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

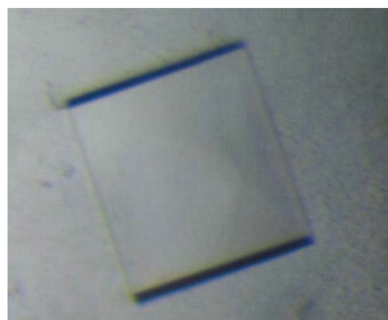
Crystallization and preliminary X-ray diffraction characterization of an essential protein from *Xanthomonas campestris* that contains a noncanonical PilZ signature motif yet is critical for pathogenicity

Tso-Ning Li,[a] Ko-Hsin Chin,[a,b] Hui-Ling Shih,[c] Andrew H.-J. Wang[c,d] and Shan-Ho Chou[a,b]*

[a]Institute of Biochemistry, National Chung-Hsing University, Taichung 40227, Taiwan, [b]National Chung-Hsing University Biotechnology Center, National Chung-Hsing University, Taichung 40227, Taiwan, [c]Core Facility for Protein Crystallography, Academia Sinica, Nankang, Taipei, Taiwan, and [d]Institute of Biological Chemistry, Academia Sinica, Nankang, Taipei, Taiwan

Correspondence e-mail: shchou@nchu.edu.tw

Received 28 July 2009
Accepted 8 September 2009



© 2009 International Union of Crystallography
All rights reserved

Recent studies have identified c-di-GMP as a novel secondary messenger molecule that is heavily involved in regulating bacterial biofilm formation, motility, production of pathogenicity factors *etc.* PilZ domain-containing proteins have been suggested and subsequently proved to be the c-di-GMP receptor. However, considering the diverse biological functions exhibited by c-di-GMP, it may be that receptors other than the PilZ domain exist. An essential protein from the plant pathogen *Xanthomonas campestris* pv. *campestris* (Xcc) that contains a noncanonical PilZ signature motif yet is critical for Xcc pathogenicity has been cloned, purified and crystallized. Detailed characterization of this protein may reveal an alternative binding mode of c-di-GMP and allow a more thorough understanding of how c-di-GMP exhibits its diverse effects.

1. Introduction

Recent studies have identified c-di-GMP as a universal secondary messenger molecule that is heavily involved in regulating bacterial pathogenicity (Romling *et al.*, 2005; Amikam & Galperin, 2006; Jenal & Malone, 2006; Ryan *et al.*, 2006; Benach *et al.*, 2007; Pratt *et al.*, 2007; Ramelot *et al.*, 2007; Hickman & Harwood, 2008). The PilZ domain (Alm *et al.*, 1996) was first identified as the receptor for the important secondary messenger c-di-GMP *via* a bioinformatics study (Amikam & Galperin, 2006). Indeed, many PilZ domain-containing receptor proteins from different bacteria have been found and bind c-di-GMP with variable affinities ranging from submicromolar to micromolar. These include PlzD (Benach *et al.*, 2007) and VCA0042 from *Vibrio cholerae* (Pratt *et al.*, 2007), YcgR from *Escherichia coli* and *Gluconacetobacter xylinus* (Ryjenkov *et al.*, 2006), DgrA from *Caulobacter crescentus* (Christen *et al.*, 2005) and PA4608 (Ramelot *et al.*, 2007) and Alg44 from *Pseudomonas aeruginosa* (Merighi *et al.*, 2007). In these proteins, several highly conserved residues for binding c-di-GMP were found, namely the R*XXX*R and the D/N*X*S*XX*G motifs. However, considering the highly diverse functions exhibited by c-di-GMP, it is possible that receptors with different c-di-GMP-binding motifs other than the PilZ domain may exist. Indeed, the PelD protein from *P. aeruginosa* that mediates the production of polysaccharide biosynthesis has been shown to be a novel c-di-GMP receptor without the classical c-di-GMP-binding motif (Lee *et al.*, 2007), as is another protein FleQ also from *P. aeruginosa* that regulates the genes for flagella biosynthesis (Hickman & Harwood, 2008). PA2960 from *P. aeruginosa* was first annotated as a PilZ-domain protein (Merighi *et al.*, 2007) and is required for type IV pilus-mediated twitching motility (Alm *et al.*, 1996; Mattick, 2002), yet lacks a c-di-GMP switch loop (Benach *et al.*, 2007). A similar situation also occurs in XC1028 from *Xanthomonas campestris* pv. *campestris*, (Xcc), which was found to adopt a five-stranded β-barrel core similar to other canonical PilZ domains but exhibits considerable differences in the N-terminal region (Li *et al.*, 2009). Clearly, further structural and functional variants of the c-di-GMP-binding proteins will be revealed in due course.

