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

Preliminary crystallographic studies of the regulatory domain of response regulator YycF from an essential two-component signal transduction system

Haiyan Zhao,[a] Annie Heroux,[b] Reuben D. Sequeira[a] and Liang Tang[a]*

[a]Department of Molecular Biosciences, University of Kansas, 1200 Sunnyside Avenue, Lawrence, KS 66045, USA, and [b]Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA

Correspondence e-mail: tangl@ku.edu

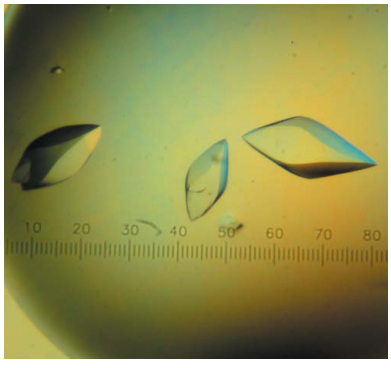
Received 19 February 2009
Accepted 12 June 2009

YycGF is a crucial signal transduction system for the regulation of cell-wall metabolism in low-G+C Gram-positive bacteria, which include many important human pathogens. The response regulator YycF receives signals from its cognate histidine kinase YycG through a phosphotransfer reaction and elicits responses through regulation of gene expression. The N-terminal regulatory domain of YycF from *Bacillus subtilis* was overproduced and purified. The protein was crystallized and X-ray data were collected to 1.95 Å resolution with a completeness of 97.7% and an overall $R_{\rm merge}$ of 7.7%. The crystals belonged to space group $P3_121$, with unit-cell parameters $a = b = 59.50$, $c = 79.06$ Å.

1. Introduction

Prokaryotes and some archaea and eukaryotes utilize a two-component regulatory system (TCS) to sense external or internal signals and implement appropriate responses through regulation of gene expression (West & Stock, 2001; Stock *et al.*, 2000). TCSs play pivotal roles in adaptation, development and pathogenesis. A paradigmatic TCS consists of a histidine kinase (HK) as the sensor and a response regulator (RR) that receives the signal from the HK through phosphotransfer and triggers downstream responses, usually through regulation of the expression of target genes. A typical response regulator consists of a regulatory domain that accepts the phosphoryl group from the cognate histidine kinase and a DNA-binding or effector domain (Galperin, 2006; Gao *et al.*, 2007). YycGF is a two-component regulatory system that is specific to low-G+C Gram-positive bacteria, which include important human pathogens such as *Staphylococcus*, *Streptococcus* and *Enterococcus* (Winkler & Hoch, 2008; Fabret & Hoch, 1998; Dubrac *et al.*, 2008). YycG is a membrane-bound histidine kinase and the response regulator YycF is a transcription factor that recognizes sequence-specific motifs on the bacterial DNA chromosome and regulates the expression of certain genes (Howell *et al.*, 2003; Dubrac & Msadek, 2004). The YycGF system has been recognized as a crucial signal transduction mechanism in the regulation of bacterial cell-wall metabolism and cell division (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). Genes under the putative control of the YycGF system encode important enzymes involved in cell-wall synthesis, such as glucosamidase, muramidase, amidase and endopeptidase (Dubrac *et al.*, 2007). The YycGF system is essential to bacterial viability and has been suggested as a target for the development of novel antibacterial drugs (Watanabe *et al.*, 2008; Qin *et al.*, 2006; Furuta *et al.*, 2005; Winkler & Hoch, 2008).

