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

Crystallization and preliminary crystallographic analysis of the second GAF domain of DevS from *Mycobacterium smegmatis*

Ha Yeon Cho,[a] Hyo Je Cho,[a] Young Min Kim,[b] Jeong Il Oh[c] and Beom Sik Kang[a]*

[a]School of Life Science and Biotechnology, Kyungpook National University, Daegu 702-701, Republic of Korea, [b]Department of Biology, Yonsei University, Seoul 120-749, Republic of Korea, and [c]Department of Microbiology, Pusan National University, Pusan 609-735, Republic of Korea

Correspondence e-mail: bskang2@knu.ac.kr

Received 15 December 2007
Accepted 23 February 2008

Mycobacterium tuberculosis is known to transform into the nonreplicating persistence state under the influence of hypoxia or nitric oxide. DevS-DevR is a two-component regulatory system that mediates the genetic response for the transformation. DevS is a histidine kinase that contains two GAF domains for sensing hypoxia or nitric oxide. The second GAF from *M. smegmatis* DevS was crystallized using the sitting-drop vapour-diffusion method in the presence of sodium citrate and 2-propanol as precipitants. X-ray diffraction data were collected from crystals containing selenomethionine to a maximum resolution of 2.0 Å on a synchrotron beamline. The crystals belong to the hexagonal space group $P6_1$. The asymmetric unit contains one molecule, corresponding to a packing density of 2.5 Å^3 Da^{-1}. The selenium substructure was determined by the single anomalous dispersion method and structure refinement is in progress.

1. Introduction

Mycobacterium tuberculosis is still one of the most dreaded pathogens and one of the reasons for its success as a pathogen lies in its ability to persist for years within the host. One-third of the world's population is estimated to carry *M. tuberculosis* in the dormant form (Parrish *et al.*, 1998). In this state, the sensitivity of *M. tuberculosis* to most available medications is diminished. *M. tuberculosis* has been shown to undergo a metabolic transformation to the nonreplicating persistence state under the influence of environmental stimuli such as hypoxia or nitric oxide (Wayne & Sohaskey, 2001).

DevS-DevR is a two-component regulatory system that mediates the genetic response to oxygen limitation and nitric oxide exposure in mycobacteria (Sherman *et al.*, 2001). DevR is a transcriptional regulator and DevS is a histidine kinase that undergoes autophosphorylation in response to environmental change and subsequently transduces the signal to the cognate response regulator DevR. DevS has two GAF domains at its N-terminus and a kinase domain is at its C-terminus (Saini *et al.*, 2004). The sensing by DevS is presumably carried out through the GAF domains.

GAF domains are small-molecule-binding domains and are named after cyclic **G**MP-regulated cyclic nucleotide phosphodiesterases (PDEs), *Anabaena* **a**denylyl cyclase and bacterial transcription factor **F**hlA (Aravind & Ponting, 1997). These domains are found in many proteins from various organisms and play important roles as regulatory elements. In mammals, GAF domains are mostly found in cyclic nucleotide PDEs and control cyclic GMP and cyclic AMP second-messenger levels. Five of 11 mammalian PDEs contain two GAF domains in tandem that have two separate functions: binding a cyclic nucleotide and dimerization (Soderling *et al.*, 1999; Fawcett *et al.*, 2000). A crystal structure of PDE2A GAF domains showed that the second GAF (GAF-B) binds a cyclic nucleotide, whereas the first GAF (GAF-A) mediates protein dimerization (Martinez *et al.*, 2002). Adenylyl cyclases from the cyanobacterium *Anabaena* contain tandem GAF domains in their N-termini (Katayama & Ohmori, 1997). The structure of adenylyl cyclase cyaB2 revealed that both GAF-A and GAF-B domains bound cyclic AMP and were involved in the dimerization (Martinez *et al.*, 2005).

