Crystallization of the N-terminal regulatory domain of the enhancer-binding protein FleQ from *Stenotrophomonas maltophilia*

Jauo-Guey Yang,^{a,†} Min-Shao Shih,^{a,†} Wei-Ting Kuo,^a Ko-Hsin Chin,^b Gwan-Han Shen^c and Shan-Ho Chou^{a,b,*}

^aInstitute of Biochemistry, National Chung Hsing University, Taichung 40227, Taiwan,

^bAgricultural Biotechnology Center, National Chung Hsing University, Taichung 40227, Taiwan, and ^cDepartment of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan

† These authors made equal contributions to this article and can be considered co-first authors.

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

Received 12 November 2013

Accepted 21 January 2014

FleQ is a master regulator that controls bacterial flagellar gene expression. It is a unique enhancer-binding protein or repressor protein comprising an N-terminal FleQ domain, an AAA⁺/ATPase σ 54-interaction domain and a helix–turn–helix DNA-binding domain. FleN is a putative ATPase with a deviant Walker A motif that works together with FleQ by binding to the FleQ N-terminal domain to fully express *pel*, *psl* and *cdr* operons in the presence of c-di-GMP to enhance biofilm formation. *Stenotrophomonas maltophilia* is an emerging human pathogen that causes fatal infections in humans. In order to understand the interaction between the FleN and FleQ domains and its effect on *S. maltophilia* biofilm formation, determination of the FleQ–c-di-GMP and FleN–FleQ–c-di-GMP complex structures was embarked upon. Towards this goal, the FleQ N-terminal domain from *S. maltophilia* was first cloned and expressed in *Escherichia coli*. Native and SeMet-labelled FleQ domains were successfully crystallized and diffracted to resolutions of 2.08 and 2.58 Å, respectively.

1. Introduction

Stenotrophomonas maltophilia is an aerobic, non-fermentative Gram-negative bacterium that is widespread in soil and the clinical environment (Ryan *et al.*, 2009). Although it is phylogenetically closely related to the genera *Xanthomonas* and *Xylella*, it is not phytopathogenic. In contrast, it is an emerging human pathogen that causes fatal infections in humans, especially in the debilitated and immunocompromised. This is possibly owing to its intrinsic resistance to many antibiotics and its capability to form biofilms, which make the selection of an optimal therapy difficult (Crossman *et al.*, 2008). Intriguingly, this bacterium can also tolerate high levels of extremely toxic metals, such as Cd, Pb, Co, Zn, Hg, Ag, selenite, tellurite and uranyl from 0.1 to 50 mM concentration (Pages *et al.*, 2008). Thus, it is also useful for environmental bioremediation.

c-di-GMP is a newly discovered second-messenger molecule that controls a plethora of bacterial activities ranging from biofilm formation and biogenesis of flagella and pili to the secretion of pathogenic factors *etc.* (Hengge, 2009; Römling *et al.*, 2005; Schirmer & Jenal, 2009). It is synthesized by the GGDEF-domain-containing diguanylate cyclases and is degraded by EAL-domain-containing or HD-GYP-domain-containing phosphodiesterases. It has also been found to interact with a wide variety of protein-based or RNA-based receptors. Some representative examples include transcriptional factor Clp (Chin *et al.*, 2010), degenerate GGDEF or EAL domains (Navarro *et al.*, 2009), PilZ domain proteins (Amikam & Galperin, 2006; Benach *et al.*, 2007; Chin *et al.*, 2012) and riboswitches (Kulshina *et al.*, 2009; Smith *et al.*, 2009). Because of its potential roles in a wide variety of bacterial functions, the search for novel c-di-GMP effectors and their functions is still ongoing (Römling, 2011; Sondermann *et al.*, 2011; Ryan *et al.*, 2012; Krasteva *et al.*, 2012; Kalia *et al.*, 2013; Gomelsky, 2011).