TCSs are prevalent in bacteria as indicated by genome-sequence analysis (West & Stock, 2001; Stock *et al.*, 2000). They are abundant in most eubacteria, typically constituting ∼1% of encoded proteins. The *Escherichia coli* genome encodes 62 TCS proteins and over 30 TCSs have been identified in *Bacillus subtilis*. These TCSs mediate important processes such as environmental stress responses, osmoregulation, chemotaxis, phototaxis, virulence, and cell growth and differentiation. Specificity is an essential feature in HK–RR recognition in TCS signaling, *i.e.* each RR receives signals only from its



cognate HK to elicit responses appropriately and there are very few examples of crosstalk between different TCSs (Skerker *et al.*, 2008; Gao *et al.*, 2008; Bourret, 2008; Howell *et al.*, 2006). The molecular basis of this high level of specificity remains to be defined by direct structural data. YycF belongs to the OmpR/PhoB subfamily of response regulators. The regulatory domains of most response regulators in this subfamily share a common $(\beta\alpha)_5$ fold, but structural deviations occur near the phosphorylation site, which presumably forms the basis for specific recognition between the cognate pairs of histidine kinases and response regulators (Schnell *et al.*, 2008; Gao *et al.*, 2007). Crystal structures of the N-terminal regulatory domain of the YycF homologue from *Streptococcus* have been described (Bent *et al.*, 2004). The crystal structure of the DNA-binding domain of YycF from *B. subtilis* has been solved (Okajima *et al.*, 2008). However, structural analysis has not been reported for the histidine kinases in this and other YycGF orthologs, preventing a full understanding of HK–RR recognition in the YycGF system. Recently, we crystallized a truncated form of the histidine kinase YycG (Zhao & Tang, 2009). Here, we report preliminary crystallographic characterization of the N-terminal regulatory domain of the response regulator YycF from *B. subtilis*. The protein has been crystallized and X-ray data have been collected to 1.95 Å resolution. This should enable complete structural analysis of this important two-component signal transduction system and particularly of the molecular determinants that define the specificity of histidine kinase and response-regulator recognition.

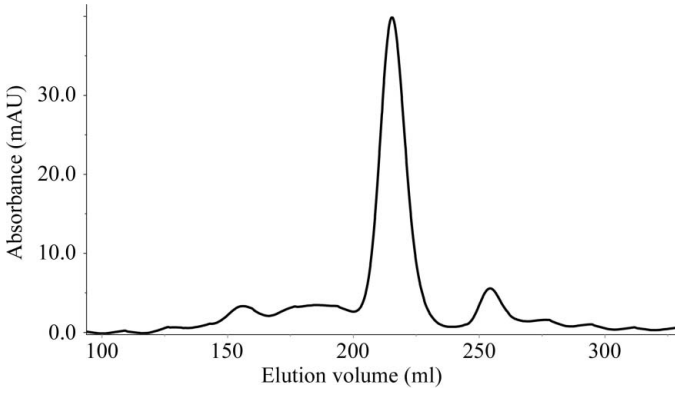


Figure 1
Size-exclusion chromatogram using a Sephacryl 26/60 S-300 column (GE Healthcare) showing a major peak at 215.5 ml corresponding to an approximate molecular mass of ∼17 kDa, consistent with a YycF120 monomer. The column was calibrated using the molecular-weight markers in the MWGF1000 Kit (Sigma) following the procedure suggested by the manufacturer.

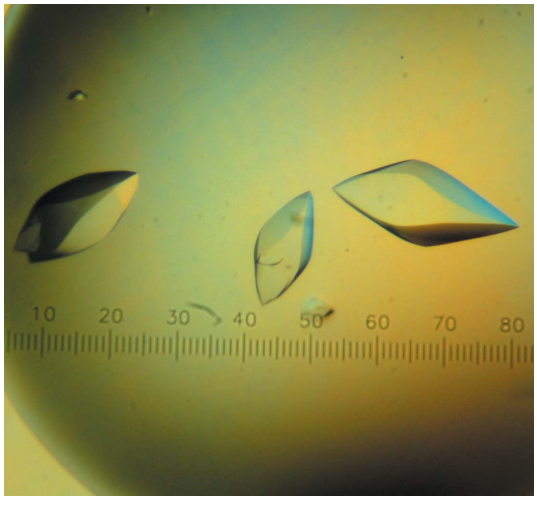


Figure 2
Crystals of the YycF N-terminal regulatory domain. The largest crystal is approximately 0.4 × 0.2 × 0.2 mm. The smallest division of the ruler corresponds to 0.022 mm.