The *X. campestris* pv. *campestris* 8004 genome contains four PilZ domain-containing proteins (Qian *et al.*, 2005; McCarthy *et al.*, 2008).

One of these proteins, XC2249, was found to be crucial in causing the pathogenicity of Xcc and mutations of XC2249 were found to reduce its virulence in Chinese radish, causing a reduction in mobility, and to affect extracellular enzyme production (McCarthy *et al.*, 2008). In this manuscript, we have cloned XC6102 from *X. campestris* pv. *campestris* 17 (the equivalent of XC2249 from *X. campestris* pv. *campestris* 8004; 100% sequence identity) and overexpressed it in *E. coli*. The final protein was purified and crystallized in a square-shaped form. Since it also lacks the canonical c-di-GMP-binding motif, its crystal structure should reveal interesting insights into how an alternative PilZ domain functions.

2. Materials and methods

2.1. Cloning, expression and purification

The XC6012 gene fragment was PCR-amplified directly from a local Xcc genome (*X. campestris* pv. *campestris* strain 17) with a forward 5′-*TACTTCCAATCCAATGCT*ATGTCCACGCTCGGCACGCT primer and a reverse 5′-*TTATCCACTTCCAATG*TCAGCGCTGGCGACGGGCG primer (the linker sequences are italicized). A ligation-independent cloning (LIC) approach (Aslanidis & de Jong, 1990) was carried out to obtain the desired construct according to a previously published protocol (Wu *et al.*, 2005). The final construct codes for an N-terminal His_6 tag, a 17-amino-acid linker and the XC6012 target protein under the control of a T7 promoter. The vector was transformed into *E. coli* BL21 (DE3) host cells, which were grown at 310 K in LB medium until an OD_{600} of 0.8 was attained. Overexpression of the His_6-tagged target protein was induced by the addition of 0.5 m*M* IPTG at 293 K for 20 h. The cells were harvested, resuspended in lysis buffer (20 m*M* Tris–HCl pH 8.0, 80 m*M* NaCl) and lysed using a microfluidizer (Microfluidics). Most of the tagged target protein was present in the soluble fraction (Fig. 1). After centrifugation, the target protein was purified by immobilized metal-affinity chromatography (IMAC) on a nickel column (Sigma) and was eluted with a gradient of 50–300 m*M* imidazole in lysis buffer. The fractions containing XC6012 were monitored by SDS–PAGE, recombined and dialyzed repeatedly against lysis buffer. After buffer-exchange and concentration, the His_6 tag and linker were cleaved from XC6012 by TEV (tobacco etch virus) protease at 295 K for 16 h and removed by immobilized metal-affinity chromatography (IMAC) on a nickel column (Sigma). For crystallization, XC6012 was further purified by FPLC (ÄKTA, Pharmacia Inc.) on a Superdex 200 gel-filtration column equilibrated with lysis buffer. The final target protein was greater than 99% pure, with only an extra tripeptide (SNA) at the N-terminal end. SeMet XC6012 was prepared in a similar way except that the cells were induced in SeMet-containing M9 minimum medium when an OD_{600} of 0.8 was reached. The overexpression and purification of SeMet XC6012 was monitored by SDS–PAGE as shown in Fig. 1.

2.2. Crystallization

For crystallization, the protein was concentrated to 6.5 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0 and 80 m*M* NaCl using an Amicon Ultra-10 (Millipore). Screening for crystallization conditions was performed by sitting-drop vapour diffusion in 96-well plates (Hampton Research) at 295 K by mixing 0.5 µl protein solution with 0.5 µl reagent solution. Initial screens including Emerald BioSystems Wizard I and II random sparse-matrix crystallization screens, Hampton Research sparse-matrix Crystal Screens 1 and 2, a systematic PEG–pH screen and the PEG/Ion Screen were performed using a Gilson C240 crystallization workstation. Square-shaped crystals appeared from condition E8 of the Wizard II random sparse-matrix crystallization screen comprising 0.1 *M* sodium/potassium phosphate pH 6.2, 0.2 *M* NaCl, 10%(*w*/*v*) PEG 8K in 3 d. This initial condition was then optimized by varying the concentrations of NaCl and PEG 8K. Crystals suitable for diffraction experiments were obtained in 0.1 *M* sodium/potassium phosphate pH 6.2, 0.25 *M* NaCl, 12%(*w*/*v*) PEG 8K using the hanging-drop vapour-diffusion method and reached maximum dimensions of 0.2 × 0.2 × 0.1 mm in two weeks (Fig. 2).

