Crystallographic studies of SarV, a global regulator from *Staphylococcus aureus*

Yang Song,^{a,b} Fan Zhang,^{a,b} Xu Li,^{a,b} Jianye Zang^{a,b*} and Xuan Zhang^{a,b*}

^aHefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, Collaborative Innovation Center of Chemistry for Life Science, University of Science and Technology of China, 96 Jinzhai Road, Hefei, Anhui 230026, People's Republic of China, and ^bKey Laboratory of Structural Biology, Chinese Academy of Sciences, Hefei, Anhui 230027, People's Republic of China. *Correspondence e-mail: zangjy@ustc.edu.cn, xuanzbin@ustc.edu.cn

Received 23 April 2015

Accepted 8 June 2015

Edited by B. Hazes, University of Alberta, Canada

Keywords: SarV; global regulator; *Staphylococcus aureus*; autolysis; crystallographic study.

SarV, a member of the SarA protein family, is a global transcriptional regulator which has been reported to be involved in the regulation of autolysis in *Staphylococcus aureus*. In this study, SarV from *S. aureus* was successfully cloned, expressed, purified and crystallized. X-ray diffraction data were collected to 2.10 Å resolution. The crystals of SarV belonged to the monoclinic space group $P2_1$, with unit-cell parameters $a = 36.40$, $b = 119.64$, $c = 66.80$ Å, $\alpha = \gamma = 90$, $\beta = 98.75^\circ$. The Matthews coefficient and the solvent content were estimated to be $2.57 \text{ Å}^3 \text{ Da}^{-1}$ and 52%, respectively, suggesting the presence of four molecules in the asymmetric unit. The results of size-exclusion chromatography (SEC) indicated that *S. aureus* SarV exists as a homodimer in solution. Unfortunately, the structure cannot be solved by molecular replacement because of the low sequence identity of *S. aureus* SarV to known structures. Further phase determination by selenomethionine single-wavelength anomalous dispersion (SAD) and the heavy-atom method is in progress.

1. Introduction

Staphylococcus aureus, a Gram-positive bacterium, is a leading cause of a wide spectrum of human illnesses ranging from superficial infections to life-threatening pneumonia, endocarditis, osteitis, toxic shock syndrome and septicaemia (Crossley *et al.*, 2009). As a result, there is a considerable need to identify and develop novel drug treatment strategies to control nosocomial *S. aureus* infections. Therefore, a better understanding and analysis of staphylococcal molecular pathogenesis is required. The pathogenicity of *S. aureus* is very complicated owing to the coordinated expression of diverse virulence factors in response to the changing host environment, which is known to be controlled by global regulatory elements and helps *S. aureus* to survive during various stages of infection (Lowy, 1998; Cheung *et al.*, 2004). To date, two major families of global regulators in *S. aureus* have been identified: two-component regulatory systems (TCRSs) and SarA homologues (Cheung *et al.*, 2004).

The SarA family can be divided into three subgroups based on sequence alignment and structural analysis: single-domain proteins (*e.g.* SarA, SarR, Rot, SarT, SarV and SarX), two-domain proteins (*e.g.* SarS, SarU and SarY) and MarR homologues (*e.g.* MgrA and SarZ) (Cheung *et al.*, 2008). Phenotypic characterization of some of the SarA family proteins has led to the discovery of distinct phenotypes and complicated transcription-regulation networks. Novel structural elucidations have revealed diverse folding and DNA-recognition features of SarA family members. More recently, the crystal structures of SarR, SarA, SarS, MgrA, SarZ and Rot have been published (Schumacher *et al.*, 2001; Liu *et al.*,

© 2015 International Union of Crystallography

Table 1

Macromolecule-production information.

The restriction sites are underlined and the vector-encoded residues are shown in lower case.

Source organism	<i>S. aureus</i>
DNA source	<i>S. aureus</i>
Forward primer	5'-CGCGGATCCATGAGTAATAAAGTTCAACGTT- TTATAGAAGCAGAAAGG-3'
Reverse primer	5'-CGGCTCGAGGTTTGTAAACGCTTAACACTTCT- ATTTCTTTAGTAATTTATC-3'
Cloning vector	pET-28a(+)
Expression vector	pET-28a(+)
Expression host	<i>E. coli</i> BL21 (DE3)
Complete amino-acid sequence of the construct produced	mssghhhhhhhgssenyqgsMSNKVQRFIEAE- RELSQLKHWLKTTHKISIEEFVLFKVYAEK- ISGKELRDLTHFEMLWDTSKIDVIIRKIYKKE- LISKLRSETDERQVFYFYSTSQKLLDKITKE- IEVLVTN1ehhhhhh

2001; Li *et al.*, 2003; Chen *et al.*, 2006; Poor *et al.*, 2009; Zhu *et al.*, 2014), which have added to our understanding of the generalized mode of gene regulation by these proteins at the molecular level. These reported members of the SarA family of proteins, which belong to the typical winged-helix DNA-binding proteins, are likely to bind to target promoter DNA using a similar motif but may have divergent activation domains.