Two GAF domains are also found in tandem in *M. tuberculosis* DevS. GAF-A (63–210) of DevS contains a haem and His149 is the proximal haem ligand (Sardiwal *et al.*, 2005). Thus, the DevS sensor might be controlled through the binding of molecular oxygen or nitric oxide by the haem. However, the role of GAF-B (231–379) of DevS is still not known. It might be involved in binding to a cyclic nucleotide or in dimerization. A previous analysis of the DevS sequence suggested the presence of two putative transmembrane helices in GAF-B (Dasgupta *et al.*, 2000). This GAF domain could also be responsible for NO sensing without haem. Transcription factor NorR in *Escherichia coli* has an NO-responsive activity and senses NO *via* a nonhaem iron centre in the GAF domain of NorR (Autreaux *et al.*, 2005). We are investigating the molecular mechanism of DevS from *M. smegmatis*, a nonpathogenic cousin of *M. tuberculosis*. Structural analysis of GAF-B would provide insights into the role of the domain in histidine kinase activation upon hypoxia and NO. In this paper, the crystallization and initial X-ray crystallographic analysis of GAF-B from *M. smegmatis* DevS are described.

2. Expression and purification

The gene encoding the second GAF domain (Asp232–Asp380) of DevS from *M. smegmatis* was amplified by the polymerase chain reaction (PCR) using the primer set 5′-CATGCCATGGACCCG-GCG-3′ and 5′-GGAATTCAGTCGGCCAGG-3′. The primers carried *Nco*I and *Eco*RI restriction-enzyme sites at their 5′ ends, respectively. The PCR product was cloned into the pGST-parallel vector (Sheffield *et al.*, 1999), a GST-fusion protein expression vector containing a recombinant TEV protease (rTEV) cleavage site. The integrity of the insert was verified by direct DNA sequencing. The expression of SeMet-labelled GAF-B was induced by 1 m*M* IPTG in *E. coli* strain B834 (Novagen) with M9 medium containing 50 µg ml^{-1} SeMet (Fig. 1). The expressed proteins were purified by affinity chromatography using Glutathione-Sepharose 4B (GE Healthcare). The recombinant protein was digested using rTEV (Invitrogen) at 283 K in the presence of 0.5 m*M* EDTA and 1 m*M* DTT. After complete digestion, the GST tag was removed using a Glutathione-Sepharose 4B column. Gel filtration was performed with a Superdex G75 column (GE Healthcare) equilibrated with 20 m*M* Tris–HCl pH 7.5 and the fractions containing GAF-B were collected and concentrated using Centriprep YM10 (Millipore) for crystallization screening. 15 mg protein was obtained from a 4 l culture, corresponding to a yield of 3.8 mg per litre of culture. The homogeneity of GAF-B was assessed by 15% SDS–PAGE and Coomassie Brilliant Blue staining (Fig. 1). The purified protein contains an additional three amino acids (GAM) at the N-terminus arising from the cloning procedure.

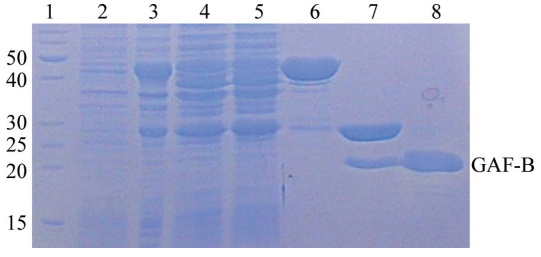

Figure 1
Expression and purification of the GAF-B domain of DevS from *M. smegmatis*. Proteins were subjected to SDS–PAGE at each purification step. Lane 1, protein size markers (kDa); lane 2, uninduced cell lysate; lane 3, induced cell lysate; lane 4, total cell lysate; lane 5, soluble fraction; lane 6, GST-fusion GAF-B from Glutathione-Sepharose column; lane 7, rTEV protease-treated protein; lane 8, purifed GAF-B.

During the gel-filtration procedure, GAF-B eluted as a single peak and the size of the protein was estimated as 32 kDa after calibration of the column. Considering the molecular weight of GAF-B (15.5 kDa), this indicates that GAF-B forms a dimer in solution and could play a role in the dimerization of DevS. Although two segments were predicted to be transmembrane helices by sequence analysis, the solubility of GAF-B suggests that these segments are parts of a soluble domain.