2. Materials and methods

2.1. Protein overproduction and purification

The construct *yycF*120 encoding residues 1–120 of YycF (YycF120) was PCR-amplified from *B. subtilis* strain 168 genomic DNA with the use of the following primers: 5′-CCACCATGGATAAAAAGATC-CTTGTAG-3′, which contains an *Nco*I site followed by the 5′-coding sequence of *yycF*120, and 5′-CTCCTCGAGTTACTGGCGGCGC-AGGTTC-3′, which harbours the 3′-coding sequence of *yycF*120, a stop codon and an *Xho*I site. After digestion of PCR-amplified DNA with *Nco*I and *Xho*I, it was ligated with pET28a (Novagen) plasmid cut with *Nco*I and *Xho*I enzymes. The protein YycF120 was overproduced in BL21 (DE3) *E. coli* cells grown at 303 K in LB medium. Frozen cell pellets were resuspended in 20 m*M* Tris–HCl pH 7.5,

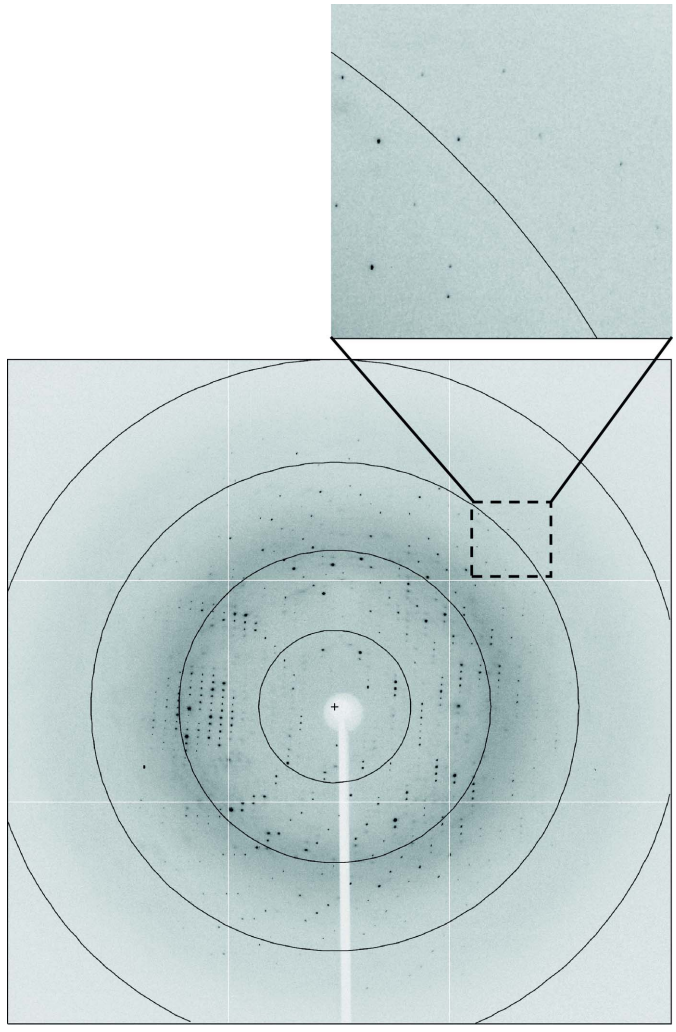


Figure 3
A diffraction pattern of a crystal of the YycF N-terminal regulatory domain. A close-up of an outer region is shown. Rings are shown for 1.7, 2.3, 3.5 and 7.0 Å resolution.