SarV, a homologue of SarA, is a 116-residue polypeptide that has been reported to be involved in the regulation of some virulence and autolysis genes. SarV shares rather low sequence identity with other SarA family proteins. It might act as an activator for the expression of *agr*, *lytSR* and *arlRS*. The transcription of *sarV* can be repressed by the binding of SarA and MgrA proteins to the *sarV* promoter region, suggesting that the *sarV* gene product may play a regulatory role in cell lysis (Manna *et al.*, 2004). Although SarV has been reported to be involved in the regulation of autolysis in *S. aureus*, functioning as a global transcriptional regulator, the details of the interaction of SarV with various genes remain unclear. Here, we report the purification, crystallization and preliminary

Figure 1

A crystal of *S. aureus* SarV with a dimension of approximately 1.0 mm along the longest axis.

X-ray analysis of SarV from *S. aureus* with the aim of establishing the three-dimensional structure to provide a better understanding of the function of SarV.

2. Materials and methods

2.1. Macromolecule production

The *sarV* gene was amplified from *S. aureus* strain NCTC 8325 by polymerase chain reaction (PCR) using PrimeStar HS DNA Polymerase (Takara) with the desired primers (Table 1). The amplified product was digested with the restriction endonucleases BamHI and XhoI, and inserted into similarly digested pET-28a(+) vector which carries an N-terminal His₈ tag and a C-terminal His₆ tag. The resulting SarV/p28 construct was verified by DNA sequencing, transformed into *Escherichia coli* BL21 (DE3) cells and grown with shaking overnight at 310 K in a 20 ml starter culture of Luria–Bertani (LB) medium containing 50 µg ml⁻¹ kanamycin. The overnight culture was then inoculated into 1 l LB medium containing 50 µg ml⁻¹ kanamycin and incubated at 310 K with shaking until an OD₆₀₀ of 0.6–0.8 was reached. Protein expression was induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.5 mM at 289 K for 20 h. The cells were harvested by centrifugation at 6760g for 8 min. The cell pellet was collected and frozen at 193 K for storage.

The cell pellets were resuspended in buffer A consisting of 50 mM Tris–HCl pH 8.0, 500 mM NaCl, 2 mM β-mercaptoethanol and lysed by sonication on ice. The lysate was centrifuged at 23 800g for 30 min at 277 K and the supernatant was subjected to a column containing 6 ml 50% (v/v) Ni–NTA agarose resin (GE Healthcare, USA) pre-equilibrated with buffer A. The column was washed with 3 × 10 column volumes of buffer A containing 20, 40 and 60 mM imidazole, respectively. The bound protein was eluted with buffer A containing 500 mM imidazole. The eluted sample was concentrated to 2 ml using a Millipore 10 kDa centrifugal device and further purified by size-exclusion chromatography (SEC) on a HiLoad 16/60 Superdex 200 size-exclusion column (GE Healthcare, USA) with buffer B (50 mM Tris–HCl pH 8.0, 500 mM NaCl, 2 mM DTT). All purification steps were performed at a temperature below 283 K. Fractions of the peak corresponding to SarV were pooled, concentrated to 15–25 mg ml⁻¹ and frozen at 193 K for storage. The absorbance measured at 280 nm (*A*₂₈₀) was used to calculate the protein concentration (a concentration of 1 mg ml⁻¹ has an *A*₂₈₀ of 1.213). The purity of the SarV protein was verified by SDS–PAGE. The retention volume corresponding to the target protein indicated that recombinant SarV is a dimer in solution.

2.2. Crystallization

Purified SarV (in buffer B) was concentrated to 20 mg ml⁻¹ prior to crystallization trials. Initial crystallization trials were performed using the sitting-drop vapour-diffusion method at 289 K. Drops were prepared by mixing 1 µl protein solution with 1 µl reservoir solution in 48-well plates and were

Table 2
Crystallization.