2.3. Data collection

The crystals were soaked in cryoprotectant solution comprising the reservoir solution plus 25%(*v*/*v*) glycerol and were then flash-cooled at 100 K in a stream of cold nitrogen. X-ray diffraction data were collected on National Synchrotron Radiation Research Center (NSRRC; Taiwan) beamline 13B1 using a Q315 area detector. A three-wavelength multiple-wavelength anomalous diffraction (MAD) data set was obtained to a resolution of 2.1 Å with 2 s exposure time, 0.5° oscillation angle and 300 mm crystal-to-detector distance. The data were indexed and integrated using the *HKL*-2000 processing

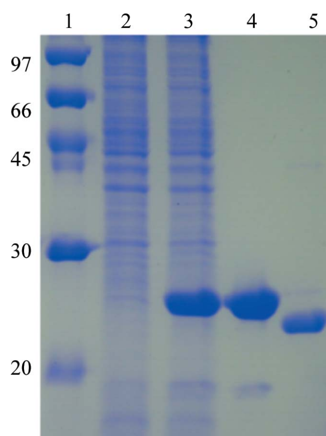


Figure 1
SDS–PAGE monitoring of the overexpression and purification of SeMet XC6012. Lane 1, protein molecular-weight markers (kDa); lane 2, whole cell lysate before IPTG induction; lane 3, whole cell lysate after IPTG induction; lane 4, supernatant of His_6-tagged XC6012; lane 5, purified XC6012 after TEV cleavage.

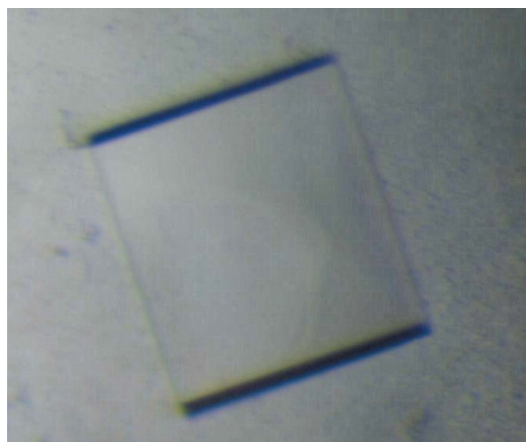


Figure 2
Crystals of SeMet XC6012 from *X. campestris* grown in 0.1 *M* sodium/potassium phosphate pH 6.2, 0.25 *M* sodium chloride, 12% PEG 8K using the hanging-drop vapour-diffusion method at room temperature. The average dimensions of these crystals reached 0.2 × 0.2 × 0.1 mm in two weeks.

software (Otwinowski & Minor, 1997), generating data that were 99.4% complete with an overall R_{merge} of 3.9–5.6% on intensities. The crystals belonged to the monoclinic space group $P2_1$. There are four XC6012 molecules in the asymmetric unit. The Matthews coefficient and solvent content (Matthews, 1968) of the crystals are 2.67 $Å^3$ Da^{-1} and 54.04%, respectively. The data-collection statistics are summarized in Table 1 and an X-ray diffraction image is shown in Fig. 3. The refinement of Se-atom positions, phase calculation and density modification were carried out using the program *SnB* v.2.3 (Rappleye *et al.*, 2002).

3. Results and discussion

c-di-GMP is synthesized from two moles of GTP by a class of enzymes (diquanylate cyclases) containing GGDEF domains and is hydrolyzed by another enzyme family (cyclic phosphodiesterases) containing an EAL or HD-GYP domain (Jenal & Malone, 2006). c-di-GMP has also been found to bind at an allosteric inhibitor site of GGDEF-domain proteins for tight regulation of the intracellular c-di-GMP concentration (Chan *et al.*, 2004; De *et al.*, 2008). Sometimes such feedback regulation involves the formation of several distinct oligomeric states, switching from an active dimer to a product-inhibited dimer *via* a tetrameric assembly, as happens in the diquanylate cyclase WspR (De *et al.*, 2008).

