

**Zhao-Yang Ye,^{a,b} Qiao-Ming
 Hou,^{a,b} Lan-Fen Li^a and
 Xiao-Dong Su^{a,b*}**

^aPeking University Shenzhen Graduate School, Shenzhen 518055, People's Republic of China, and ^bNational Laboratory of Protein Engineering and Plant Genetic Engineering, College of Life Sciences, Peking University, Beijing 100871, People's Republic of China

Correspondence e-mail: xdsu@pku.edu.cn

Received 8 January 2009

Accepted 14 March 2009

Crystallization and preliminary X-ray crystallographic analysis of SMU.412c protein from the caries pathogen *Streptococcus mutans*

The *smu.412c* gene encodes a putative histidine triad-like protein (SMU.412c) with 139 residues that is involved in cell-cycle regulation in *Streptococcus mutans*. The gene was cloned into the expression vector pET28a and subsequently expressed in *Escherichia coli* strain BL21 (DE3) to give a substantially soluble form of SMU.412c with a His₆ tag at its N-terminus. The recombinant protein was purified to homogeneity in a two-step procedure involving Ni²⁺-chelating and size-exclusion chromatography. Crystals suitable for X-ray diffraction were obtained using the sitting-drop vapour-diffusion method and diffracted to 1.8 Å resolution on beamline BL6A at Photon Factory, Tsukuba, Japan. The crystal belonged to space group *P*4₁2₁2, with unit-cell parameters $a = b = 53.5$, $c = 141.1$ Å.

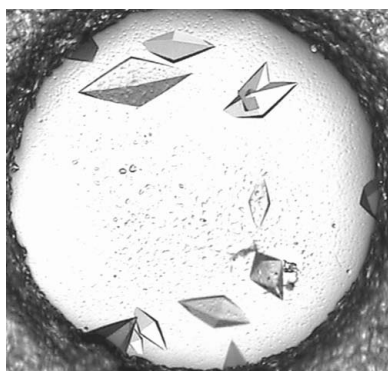
1. Introduction

Streptococcus mutans, a Gram-positive bacterium, has been found to be responsible for human dental caries and sub-acute infective endocarditis (Locsche, 1986; Ulman *et al.*, 1988; Liu, 1999). The details of *S. mutans* pathogenesis largely remain unclear, despite several studies having been carried out including those on biofilm formation, acid tolerance and microbial accumulation (Bleiweis *et al.*, 1992; Yamashita *et al.*, 1993; Quivey *et al.*, 2001).

The genome sequence of *S. mutans* UA159 (a serotype c strain; Ajdić, 2002) contains 1963 ORFs (open reading frames), of which about 63% have been assigned putative functions, 16% are unique to *S. mutans* and 21% have unknown function. The *smu.412c* gene (gi:24376788 or gi:24378905) in the *S. mutans* genome encodes a putative histidine triad-like (HIT) protein. The SMU.412c protein, with a molecular mass of 15.6 kDa and a calculated isoelectric point (pI) of 5.0, has 139 residues and contains an HAHVHLV motif at its 97th amino-acid sequence site.

SMU.412c belongs to the HIT family (all members of which share a common His-*X*-His-*X*-His-*XX* sequence motif, where *X* is a hydrophobic amino-acid residue; Seraphin, 1992; Brenner *et al.*, 1999; Brenner, 2002; Huber & Weiske, 2008). Its closest homologous protein identified to date is a cell-cycle regulation HIT protein (gi:55823514) from *S. thermophilus* CNRZ1066 (67% sequence identity; Bolotin *et al.*, 2004).

Five HIT-protein subfamilies have been identified to date, mostly from mammalian species: Hint (histidine-triad nucleotide-binding protein), Fhit (fragile histidine triad), GalT (galactose-1-phosphate-uridylyltransferase), DcpS (scavenger-decapping protein) and aprataxin (Ahel *et al.*, 2006; Kijas *et al.*, 2006). Hint proteins are able to interact with cyclin-dependent kinase 7 (CDK7) for transcriptional regulation and growth control (Korsisaari & Makela, 2000). Fhit proteins act as tumour suppressors: their nonfunctional expression or lack of expression has been associated with the occurrence of cancer and re-expression of Fhit in Fhit-negative cells can suppress cell growth and reduce tumour formation (Dumon *et al.*, 2001; Roz *et al.*, 2002; Semba *et al.*, 2006). Mutations in aprataxin have been reported in ataxia-oculomotor aprataxia 1 (AOA1, a neurodegenerative disease; Date *et al.*, 2001). Although HIT-family proteins are also



© 2009 International Union of Crystallography
 All rights reserved

widespread in bacteria, their functions have been little investigated to date and it is hoped that structure determination and functional studies of SMU.412c will help to identify its biological function in prokaryotes and contribute to the treatment of human dental caries.