Method	Vapour diffusion
Plate type	48-well plates
Temperature (K)	289
Protein concentration (mg ml ⁻¹)	20
Buffer composition of protein solution	50 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM DTT
Composition of reservoir solution	0.5 M ammonium sulfate, 0.05 M bis-tris pH 6.5, 30%(v/v) pentaerythritol ethoxylate (15/4 EO/OH)
Volume and ratio of drop	1 µl protein solution and 1 µl reservoir solution
Volume of reservoir (µl)	100

equilibrated against a total volume of 100 µl well solution. The initial crystallization conditions were determined using the commercially available Crystal Screen, Crystal Screen 2, Index, SaltRx 1, SaltRx 2, Grid Screen (Hampton Research, USA) and ProPlex (Molecular Dimensions, UK) kits. The best condition was Index condition No. 57, which consists of 0.5 M ammonium sulfate, 0.05 M bis-tris pH 6.5, 30%(v/v) pentaerythritol ethoxylate (15/4 EO/OH). The crystals were reproduced using the hanging-drop vapour-diffusion method and grew in about 7 d to a maximum size of ~1.0 mm in the longest axis (Fig. 1). Crystallization information is summarized in Table 2.

2.3. Data collection and processing

The crystal used for data collection was flash-cooled in liquid nitrogen using a cryoprotectant solution composed of the mother liquor supplemented with 20%(v/v) glycerol. X-ray diffraction data were collected on beamline BL17U1 at Shanghai Synchrotron Radiation Facility (SSRF) using an ADSC Q315 CCD detector with a crystal-to-detector distance of 250 mm. The wavelength was 0.9792 Å and the temperature for data collection was 100 K. Individual frames were

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	BL17U1, SSRF
Wavelength (Å)	0.9792
Temperature (K)	100
Detector	ADSC Q315 CCD
Crystal-to-detector distance (mm)	250
Rotation range per image (°)	1
Total rotation range (°)	220
Exposure time per image (s)	1
Space group	<i>P</i> 2 ₁
Unit-cell parameters (Å, °)	<i>a</i> = 36.40, <i>b</i> = 119.64, <i>c</i> = 66.80, $\alpha = \gamma = 90$, $\beta = 98.75$
Mosaicity (°)	0.90
Resolution range (Å)	50.00–2.10 (2.21–2.10)
Total No. of reflections	121934 (30535)
No. of unique reflections	16749 (4186)
Completeness (%)	93.0 (87.8)
Multiplicity	4.0 (4.0)
$\langle I/\sigma(I) \rangle$	10.8 (3.1)
<i>R</i> _{meas} (%)	8.1 (47.2)
Overall <i>B</i> factor from Wilson plot (Å ²)	24.9

collected with an exposure time of 1 s for each 1.0° oscillation over a range of 220°. X-ray diffraction data were indexed, integrated and scaled using *iMosflm* and *SCALA* in the *CCP4* package (Winn *et al.*, 2011). Data-collection and processing statistics are shown in Table 3.

3. Results and discussion

Full-length recombinant *S. aureus* SarV protein was successfully overexpressed in *E. coli* and purified with a yield of 10 mg from 1 l of LB culture medium. The theoretical molecular weight of recombinant SarV is 17.68 kDa. The purity of the

Figure 2
Gel-filtration profile of SarV and SDS-PAGE of SarV after gel filtration. The gel-filtration profile of SarV is coloured red and that of the protein marker standard for SEC is coloured black. Molecular weights for gel-filtration peaks are labelled in kDa.

Figure 3
A diffraction pattern of the *S. aureus* SarV crystal.

protein sample was verified by Coomassie-stained SDS-PAGE. Two protein bands corresponding to molecular weights of between 14.4 and 18.4 kDa were observed on the SDS-PAGE gel (Fig. 2). The dissolved crystals also showed the same bands on SDS-PAGE. For further investigation, the two bands were cut out from the gel and sent to the Core Facility Center for Life Sciences (University of Science and Technology of China) for liquid chromatography-mass spectrometry (LC-MS) analysis. The LC-MS result identified that both bands contained SarV protein; the band with the lower molecular weight could be a degradation fragment of full-length SarV. The molecular weight of SarV calculated by size-exclusion chromatography (SEC) was 31.2 kDa, suggesting a homodimer in solution (Fig. 2). Other single-domain proteins of the SarA family also form a homodimer to bind promoter DNA, while the two-domain proteins function as a monomer with a complete DNA-binding site formed by two domains.