Prokaryotic signal transduction is mainly triggered by two-component systems (TCSs) comprising a histidine kinase domain and a response-regulator domain *via* a phosphorelay mechanism (West & Stock, 2001). However, phosphotransfer is not the only signal transduction mechanism. In addition to ligand binding, signal transduction is also mediated in some cases by protein–protein interaction rather than by covalent modification, such as that for the NifL–NifA

pair (Barrett *et al.*, 2011). FleQ is an enhancer-binding protein comprising an N-terminal FleQ domain, an AAA⁺/ATPase σ 54-interaction domain and a helix–turn–helix DNA-binding domain. Yet it lacks both a typical phosphoacceptor domain with a highly conserved Asp residue and a histidine kinase gene in the same operon, suggesting that FleQ activation and regulation are possibly not achieved *via* the typical phosphorelay mechanism (Arora *et al.*, 1997). Interestingly, it binds to two FleQ boxes on the *pelA* promoter very well in the absence of FleN or c-di-GMP to repress *pelA* expression (Hickman & Harwood, 2008). In the presence of ATP and FleN, which is a putative ATPase with a deviant Walker A motif, a distortion of the *pel* promoter was observed, and FleN was proposed to form a bridge between the two FleQ proteins bound to DNA-binding sites to cause DNA bending. In addition, further c-di-GMP binding at the AAA⁺/ATPase σ 54-interaction domain was found to relieve this DNA distortion and activate *pelA* expression instead (Baraquet *et al.*, 2012; Baraquet & Harwood, 2013). FleQ is thus unusual in that it can convert from a repressor to an activator depending on the presence of c-di-GMP. Yet, how such a conversion occurs has remained unclear to date, and structure determination of

FleQ and its binding complexes with FleN, ATP, c-di-GMP and/or the *pelA* promoter is necessary to clarify this distinct mechanism. In the first step towards this goal, we have successfully cloned the N-terminal FleQ domain from *S. maltophilia* and expressed it in *Escherichia coli*. Native and SeMet-labelled proteins were later crystallized and diffracted to resolutions of 2.08 and 2.58 Å, respectively.

2. Materials and methods

2.1. Cloning, expression and purification of Sm2620^{1–143}

The coding regions corresponding to Sm2620^{1–131} and Sm2620^{1–143} were PCR-amplified directly from the *S. maltophilia* genome using the forward primer 5'-TACTTCCAATCCAATGCTGTGAGC-GAATCGCGCATCCTTGTGCTG-3' and the reverse primers 5'-CGGGCCGTTGCCGGTCCGGTC-3' and 5'-TTATCCACTTCC-AATGACCGGCCTGGTGCTCGGCATCC-3'. The PCR products exhibited correct sizes in agarose gel electrophoresis and were further confirmed by DNA sequencing. A ligation-independent cloning

Figure 1

Sequence alignment of the PaFleQ and SmFleQ proteins. Both proteins comprise a FleQ domain (marked by a red bar above the sequence), an AAA⁺ σ -interacting domain (marked by a green bar) and a HTH DNA-binding domain (marked by a blue bar). Identical residues are highlighted in orange and similar residues in light blue. The phosphoacceptor residues DD-D-K in the typical REC domain were changed to DD-S⁵⁹-M in PaFleQ and DN-S⁵⁶-T in SmFleQ (highlighted in red). The ATP-binding site is highlighted in green and the DNA-binding HTH helical motif in brown.

(LIC) approach was used to obtain the desired constructs (Wu *et al.*, 2005). The final construct codes for an N-terminal His₁₂ tag, a 17-amino-acid linker and the *Sm*2620^{1–131} (or *Sm*2620^{1–143}) targets under the control of a T7 promoter. Overexpression of the His₁₂-tagged target protein was induced by the addition of 0.5 mM IPTG to the medium at 16°C for 20 h. The target proteins were purified by immobilized metal-affinity chromatography (IMAC) on a nickel column (Sigma). The His₁₂ tag and linker were then cleaved from *Sm*2620^{1–131} (or *Sm*2620^{1–143}) by *Tobacco etch virus* (TEV) protease at 4°C for 48 h. For crystallization, *Sm*2620^{1–131} and *Sm*2620^{1–143} proteins were further purified on a Superdex 75 column (ÄKTA, Pharmacia). The final target proteins exhibited a purity greater than 99%. SeMet-labelled *Sm*2620^{1–143} was prepared in a similar way and was produced using *E. coli* strain BL21 (DE3) as the host in the absence of methionine but with ample amounts of selenomethionine (100 mg l^{–1}) in M9 medium. Induction was conducted at 16°C for 20 h by the addition of 0.5 mM IPTG. Purification of the SeMet-labelled *Sm*2620^{1–143} protein was carried out using the same protocols as those established for the native protein.

