Purification, crystallization and preliminary X-ray analysis of a putative nucleotide phosphohydrolase, YpgQ, from *Bacillus subtilis*

Ye Ji Jeon,^{a,b} Wan Seok Song^{a,b}
and Sung-il Yoon^{a,b*}

^aDepartment of Systems Immunology, College of Biomedical Science, Kangwon National University, Chuncheon 200-701, Republic of Korea, and ^bInstitute of Bioscience and Biotechnology, Kangwon National University, Chuncheon 200-701, Republic of Korea

Correspondence e-mail: sungil@kangwon.ac.kr

Received 20 November 2013

Accepted 26 March 2014

The histidine-aspartate (HD) domain exerts phosphohydrolase activity on nucleotides and functions in nucleotide metabolism. Sequence analysis suggested that YpgQ from *Bacillus subtilis* contains the HD domain, but the structure and function of YpgQ remain to be revealed. The recombinant YpgQ protein was overexpressed in an *Escherichia coli* cell expression system and was purified to homogeneity by Ni-NTA affinity and anion-exchange chromatography. Crystals in space group $P2_1$ were obtained in PEG 600 solutions and diffracted X-rays to 2.3 Å resolution. Moreover, X-ray fluorescence scans on YpgQ crystals demonstrated the metal-binding ability of YpgQ.

1. Introduction

The histidine-aspartate (HD) domain is found as a part of multi-domain proteins or as standalone proteins in diverse organisms of archaea, bacteria and eukaryotes (Aravind & Koonin, 1998). The HD domain generally exhibits phosphohydrolase activity for nucleotides and nucleic acids, and plays a key role in nucleotide metabolism and signal transduction. To exert its enzymatic activity, the HD domain has been shown to require a metallic cofactor that is coordinated by conserved histidine and aspartate residues (H...HD...D).

HD-domain-containing enzymes exhibit different specificities for substrates and metals to catalyze distinct reactions of nucleotide hydrolysis. For example, the HD domain of RelA/SpoT in complex with Mn^{2+} exhibits 3' hydrolysis activity on (p)ppGpp, an intracellular signalling alarmone, to generate GDP (GTP) and pyrophosphate (Hogg *et al.*, 2004), whereas *Escherichia coli* YfbR drives 5' hydrolysis of dNMP into a nucleoside and a phosphate in the presence of Co^{2+} (Zimmerman *et al.*, 2008). dNTP triphosphohydrolases that contain the HD domain hydrolyze dNTP into a nucleoside and a triphosphate (Vorontsov *et al.*, 2011; Kondo *et al.*, 2004, 2007). Additionally, the HD domain of GlnD, a bifunctional uridylyltransferase/uridylyl-removing enzyme, in enterobacteria was demonstrated to remove the UMP moiety from uridylylated PII proteins such as GlnB and GlnK (Zhang *et al.*, 2010). Such examples indicate that HD enzymes are responsible for diverse biological functions although they are grouped into the same superfamily by sequence conservation. Thus, functional annotation for each HD enzyme requires a careful biochemical and structural analysis that defines its substrate specificity and catalytic activity.

The genome of *Bacillus subtilis* encodes a YpgQ protein that is conserved mainly among Gram-positive bacteria. PSI-BLAST searches using the amino-acid sequence of YpgQ indicate that YpgQ contains the canonical histidine and aspartate residues of the HD domain, suggesting that YpgQ would exert phosphohydrolase activity on nucleotides. However, the biological function and enzymatic activity of YpgQ have not been characterized and YpgQ was classified in a cluster of unknown function in the HD superfamily (Aravind & Koonin, 1998). Moreover, structural studies on YpgQ have not been reported to date, although unpublished structures of YpgQ homologues (with a highest sequence identity of ~39%) are available in the Protein Data Bank. Here, we report the expression, purification, crystallization and X-ray diffraction analyses of YpgQ as the first step in the study of the structure and function of YpgQ.

Furthermore, the metal-binding ability of YpgQ was demonstrated by X-ray fluorescence scattering of YpgQ crystals.