The purified dimeric protein was successfully crystallized under several conditions from initial crystallization screening trials using the sitting-drop vapour-diffusion method. Larger crystals were reproduced in Index condition No. 57 using the hanging-drop vapour-diffusion method (Fig. 1). Good-quality crystals were sent to an X-ray beamline for data collection and the best data were collected to a resolution of 2.10 Å (Fig. 3, Table 3). Based on the diffraction data, the SarV crystal belonged to the monoclinic space group $P2_1$, with unit-cell parameters $a = 36.40$, $b = 119.64$, $c = 66.80$ Å, $\alpha = \gamma = 90$, $\beta = 98.75^\circ$. The statistics of data collection are shown in Table 3. Calculation of the Matthews coefficient (Matthews, 1968) indicated the presence of four protein molecules in the asymmetric unit ($V_M = 2.57$ Å³ Da⁻¹), with a solvent content of 52%. Self-rotation function analysis showed the presence of two noncrystallographic twofold axes in the $\chi = 180.0^\circ$ section, which is consistent with the Matthews coefficient.

To determine the structure of SarV, both monomeric and dimeric structures of SarA (PDB entry 2frh; 28% sequence identity to SarV; Liu *et al.*, 2006), OhrRC15S (PDB entry 1z91; 25% sequence identity to SarV; Hong *et al.*, 2005), MgrA (PDB entry 2bv6; 24% sequence identity to SarV; Chen *et al.*, 2006) and SarS (PDB entry 1p4x; 25% sequence identity to SarV; Li *et al.*, 2003) were used as search models for molecular replacement. However, owing to low sequence identity, it was difficult to determine the structure of SarV directly by the molecular-replacement method. Experimental phasing using

selenomethionine single-wavelength anomalous dispersion (SAD) and the heavy-atom method is in process.

Acknowledgements

We thank the staff at beamline BL17U1 of the SSRF for assistance with data collection. We thank Dr Baolin Sun (University of Science and Technology of China) for kindly providing *S. aureus* genomic DNA. This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (grant No. XDB 08010101). This work was also supported by grants from the Chinese Ministry of Science and Technology (No. 2012CB917202) and the National Natural Science Foundation of China (Nos. 31370756, 31171241 and 31361163002) and a Scientific Research Grant from the Hefei Science Center of CAS (No. 2015SRG-HSC043) to JZ.

References

- Chen, P. R., Bae, T., Williams, W. A., Duguid, E. M., Rice, P. A., Schneewind, O. & He, C. (2006). *Nature Chem. Biol.* **2**, 591–595.
- Cheung, A. L., Bayer, A. S., Zhang, G., Gresham, H. & Xiong, Y.-Q. (2004). *FEMS Immunol. Med. Microbiol.* **40**, 1–9.
- Cheung, A. L., Nishina, K. A., Trotonda, M. P. & Tamber, S. (2008). *Int. J. Biochem. Cell Biol.* **40**, 355–361.
- Crossley, K. B., Archer, G., Jefferson, K. & Fowler, V. (2009). Editors. *Staphylococci in Human Disease*, 2nd ed. Chichester: John Wiley & Sons.
- Li, R., Manna, A. C., Dai, S., Cheung, A. L. & Zhang, G. (2003). *J. Bacteriol.* **185**, 4219–4225.
- Hong, M., Fuangthong, M., Helmann, J. D. & Brennan, R. G. (2005). *Mol. Cell.* **20**, 131–141.
- Liu, Y., Manna, A., Li, R., Martin, W. E., Murphy, R. C., Cheung, A. L. & Zhang, G. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 6877–6882.
- Liu, Y., Manna, A. C., Pan, C.-H., Kriksunov, I. A., Thiel, D. J., Cheung, A. L. & Zhang, G. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 2392–2397.
- Lowy, F. D. (1998). *N. Engl. J. Med.* **339**, 520–532.
- Manna, A. C., Ingavale, S. S., Maloney, M., van Wamel, W. & Cheung, A. L. (2004). *J. Bacteriol.* **186**, 5267–5280.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Poor, C. B., Chen, P. R., Duguid, E., Rice, P. A. & He, C. (2009). *J. Biol. Chem.* **284**, 23517–23524.
- Schumacher, M. A., Hurlburt, B. K. & Brennan, R. G. (2001). *Nature (London)*, **409**, 215–219.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Zhu, Y., Fan, X., Zhang, X., Jiang, X., Niu, L., Teng, M. & Li, X. (2014). *Acta Cryst.* **D70**, 2467–2476.