50 m*M* NaCl, 1 m*M* EDTA and 1 m*M* DTT (buffer *A*), sonicated and centrifuged to remove cell debris. The supernatant was loaded onto a HiTrap Q FF (GE Healthcare) column, washed and eluted with buffer *A* plus 500 m*M* NaCl (buffer *B*). The fractions containing YycF120 were pooled and dialyzed against 20 m*M* sodium phosphate pH 6.2, 50 m*M* NaCl, 1 m*M* EDTA and 1 m*M* DTT (buffer *C*). This preparation was loaded onto a MonoQ column (1 ml; GE Healthcare) and eluted with buffer *C* plus 500 m*M* NaCl (buffer *D*). The major fractions containing YycF120 were then loaded onto a size-exclusion chromatography column (Sephacryl S-300, GE Healthcare) and eluted with 20 m*M* Tris–HCl pH 8.5, 150 m*M* NaCl, 1 m*M* EDTA and 1 m*M* DTT. The fractions containing YycF120 were combined and concentrated with Amicon Ultra-4 10 000 MWCO (Millipore). The final preparation of YycF120 was concentrated to 15 mg ml^{-1} for subsequent crystallization trials. Electrospray ionization mass spectrometry showed a molecular mass of 13 765.5 Da, which is in agreement with the calculated mass. Size-exclusion chromatography showed that the protein existed as a monomer in solution (Fig. 1).

2.2. Crystallization

Initial crystallization screening was performed using the Wizard I, II and III random sparse-matrix crystallization screen kits commercially available from Emerald BioStructures. The conditions were optimized and the optimal crystals were grown at 293 K by the hanging-drop vapour-diffusion method. 1 µl protein solution at 15 mg ml^{-1} was mixed with an equal volume of reservoir solution containing 0.1 *M* Tris–HCl pH 8.5, 12% glycerol and 1.5 *M* ammonium sulfate and was equilibrated against 500 µl reservoir solution. Crystals grew to maximal dimensions of approximately 0.4 × 0.2 × 0.2 mm within four weeks (Fig. 2).

2.3. Data collection and processing

Crystals were flash-cooled in liquid nitrogen using 25% glycerol as a cryoprotectant. X-ray data were collected at a cryogenic temperature of 100 K on beamline X29 at the National Synchrotron Light Source (NSLS, New York) using an ADSC Quantum-315r CCD detector. A total of 200 frames were collected with an oscillation range of 1° and an exposure time of 0.5 s per frame (Fig. 3). The diffraction data were processed with *HKL*-2000 (Otwinowski & Minor, 1997). Data-collection and processing statistics are summarized in Table 1.

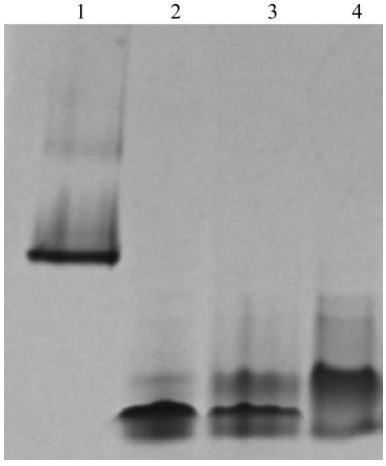


Figure 4
Native polyacrylamide gel electrophoresis of purified phosphorylated and unphosphorylated YycF120. Lane 1, bovine serum albumin; lane 4, unphosphorylated YycF120; lane 2, phosphorylated YycF120; lane 3, 10 µl unphosphorylated YycF120 at 1 mg ml^{-1} in 50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 10 m*M* $MgCl_2$, 1 m*M* DTT and 0.1 m*M* EDTA was mixed with 3 µl 100 m*M* lithium carbamoyl phosphate, a small-molecule phosphorylating agent (Zapf *et al.*, 1996), and incubated for 2 h at room temperature. The unphosphorylated YycF120 was nearly fully converted to the phosphorylated form.

Table 1
Statistics of X-ray data collection and processing.

Values in parentheses are for the outermost resolution shell.