Although c-di-GMP appears to be a conserved signal, the outputs it regulates are quite diverse and vary between different systems. Currently, several different classes of c-di-GMP-binding receptors have been characterized. The first and most common class are the PilZ-domain proteins, which contain conserved R*XXX*R and D/N*X*S*XX*G motifs (Alm *et al.*, 1996; Amikam & Galperin, 2006; Ryjenkov *et al.*, 2006; Benach *et al.*, 2007; Merighi *et al.*, 2007; Pratt *et al.*, 2007). The second class share the amino-acid motif R*XX*D, as characterized by PelD from *P. aeruginosa* (Lee *et al.*, 2007). Lastly, a third class is represented by the transcription factor FleQ from *P. aeruginosa* (Lee *et al.*, 2007) and XC1028 from *X. campestris* (Li *et al.*, 2009), which contain none of the conserved motifs described above. In the first class of PilZ-domain proteins, binding of c-di-GMP was found to induce significant conformational changes in the receptor (Benach *et al.*, 2007). Obviously, further structural and functional studies of c-di-GMP receptors are necessary in order to fully understand the diverse functions that they exhibit (Galperin, 2004; Jenal & Malone, 2006).

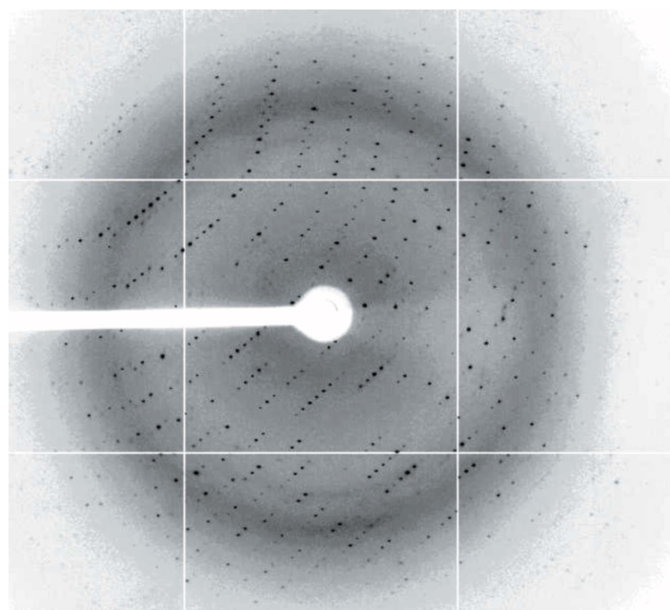


Figure 3
The diffraction pattern of SeMet XC6012 collected from a flash-frozen crystal at the Taiwan synchrotron facility (NSRRC 13B1). The exposure time was 2 s, the oscillation range was 0.5° per frame and the crystal-to-detector distance was 300 mm.

Table 1
Summary of the Se-MAD crystallographic data of XC6012.

Values in parentheses are for the outermost shell.

	Inflection	High remote	Peak
Beamline	NSRRC BL13B1		
Wavelength (Å)	0.97910	0.97622	0.97884
Space group	$P2_1$	$P2_1$	$P2_1$
Unit-cell parameters (Å, °)	a = 58.84, b = 89.37, c = 87.35, β = 105.4	a = 58.84, b = 89.37, c = 87.35, β = 105.4	a = 58.84, b = 89.37, c = 87.35, β = 105.4
Resolution range (Å)	30–2.11 (2.19–2.11)	30–2.1 (2.18–2.1)	30–2.2 (2.28–2.2)
Unique observations	49875 (4940)	50523 (5020)	44136 (4415)
Redundancy	3.8 (3.4)	3.9 (3.4)	7.7 (7.0)
Completeness (%)	99.4 (98.6)	99.5 (99.0)	99.7 (99.8)
R_{merge} (%)	5.6 (43.8)	3.9 (45.4)	5.5 (43.2)
$I/\sigma(I)$	20.0 (2.8)	21.2 (3.4)	32.8 (4.5)

XC2249 from the *X. campestris* pv. *campestris* 8004 genome was found to be a PilZ-domain protein that is heavily involved in the virulence of Xcc. In this manuscript, we have successfully cloned XC6012 from Xcc strain 17 (the equivalent of XC2249 from Xcc strain 8004) and overexpressed it in *E. coli*. Interestingly, it was found to adopt a stable tetramer, as characterized by gel-filtration chromatography and by analytical ultracentrifugal chromatography (data not shown). As far as we know, this is the first report of a PilZ-domain receptor protein existing in a tetrameric state. Further structural and functional studies of XC6012 may highlight a new direction towards the alternative c-di-GMP-binding mode. As we have obtained high-resolution X-ray diffraction data from SeMet-substituted XC6012 crystals (there are two methionines in the XC6012), the structure should be solvable by the MAD approach (Terwilliger & Berendzen, 1999). Indeed, a preliminary tetrameric structure of XC6012 has been obtained and we are currently refining its tertiary structure.