2.2. Crystallization of *Sm*2620^{1–131} and Se-Met-labelled *Sm*2620^{1–143} domains

For crystallization, the *Sm*2620^{1–131} and SeMet-labelled *Sm*2620^{1–143} proteins were concentrated to 15 mg ml^{–1} in 40 mM Tris–HCl pH 8.0, 500 mM NaCl using an Amicon Ultra-10 (Millipore). Screening for crystallization conditions was performed by

Table 1
Summary of the X-ray crystallographic data for the SeMet-labelled *Sm*FleQ and native *Sm*FleQ targets.

Values in parentheses are for the outermost shell.

	SeMet-labelled <i>Sm</i> FleQ	Native <i>Sm</i> FleQ
Beamline	BL13B1	BL13C1
Wavelength (Å)	0.97925	0.97622
Space group	<i>P</i> 3 ₁ 2 ₁	<i>P</i> 2 ₁
Unit-cell parameters (Å, °)	<i>a</i> = <i>b</i> = 63.025, <i>c</i> = 124.506, α = β = 90, γ = 120	<i>a</i> = 29.321, <i>b</i> = 30.790, <i>c</i> = 56.946, α = 90, β = 98.755, γ = 90
Resolution range (Å)	30–2.58 (2.67–2.58)	30–2.08 (2.15–2.08)
Total observations	82857 (3929)	20686 (1570)
Unique observations	17425 (893)	6088 (561)
Multiplicity	4.8 (4.4)	3.4 (2.8)
Completeness (%)	99.9 (99.7)	98.2 (92.3)
<i>R</i> _{merge} [†] (%)	7.0 (33.6)	4.6 (16.7)
<i>I</i> / σ (<i>I</i>)	23.7 (3.3)	23.8 (5.4)
Matthews coefficient (Å ³ Da ^{–1})	2.22	1.81
Solvent content (%)	44.6	32.3

[†] *R*_{merge} = $\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$.

using the sitting-drop vapour-diffusion method in 96-well plates (Hampton Research) at 4°C by mixing 0.5 µl protein solution with 0.5 µl reagent solution. Initial screens including Clear Strategy Screen 1, Structure Screens 1 and 2 (Molecular Dimensions), a systematic PEG–pH screen and a PEG/ion screen were carried out using a Phoenix RE crystallization workstation (Rigaku).

Figure 2
SDS-PAGE purification of the native *Sm*FleQ^{1–131} and SeMet-labelled *Sm*FleQ^{1–143} domains. (a) The domain architecture and constructs used in these studies. Soluble domains are shown as red bars and insoluble domains as grey bars. (b) SDS-PAGE (13%) monitoring of the overexpression and purification of the *Sm*FleQ^{1–131} domain. Lane M, protein marker (labelled in kDa); lane 1, whole cell lysate before IPTG induction; lane 2, whole cell lysate after IPTG induction; lane 3, nickel-column-purified domain after TEV cleavage. (c) Similar SDS-PAGE monitoring of the SeMet-labelled *Sm*FleQ^{1–143} domain.

Rod-like crystals of $Sm2620^{1-131}$ appeared from a reservoir solution comprising 0.1 M trisodium citrate, 2.4 M ammonium sulfate in 6 d. Diamond-shaped crystals of SeMet-labelled $Sm2620^{1-143}$ appeared from a reservoir solution comprising 1.4 M trisodium citrate pH 7.5, 0.1 M Na HEPES in one week.

2.3. Data collection and refinement

Crystals were flash-cooled at -173°C under a stream of cold nitrogen. X-ray diffraction data were collected on beamlines 13C1 and 13B1 at the National Synchrotron Radiation Research Center (NSRRC), Taiwan. The diffraction of the $Sm2620^{1-131}$ crystals reached a resolution of 2.08 Å and they belonged to space group $P2_1$, with unit-cell parameters $a = 29.321$, $b = 30.790$, $c = 56.945$ Å, $\alpha = 90$, $\beta = 98.755$, $\gamma = 90^{\circ}$. The data for the SeMet-labelled $Sm2620^{1-143}$ crystals were collected in a similar way to a resolution of 2.58 Å. The detailed X-ray diffraction data statistics are listed in Table 1.