Beamline	NSLS X29
Wavelength (Å)	1.0000
Resolution (Å)	50–1.95 (2.02–1.95)
No. of measurements	124916
No. of unique reflections	11994
Completeness (%)	97.7 (95.0)
$I/\sigma(I)$	57.2 (4.0)
$R_{\rm merge}$† (%)	7.7 (46.9)
Space group	$P3_121$
Unit-cell parameters (Å)	$a = b = 59.50$, $c = 79.06$

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

3. Results and discussion

The N-terminal regulatory domain of YycF encompassing residues 1–120 was purified using three chromatographic columns. We observed the presence of both unphosphorylated and phosphorylated versions of YycF120 on overproduction in *E. coli*, although the unphosphorylated protein dominated (Fig. 4). The use of a MonoQ column was essential for separation of the two versions and successful crystallization of the purified unphosphorylated YycF120. This also enabled purification of the phosphorylated YycF120, which may allow comparative structural studies to shed light on plausible conformational changes induced by phosphorylation as a mechanism for YycF activation.

Crystals of YycF120 belonged to space group $P3_121$, with unit-cell parameters $a = b = 59.50$, $c = 79.06$ Å. The data set contained a total of 11 994 unique reflections in the resolution range 50.0–1.95 Å, with a completeness of 97.7% and an $R_{\rm merge}$ of 7.7%. The data were analyzed with the program *PHENIX* (Adams *et al.*, 2002) and no twinning was detected. The crystal contained one molecule in the crystallographic asymmetric unit, with 58% solvent content and a $V_{\rm M}$ value of 2.9 Å^3 Da^{-1} (Matthews, 1968). The structure has been determined using the molecular-replacement method and deposited in the RCSB Protein Data Bank with accession code 3f6p. Details of the structure will be reported elsewhere.

The X-ray data for this study were measured on beamline X29 of the National Synchrotron Light Source. Financial support came principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the US Department of Energy and from the National Center for Research Resources of the National Institutes of Health. We thank 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

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Bent, C. J., Isaacs, N. W., Mitchell, T. J. & Riboldi-Tunnicliffe, A. (2004). *J. Bacteriol.* **186**, 2872–2879.

Bourret, R. B. (2008). *Mol. Microbiol.* **69**, 1336–1340.
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.
Galperin, M. Y. (2006). *J. Bacteriol.* **188**, 4169–4182.
Gao, R., Mack, T. R. & Stock, A. M. (2007). *Trends Biochem. Sci.* **32**, 225–234.
Gao, R., Tao, Y. & Stock, A. M. (2008). *Mol. Microbiol.* **69**, 1358–1372.
Howell, A., Dubrac, S., Andersen, K. K., Noone, D., Fert, J., Msadek, T. & Devine, K. (2003). *Mol. Microbiol.* **49**, 1639–1655.
Howell, A., Dubrac, S., Noone, D., Varughese, K. I. & Devine, K. (2006). *Mol. Microbiol.* **59**, 1199–1215.
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.
Okajima, T., Doi, A., Okada, A., Gotoh, Y., Tanizawa, K. & Utsumi, R. (2008). *FEBS Lett.* **582**, 3434–3438.
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.
Schnell, R., Ågren, D. & Schneider, G. (2008). *Acta Cryst.* F**64**, 1096–1100.
Skerker, J. M., Perchuk, B. S., Siryaporn, A., Lubin, E. A., Ashenberg, O., Goulian, M. & Laub, M. T. (2008). *Cell*, **133**, 1043–1054.
Stock, A. M., Robinson, V. L. & Goudreau, P. N. (2000). *Annu. Rev. Biochem.* **69**, 183–215.
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.
Zapf, J. W., Hoch, J. A. & Whiteley, J. M. (1996). *Biochemistry*, **35**, 2926–2933.
Zhao, H. & Tang, L. (2009). *Acta Cryst.*, F**65**, 346–349.