2. Materials and methods

2.1. Cloning of the YpgQ expression vector

The expression vector for YpgQ (UniProt E0U0I6; residues 1–205) was constructed by cloning the YpgQ gene into a modified pET49b vector (pET49bm) that contains an N-terminal His₆ tag and thrombin cleavage site (Song & Yoon, 2014; Hong *et al.*, 2012). The YpgQ gene was amplified by PCR using *B. subtilis* genomic DNA as template with a forward primer (5'-TAAGGATCCGATGACGGAATTGAA-GCAGGCAG-3'; *Bam*HI site in bold) and a reverse primer (5'-GCCGATGTCGACTTACGACGTTTTGGCATCAATCC-3'; *Sal*I site in bold). The PCR products and the pET49bm vector DNAs were digested with the restriction enzymes *Bam*HI and *Sal*I and were ligated by T4 DNA ligase. The resultant ligation products were transformed into *E. coli* strain DH5 α and positive transformants were selected on an LB agar plate that contained 50 μ g ml⁻¹ kanamycin. Insertion of the YpgQ gene into the vector backbone was confirmed by restriction-enzyme digestion and DNA sequencing.

2.2. Expression of YpgQ

YpgQ-pET49bm vector DNA was transformed into *E. coli* strain BL21 (DE3) cells and the cells were grown on a kanamycin-containing LB agar plate. Cells from a single colony were further grown at 37°C in LB broth containing 50 μ g ml⁻¹ kanamycin. Protein expression was induced at an OD₆₀₀ (optical density at 600 nm) of 0.6–1.0 using 1 mM IPTG and continued overnight at 18°C. Cells were harvested by centrifugation at 5000g.

2.3. Purification of YpgQ

The cells in the pellet were resuspended in a lysis buffer consisting of 50 mM Tris pH 8.0, 200 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF) and lysed by sonication. The cell lysate was centrifuged at ~25 000g and the resulting supernatant was incubated with Ni-NTA agarose resin (Qiagen) for 2 h in the presence of 10 mM

Table 1

Data-collection statistics for *B. subtilis* YpgQ.

Values in parentheses are for the highest resolution shell.

X-ray source	Beamline 5C, PAL
Wavelength (Å)	0.9796
Space group	<i>P</i> 2 ₁
Unit-cell parameters (Å, °)	<i>a</i> = 47.42, <i>b</i> = 53.03, <i>c</i> = 82.86, α = 90.00, β = 93.74, γ = 90.00
Resolution (Å)	20.00–2.30 (2.38–2.30)
No. of observations	52288
No. of unique reflections	17271
Multiplicity	3.0 (3.0)
Completeness (%)	94.2 (96.8)
<i>R</i> _{merge} † (%)	7.9 (42.5)
$\langle I/\sigma(I) \rangle$	24.1 (3.8)

$$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}.$$

imidazole. The resin was washed with a washing buffer consisting of 50 mM Tris pH 8.0, 200 mM NaCl containing 10 mM imidazole. His₆-tagged YpgQ protein was eluted using 250 mM imidazole in 50 mM Tris pH 8.0, 200 mM NaCl and was dialyzed against 20 mM HEPES pH 7.4, 150 mM NaCl. The N-terminal His₆ tag was removed using 1 U of thrombin per ~200 μ g of YpgQ protein (3 h, 18°C). The YpgQ protein was further purified by anion-exchange chromatography using a Mono Q 10/100 GL column (GE Healthcare). The anion-exchange resin was washed with 20 mM HEPES pH 7.4 and YpgQ eluted as a single peak in an NaCl gradient.

2.4. Crystallization of YpgQ

Crystallization conditions for the YpgQ protein were screened using The JCSG Core Suites screening kit (Qiagen) by the sitting-drop vapour-diffusion method at 18°C. Crystals were initially observed in 40% (w/v) PEG 600, 0.1 M sodium citrate pH 5.5. The initial crystallization condition was optimized and the best crystals were obtained in a drop consisting of 0.5 μ l 10 mg ml⁻¹ protein solution and 0.5 μ l 38% (w/v) PEG 600, 0.1 M MES pH 6.5 at 18°C.

2.5. X-ray data collection and analysis

YpgQ crystals were harvested in a nylon loop from the crystallization drop and directly flash-cooled at –173°C under a cryo-stream without the addition of other cryoprotectants. A total of 180 X-ray diffraction images were collected with 1° oscillation and 0.5 s expo-

Figure 1
SDS-PAGE analysis of purified YpgQ protein (lane 2) with molecular-mass markers (lane 1; labelled in kDa). The calculated molecular mass of YpgQ is ~23.6 kDa.