3. Crystallization

Crystallization of the purified protein was initially performed with commercially available sparse-matrix screens from Hampton Research and Emerald Biostructures using the sitting-drop vapour-diffusion method at 294 K. Each experiment consisted of mixing 1 µl protein solution (9 mg ml^{-1} in 20 m*M* Tris–HCl pH 7.5) with 1 µl reservoir solution and then equilibrating it against reservoir solution. Needle-shaped crystals were obtained after 12 h incubation in a drop containing 1.4 *M* sodium citrate as a precipitant. The addition of 2-propanol improved the quality of the crystals in a screening with Additive Screen (Hampton Research). During the optimization procedure, the drop size and the concentrations of protein and precipitants were adjusted. The best crystals were obtained from a drop made by mixing 2 µl protein solution and 0.5 µl reservoir solution containing 1.25 *M* sodium citrate and 0.6% 2-propanol in 0.1 *M* Tris–HCl pH 9.0 (Fig. 2).

4. Diffraction data analysis and substructure determination

The crystals were transferred to a cryoprotectant solution containing 1.25 *M* sodium citrate, 0.6%(*v*/*v*) 2-propanol and 18%(*v*/*v*) glycerol, fished out with a loop that was larger than the crystals and flash-frozen by immersion in liquid nitrogen at 100 K. After a fluorescence scan, single anomalous X-ray dispersion (SAD) data were collected at a wavelength corresponding to the Se absorption peak (0.9794 Å) using a Quantum 210 CCD detector (ASDC, USA) on beamline 4A (MXW) at the Pohang Accelerator Laboratory (PAL, Pohang, Republic of Korea). The crystal-to-detector distance was set to 250 mm. A total of 360 images were collected with an oscillation

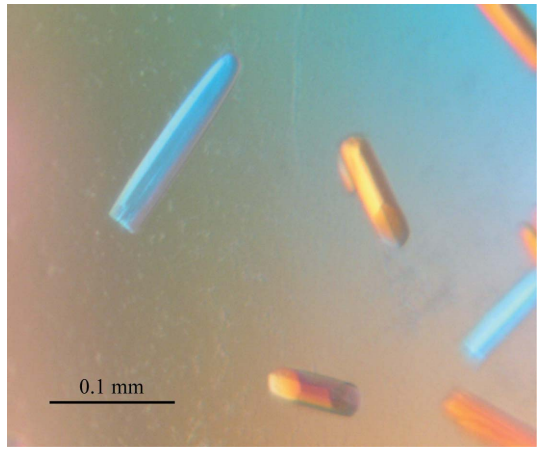

Figure 2
Hexagonal crystals of the GAF-B domain of DevS from *M. smegmatis*. The best quality crystals reached maximal dimensions in 3 d.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

Crystal	1	2
X-ray source	4A (MXW), PAL	4A (MXW), PAL
Wavelength (Å)	0.9794	1.0000
Crystal-to-detector distance (mm)	250	180
Rotation range per image (°)	1.0	1.0
Total rotation range (°)	360	120
Exposure time (s)	5	20
Space group	$P6_1$	$P6_1$
Unit-cell parameters (Å)	$a = b = 73.3$, $c = 49.6$	$a = b = 73.2$, $c = 48.5$
Resolution limit (Å)	50–2.5 (2.56–2.47)	40–2.0 (2.07–2.00)
Total reflections	121830	72234
Unique reflections	5616 (529)	10102 (968)
Redundancy	21.7 (18.3)	7.2 (6.5)
Completeness (%)	99.6 (95.7)	99.5 (97.1)
$R_{\rm merge}$†	0.134 (0.371)	0.082 (0.376)
$R_{\rm p.i.m.}$†	0.026 (0.045)	0.034 (0.076)
Average $I/\sigma(I)$	65.4 (14.6)	31.7 (5.3)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$ and $R_{\rm p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith observation, $\langle I(hkl)\rangle$ is the mean intensity of the reflections and N is the redundancy.