This work was supported by an Academic Excellence Pursuit grant from the Ministry of Education and by the National Science Council, Taiwan (grant 97-2113-M005-005-MY3) to S-HC. We appreciate the service of Structural Genomics Databases provided by the GMBD Bioinformatics Core (http://www.tbi.org.tw), NRPGM, Taiwan. We would also like to thank the Core Facilities for Protein X-ray Crystallography in the Academia Sinica, Taiwan for help in crystal screening, the National Synchrotron Radiation Research Center (NSRRC) in Taiwan and the SPring-8 Synchrotron facility in Japan for assistance in X-ray data collection. The National Synchrotron Radiation Research Center is a user facility supported by the National Science Council, Taiwan and the Protein Crystallography Facility is supported by the National Research Program for Genomic Medicine, Taiwan.

References

Alm, R. A., Bodero, A. J., Free, P. D. & Mattick, J. S. (1996). *J. Bacteriol.* **178**, 46–53.
Amikam, D. & Galperin, M. Y. (2006). *Bioinformatics*, **22**, 3–6.
Aslanidis, C. & de Jong, P. J. (1990). *Nucleic Acids Res.* **18**, 6069–6074.

Benach, J., Swaminathan, S. S., Ramayo, R., Handelman, S. K., Folta-Stogniew, E., Ramos, J. E., Forouhar, F., Neely, H., Seetharaman, J., Camilli, A. & Hunt, J. F. (2007). *EMBO J.* **26**, 5153–5166.
Chan, C., Paul, R., Samoray, D., Amiot, N. C., Giese, B., Jenal, U. & Schirmer, T. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 17084–17089.
Christen, M., Christen, B., Folcher, M., Schauerte, A. & Jenal, U. (2005). *J. Biol. Chem.* **280**, 30829–30837.
De, N., Pirruccello, M., Krasteva, P. V., Bae, N., Raghavan, R. V. & Sondermann, H. (2008). *PloS Biol.* **6**, e67.
Galperin, M. Y. (2004). *Environ. Microbiol.* **6**, 552–567.
Hickman, J. W. & Harwood, C. S. (2008). *Mol. Microbiol.* **69**, 376–389.
Jenal, U. & Malone, J. (2006). *Annu. Rev. Genet.* **40**, 385–407.
Lee, V. T., Matewish, J. M., Kessler, J. L., Hyodo, M., Hayakawa, Y. & Lory, S. (2007). *Mol. Microbiol.* **65**, 1474–1484.
Li, T.-N., Chin, K.-H., Liu, J.-H., Wang, A. H.-J. & Chou, S.-H. (2009). *Proteins*, **75**, 282–288.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mattick, J. S. (2002). *Annu. Rev. Microbiol.* **56**, 289–314.
McCarthy, Y., Ryan, R. P., O'Donovan, K., He, Y.-Q., Jiang, B.-L., Feng, J.-X., Tang, J.-L. & Dow, J. M. (2008). *Mol. Plant Pathol.* **9**, 819–824.
Merighi, M., Lee, V. T., Hyodo, M., Hayakawa, Y. & Lory, S. (2007). *Mol. Microbiol.* **65**, 876–895.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pratt, J. T., Tamayo, R., Tischler, A. D. & Camilli, A. (2007). *J. Biol. Chem.* **282**, 12860–12870.
Qian, W. *et al.* (2005). *Genome Res.* **15**, 757–767.
Ramelot, T. A., Yee, A., Cort, J. R., Semesi, A., Arrowsmith, C. H. & Kennedy, M. A. (2007). *Proteins*, **66**, 266–271.
Rappleye, J., Innus, M., Weeks, C. M. & Miller, R. (2002). *J. Appl. Cryst.* **35**, 374–376.
Romling, U., Gomelsky, M. & Galperin, M. Y. (2005). *Mol. Microbiol.* **57**, 629–639.
Ryan, R. P., Fouhy, Y., Lucey, J. F., Crossman, L. C., Spiro, S., He, Y.-W., Zhang, L.-H., Heeb, S., Camara, M., Williams, P. & Dow, J. M. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 6712–6717.
Ryjenkov, D. A., Simm, R., Romling, U. & Gomelsky, M. (2006). *J. Biol. Chem.* **281**, 30310–30314.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Wu, Y.-Y., Chin, K.-H., Chou, C.-C., Lee, C.-C., Shr, H.-L., Gao, F. P., Lyu, P.-C., Wang, A. H.-J. & Chou, S.-H. (2005). *Acta Cryst.* F**61**, 902–905.