2004). *J. Bacteriol.* **186**, 1175–1181.
Fabret, C. & Hoch, J. A. (1998). *J. Bacteriol.* **180**, 6375–6383.
Fukuchi, K., Kasahara, Y., Asai, K., Kobayashi, K., Moriya, S. & Ogasawara, N. (2000). *Microbiology*, **146**, 1573–1583.
Fukushima, T., Szurmant, H., Kim, E. J., Perego, M. & Hoch, J. A. (2008). *Mol. Microbiol.* **69**, 621–632.
Furuta, E., Yamamoto, K., Tatebe, D., Watabe, K., Kitayama, T. & Utsumi, R. (2005). *FEBS Lett.* **579**, 2065–2070.
Hancock, L. E. & Perego, M. (2004). *J. Bacteriol.* **186**, 7951–7958.
Hao, B., Isaza, C., Arndt, J., Soltis, M. & Chan, M. K. (2002). *Biochemistry*, **41**, 12952–12958.
Howell, A., Dubrac, S., Andersen, K. K., Noone, D., Fert, J., Msadek, T. & Devine, K. (2003). *Mol. Microbiol.* **49**, 1639–1655.
Marina, A., Waldburger, C. D. & Hendrickson, W. A. (2005). *EMBO J.* **24**, 4247–4259.

Martin, P. K., Li, T., Sun, D., Biek, D. P. & Schmid, M. B. (1999). *J. Bacteriol.* **181**, 3666–3673.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mohedano, M. L., Overweg, K., de la Fuente, A., Reuter, M., Altabe, S., Mulholland, F., de Mendoza, D., Lopez, P. & Wells, J. M. (2005). *J. Bacteriol.* **187**, 2357–2367.
Ng, W. L., Tsui, H. C. & Winkler, M. E. (2005). *J. Bacteriol.* **187**, 7444–7459.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Qin, Z., Zhang, J., Xu, B., Chen, L., Wu, Y., Yang, X., Shen, X., Molin, S., Danchin, A., Jiang, H. & Qu, D. (2006). *BMC Microbiol.* **6**, 96.
Stock, A. M., Robinson, V. L. & Goudreau, P. N. (2000). *Annu. Rev. Biochem.* **69**, 183–215.
Varughese, K. I. (2002). *Curr. Opin. Microbiol.* **5**, 142–148.
Varughese, K. I., Zhao, H., Veldore, V. H. & Zapf, J. (2007). *Methods Enzymol.* **422**, 102–122.
Watanabe, T., Okada, A., Gotoh, Y. & Utsumi, R. (2008). *Adv. Exp. Med. Biol.* **631**, 229–236.
West, A. H. & Stock, A. M. (2001). *Trends Biochem. Sci.* **26**, 369–376.
Winkler, M. E. & Hoch, J. A. (2008). *J. Bacteriol.* **190**, 2645–2648.
Yamada, S., Akiyama, S., Sugimoto, H., Kumita, H., Ito, K., Fujisawa, T., Nakamura, H. & Shiro, Y. (2006). *J. Mol. Biol.* **362**, 123–139.