3. Results and discussion

S. maltophilia is an emerging human pathogen that causes fatal infections in humans. Thus, understanding its pathogenic mechanism and biofilm formation is important in human health. FleQ has been found to serve as an activator of σ_{54} for several σ_{54} -dependent flagellar genes and has been well studied in *Pseudomonas aeruginosa* (*PaFleQ*; Baraquet & Harwood, 2013; Baraquet *et al.*, 2012; Arora *et al.*, 1997) and its close relative *Xanthomonas campestris* pv. *campestris*

(Hu *et al.*, 2005). The homologue of FleQ in *S. maltophilia* had not been identified. By using the FleQ gene sequence from *P. aeruginosa* in a BLAST search against the NCBI website, an analogous FleQ gene was found in *S. maltophilia*. Fig. 1 shows the sequence alignment of *SmFleQ* with *PaFleQ*. From the figure, it is clear that *SmFleQ* adopts an identical structural organization comprising three domains,

Figure 3 Crystal pictures of the FleQ domains. (a) Rod-like crystals of $Sm2620^{1-131}$ from a reservoir solution consisting of 0.1 M trisodium citrate, 2.4 M ammonium sulfate in 6 d. (b) Diamond-like crystals of SeMet- $Sm2620^{1-143}$ from a reservoir solution consisting of 1.4 M trisodium citrate pH 7.5, 0.1 M Na HEPES in one week.

Figure 4 X-ray diffraction data for $Sm2620^{1-131}$ (a) and SeMet- $Sm2620^{1-143}$ (b). Diffraction patterns were collected on a MAR CCD detector on the BL13C1 and BL13B1 beamlines at the NSRRC in Taiwan with exposure times of 45 and 8 s, respectively, for the native $Sm2620^{1-131}$ and SeMet-labelled $Sm2620^{1-143}$. The oscillation range is 1° per frame and the crystal-to-detector distances are 300 and 400 mm, respectively. The edge of the detector corresponds to resolutions of 2.58 and 2.08 Å, respectively.

namely the FleQ domain, the AAA⁺ domain and the DNA-binding domain. Similar to the *PaFleQ* domain, the *SmFleQ* domain also does not contain the characteristic Asp59 residue, which is the acceptor amino acid in REC domains for phosphate transfer via the classical phosphorelay mechanism in regular bacterial two-component signal transduction systems. Ser59 is found in *S. maltophilia* instead, indicating that *SmFleQ* is possibly also not activated via the canonical histidine kinase–receiver domain pathway.

Fig. 2 shows the SDS–PAGE purification of the native *SmFleQ*^{1–131} and SeMet-labelled *SmFleQ*^{1–143} domains. Four different constructs for the native *SmFleQ* domain consisting of residues 1–131, 1–143, 4–131 and 4–143 were cloned into *E. coli* for expression, but only *SmFleQ* domains expressed from residues 1–131 and 1–143 were soluble (marked by red bars). The two *SmFleQ* domains were further purified to greater than 99% purity using a His-tag affinity column. The His₁₂ tag and linker were further cleaved by TEV protease and were purified by gel-filtration chromatography. A single band of the expected molecular weight was identified for the native *SmFleQ*^{1–131} (marked by a red arrow in Fig. 2*b*) and SeMet-labelled *SmFleQ*^{1–143} (marked by a red arrow in Fig. 2*c*). The final domains were then subjected to crystal screening. Interestingly, only the native *SmFleQ*^{1–131} and SeMet-labelled *SmFleQ*^{1–143} domains were found to give good crystals that were suitable for X-ray diffraction (Fig. 3). These crystals were then employed to collect X-ray diffraction data at the NSRRC, Taiwan (Fig. 4). The SeMet-labelled *SmFleQ*^{1–143} data were used to determine the phases using a single-wavelength anomalous dispersion (SAD) approach. The structure of more than 80% of the sequence was observed, which will be completed by manually fitting the *SmFleQ*^{1–143} sequence into the electron-density map. The obtained SeMet-labelled *SmFleQ*^{1–143} structure will be refined against the native *SmFleQ*^{1–143} data to obtain the final structure. This project is now under way.

This work was supported in part by the Ministry of Education, Taiwan, ROC under the ATU plan, and by the National Science Council, Taiwan, ROC (grants 102-2113-M-005-006-MY3) to S-HC. We appreciate the Structural Genomics Databases service provided by the GMBD Bioinformatics Core (<http://www.tbi.org.tw>), NRPGM, Taiwan, ROC. We would also like to thank the Core Facilities for Protein X-ray Crystallography in the Academia Sinica, Taiwan, ROC for help in crystal screening and the National Synchrotron Radiation Research Center (NSRRC) in Taiwan and the SPring-8 Synchrotron facility in Japan for assistance with X-ray data collection. The National Synchrotron Radiation Research Center is a user facility supported by the National Science Council, Taiwan, ROC and the

Protein Crystallography Facility is supported by the National Research Program for Genomic Medicine, Taiwan, ROC.