Figure 2
YpgQ crystals in a crystallization drop.

Figure 3
An X-ray diffraction image of the YpgQ crystal. A resolution ring is shown at 2.3 Å resolution.

sure per image using an ADSC Quantum 315r detector on beamline 5C of the Pohang Accelerator Laboratory, Republic of Korea. The crystal-to-detector distance was set to 260 mm and the wavelength was set to 0.9796 Å. Diffraction data were reduced and scaled using the *HKL-2000* package (Otwinowski & Minor, 1997).

3. Results and discussion

B. subtilis YpgQ protein was expressed in an *E. coli* expression system and purified to at least 95% purity by Ni-NTA affinity and anion-exchange chromatography (Fig. 1). The final yield of YpgQ was ~30 mg from a 1 l culture. YpgQ crystals were obtained in PEG 600 solutions at pH 6.5 by the sitting-drop vapour-diffusion method (Fig. 2) and yielded X-ray diffraction data to 2.3 Å resolution (Fig. 3 and Table 1). The YpgQ crystal belonged to space group $P2_1$, with unit-cell parameters $a = 47.42$, $b = 53.03$, $c = 82.86$ Å, $\beta = 93.74^\circ$. The asymmetric unit potentially contains two YpgQ molecules according to Matthews coefficient probability analysis (Matthews coefficient V_M of $2.2 \text{ Å}^3 \text{ Da}^{-1}$, solvent content of 44.2%; Matthews, 1968; Kantardjieff & Rupp, 2003). The dimeric assembly of YpgQ is consistent with the noncrystallographic dimer structure of a YpgQ homologue (PDB entry 2qgs; Northeast Structural Genomics Consortium, unpublished work). The YpgQ structure, which was determined by experimental phasing using selenomethionine-substituted protein, will be described elsewhere.

Figure 4
Fluorescence scan around the *K* edge of nickel using the YpgQ crystal.

To examine the presence of bound metal ions in YpgQ, we obtained an X-ray fluorescence scattering scan for the YpgQ crystal. A strong fluorescence peak was observed at ~8348 eV near the *K* absorption edge (8333 eV) of nickel, suggesting that the purified YpgQ protein contains nickel ions that may have been captured either during expression or Ni-NTA affinity chromatography (Fig. 4). Although it is unclear whether YpgQ requires nickel ions for enzymatic activity, the data indicate that YpgQ has the ability to bind metals.

X-ray diffraction data sets were collected on the Pohang Accelerator Laboratory beamline 5C, and we thank beamline staff members. This study was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2012R1A1A1003701 to S-IY) and by a 2012 Research Grant from Kangwon National University (to S-IY).

References

- Aravind, L. & Koonin, E. V. (1998). *Trends Biochem. Sci.* **23**, 469–472.
- Hogg, T., Mechold, U., Malke, H., Cashel, M. & Hilgenfeld, R. (2004). *Cell*, **117**, 57–68.
- Hong, M., Yoon, S.-I. & Wilson, I. A. (2012). *Biochem. Biophys. Res. Commun.* **427**, 119–124.
- Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
- Kondo, N., Kuramitsu, S. & Masui, R. (2004). *J. Biochem.* **136**, 221–231.
- Kondo, N., Nakagawa, N., Ebihara, A., Chen, L., Liu, Z.-J., Wang, B.-C., Yokoyama, S., Kuramitsu, S. & Masui, R. (2007). *Acta Cryst. D* **63**, 230–239.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Song, W. S. & Yoon, S.-I. (2014). *Biochem. Biophys. Res. Commun.* **444**, 109–115.
- Vorontsov, I. I., Minasov, G., Kiryukhina, O., Brunzelle, J. S., Shuvalova, L. & Anderson, W. F. (2011). *J. Biol. Chem.* **286**, 33158–33166.
- Zhang, Y., Pohlmann, E. L., Serate, J., Conrad, M. C. & Roberts, G. P. (2010). *J. Bacteriol.* **192**, 2711–2721.
- Zimmerman, M. D., Proudfoot, M., Yakunin, A. & Minor, W. (2008). *J. Mol. Biol.* **378**, 215–226.