angle of 1° and an exposure time of 5 s per image. The data were indexed, integrated and scaled using the *HKL*-2000 package (Otwinowski & Minor, 1997). The GAF-B crystal belongs to space group $P6_1$, with unit-cell parameters $a = b = 73.3$, $c = 49.6$ Å. An additional data set was collected from another crystal to a resolution of 2.0 Å at a wavelength of 1.0000 Å, for which higher photon flux could be obtained at the beamline. 120° of data were collected with an exposure time of 20 s per image and a crystal-to-detector distance of 180 mm. The unit-cell parameters of the crystal are $a = b = 73.2$, $c = 48.5$ Å. Crystallographic data statistics are summarized in Table 1.

The selenium substructure was determined using *SOLVE* (Terwilliger & Berendzen, 1999). Using the first data set, three selenium sites were found in the asymmetric unit. Assuming that the asymmetric unit contains one molecule of GAF-B, the value of the Matthews coefficient is 2.49 Å^3 Da^{-1}, corresponding to a solvent content of 50.6%. The overall figure of merit (FOM) was 0.449. Density modification and subsequent automated model building were performed with *RESOLVE* (Terwilliger, 2000), increasing the FOM to 0.729 with 74% (113 of 152 amino acids) of the residues built. Further model building was performed manually into the density-modified electron-density map using the programs *O* (Jones *et al.*, 1991) and *Coot* (Emsley & Cowtan, 2004) and refinement with isotropic displacement parameters was performed with *REFMAC*5 (Murshudov *et al.*, 1997) within the *CCP*4 suite. The initial model was used as a template for molecular replacement with *AMoRe* (Navaza, 1994) with the 2.0 Å resolution data. Crystallographic model building and refinement of the structure at 2.0 Å resolution are in progress.

This work was supported by a Korea Science and Engineering Foundation (KOSEF) grant from the Korean government (MOST; R01-2006-000-10846-0).

References

Aravind, L. & Ponting, C. P. (1997). *Trends Biochem. Sci.* **22**, 458–459.
Autreaux, B. D., Tucker, N. P., Dixon, R. & Spiro, S. (2005). *Nature (London)*, **437**, 769–772.
Dasgupta, N., Kapur, V., Singh, K., Das, T. K., Sachdeva, S., Jyothisri, K. & Tyagi, J. S. (2000). *Tuber. Lung Dis.* **80**, 141–159.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fawcett, L., Baxendale, R., Stacey, P., McGrouther, C., Harrow, I., Soderling, S., Hetman, J., Beavo, J. A. & Phillips, S. C. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 3702–3707.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Katayama, M. & Ohmori, M. (1997). *J. Bacteriol.* **179**, 3588–3593.
Martinez, S. E., Bruder, S., Schultz, A., Zheng, N., Schultz, J. E., Beavo, J. A. & Linder, J. U. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 3082–3087.
Martinez, S. E., Wu, A., Glavas, N., Tang, X., Turley, S., Hol, W. & Beavo, J. A. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 13260–13265.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Parrish, N. M., Dick, J. D. & Dishai, W. R. (1998). *Trends Microbiol.* **6**, 107–112.
Saini, D. K., Malhotra, V. & Tyagi, J. S. (2004). *FEBS Lett.* **565**, 75–80.
Sardiwal, S., Kendall, S. L., Movahedzadeh, F., Rison, S. C. G., Stoker, N. G. & Djordjevic, S. (2005). *J. Mol. Biol.* **353**, 929–936.
Sheffield, P., Garrard, S. & Derewenda, Z. (1999). *Protein Expr. Purif.* **15**, 34–39.
Sherman, D. R., Voskuil, M., Schnappinger, D., Liao, R., Harrell, M. I. & Schoolnik, G. K. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 7534–7539.
Soderling, S. H., Bayuga, S. J. & Beavo, J. A. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 7071–7076.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Wayne, L. G. & Sohaskey, C. D. (2001). *Annu. Rev. Microbiol.* **55**, 139–163.