References

- Amikam, D. & Galperin, M. Y. (2006). *Bioinformatics*, **22**, 3–6.
- Arora, S. K., Ritchings, B. W., Almira, E. C., Lory, S. & Ramphal, R. (1997). *J. Bacteriol.* **179**, 5574–5581.
- Baraquet, C. & Harwood, C. S. (2013). *Proc. Natl Acad. Sci. USA*, **110**, 18478–18483.
- Baraquet, C., Murakami, K., Parsek, M. R. & Harwood, C. S. (2012). *Nucleic Acids Res.* **40**, 7207–7218.
- Barrett, J., Ray, P., Sobczyk, A., Little, R. & Dixon, R. (2011). *Mol. Microbiol.* **39**, 480–493.
- Benach, J., Swaminathan, S. S., Tamayo, R., Handelman, S. K., Foltá-Stogniew, E., Ramos, J. E., Forouhar, F., Neely, H., Seetharaman, J., Camilli, A. & Hunt, J. F. (2007). *EMBO J.* **26**, 5153–5166.
- Chin, K.-H., Kuo, W.-T., Yu, Y.-J., Liao, Y.-T., Yang, M.-T. & Chou, S.-H. (2012). *Acta Cryst. D* **68**, 1380–1392.
- Chin, K.-H., Lee, Y.-C., Tu, Z.-L., Chen, C.-H., Tseng, Y.-H., Yang, J.-M., Ryan, R. P., McCarthy, Y., Dow, J. M., Wang, A. H.-J. & Chou, S.-H. (2010). *J. Mol. Biol.* **396**, 646–662.
- Crossman, L. C. *et al.* (2008). *Genome Biol.* **9**, R74.
- Gomelsky, M. (2011). *Mol. Microbiol.* **79**, 562–565.
- Hengge, R. (2009). *Nature Rev. Microbiol.* **7**, 263–273.
- Hickman, J. W. & Harwood, C. S. (2008). *Mol. Microbiol.* **69**, 376–389.
- Hu, R.-M., Yang, T.-C., Yang, S.-H. & Tseng, Y.-H. (2005). *Biochem. Biophys. Res. Commun.* **335**, 1035–1043.
- Kalia, D., Merey, G., Nakayama, S., Zheng, Y., Zhou, J., Luo, Y., Guo, M., Roembke, B. T. & Sintim, H. O. (2013). *Chem. Soc. Rev.* **42**, 305–341.
- Krasteva, P. V., Giglio, K. M. & Sondermann, H. (2012). *Protein Sci.* **21**, 929–948.
- Kulshina, N., Baird, N. J. & Ferré-D'Amaré, A. R. (2009). *Nature Struct. Mol. Biol.* **16**, 1212–1217.
- Navarro, M. V. A. S., De, N., Bae, N., Wang, Q. & Sondermann, H. (2009). *Structure*, **17**, 1104–1116.
- Pages, D., Rose, J., Conrod, S., Cuine, S., Carrier, P., Heulin, T. & Achouak, W. (2008). *PLoS One*, **3**, e1539.
- Römling, U. (2011). *Environ. Microbiol.* **14**, 1817–1829.
- Römling, U., Gomelsky, M. & Galperin, M. Y. (2005). *Mol. Microbiol.* **57**, 629–639.
- Ryan, R. P., Monchy, S., Cardinale, M., Taghavi, S., Crossman, L., Avison, M. B., Berg, G., van der Lelie, D. & Dow, J. M. (2009). *Nature Rev. Microbiol.* **7**, 514–525.
- Ryan, R. P., Tolker-Nielsen, T. & Dow, J. M. (2012). *Trends Microbiol.* **20**, 235–242.
- Schirmer, T. & Jenal, U. (2009). *Nature Rev. Microbiol.* **7**, 724–735.
- Smith, K. D., Lipchick, S. V., Ames, T. D., Wang, J., Breaker, R. R. & Strobell, S. A. (2009). *Nature Struct. Mol. Biol.* **16**, 1218–1223.
- Sondermann, H., Shikuma, N. J. & Yildiz, F. H. (2011). *Curr. Opin. Microbiol.* **15**, 140–146.
- West, A. H. & Stock, A. M. (2001). *Trends Biochem. Sci.* **26**, 369–376.
- 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.