2. Experimental procedures and results

2.1. Cloning and expression

To construct the SMU.412c expression plasmid, two primers containing *Bam*HI and *Xho*I restriction sites were designed as follows: *smu.412c*-F, 5'-CGCGGATCCATGAACGATTGTCTTTTTTGTA-3', and *smu.412c*-R, 5'-CCGCTCGAGTCATGCTTCAATCTCCTT-TTG-3'. The *smu.412c* gene was cloned from the genomic DNA of *S. mutans* by the polymerase chain reaction (PCR). After digestion with *Bam*HI and *Xho*I overnight at room temperature, the PCR product was ligated into the vector pET28a(+) (Novagen) with an N-terminal fusion His₆ tag (MGSSHHHHHSSGLVPRGSHMAS-MTGGQQMGRGS) digested with the same restriction enzymes. The recombinant plasmid containing the target gene *smu.412c*, which was verified by DNA sequencing, was transformed into *Escherichia coli* strain BL21 (DE3) for protein expression.

The transformed *E. coli* cells were cultured overnight at 310 K in ~20 ml Luria–Bertani (LB) medium containing 50 µg ml⁻¹ kanamycin and then inoculated into 1 l LB medium and further cultured under the same conditions until an OD₆₀₀ value of 0.6–0.8 was reached. An aliquot of 1.0 ml 1.0 M isopropyl β-D-1-thiogalactopyranoside was added to the culture to induce protein expression for a further 2–4 h at 310 K. The cells were harvested by centrifugation (5000 rev min⁻¹, 8 min, 277 K) and resuspended in 20–30 ml lysis buffer (20 mM Tris–HCl, 500 mM NaCl pH 7.5).

2.2. Protein purification

The resuspended cells were lysed by sonication on ice and centrifuged at 277 K and 18 000 rev min⁻¹ for 30 min to remove cell debris. The supernatant was loaded onto a 5 ml Ni²⁺-chelating affinity column (HiTrap, GE Healthcare, USA) equilibrated with lysis buffer. Unbound proteins were first washed out with lysis buffer. Proteins with low Ni-affinity were also washed out with 10% (v/v) elution buffer (20 mM Tris–HCl, 500 mM NaCl, 500 mM imidazole pH 7.5) in lysis buffer. Finally, the bound target protein was eluted with a linear gradient of elution buffer from 10% to 100% in lysis buffer in about 15 min. The fractions containing the target protein were pooled and

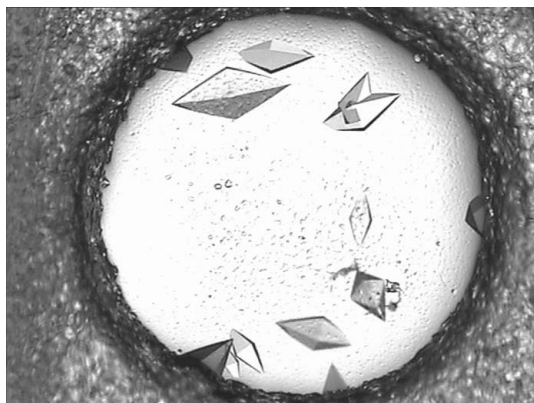


Figure 1
Photograph of crystals of SMU.412c protein. The dimensions of the rhombic crystals are about 0.58 × 0.23 × 0.15 mm.

concentrated by ultrafiltration using a Millipore centrifugal filter device (Amicon Ultra, 10 kDa cutoff) at 277 K.

A gel-filtration column (HiLoad Superdex 75, 120 ml, GE Healthcare, USA) was employed to further purify the SMU.412c protein using an eluant containing 20 mM Tris–HCl, 200 mM NaCl pH 7.5. The target protein was examined by SDS–PAGE at each step. The results showed that the purified protein had a molecular mass of 19.5 kDa (data not shown), which is consistent with the calculated molecular mass for the SMU.412c protein of 15.6 kDa plus a 4 kDa His tag.

The purified protein was directly concentrated using the same ultrafiltration device to a final concentration of about 10 mg ml⁻¹ for crystallization experiments.

2.3. Protein crystallization

The sitting-drop vapour-diffusion method was used to screen for crystallization conditions for SMU.412c at 293 K using several commercially available kits such as Crystal Screen, Crystal Screen 2 and Index Screen (Hampton Research, USA) as well as a BCR (Bio-Crystal Research) kit (XtalQuest Inc., Beijing, China) as screening conditions in 48-well SBS (ANSI/SBS 1-2004; American National Standards Institute/Society for Biomolecular Sciences, USA) microplates (XtalQuest Inc., Beijing, China). An aliquot of 1 µl protein solution was mixed with an equal volume of reservoir solution and equilibrated against 100 µl reservoir solution. After 3 d, several rhombic crystals suitable for X-ray diffraction experiments appeared in the condition 2.8 M sodium acetate pH 7.0 (Hampton Research Index Screen condition No. 24). The crystal dimensions were about 0.6 × 0.2 × 0.15 mm (Fig. 1).

2.4. X-ray diffraction data collection and processing

Diffraction data were collected using an ADSC Quantum 4R Detector (188 × 188 mm) on beamline BL6A, Photo Factory, Tsukuba, Japan. The crystals (Fig. 1) from the original condition (2.8 M sodium acetate pH 7.0) were directly flash-cooled without

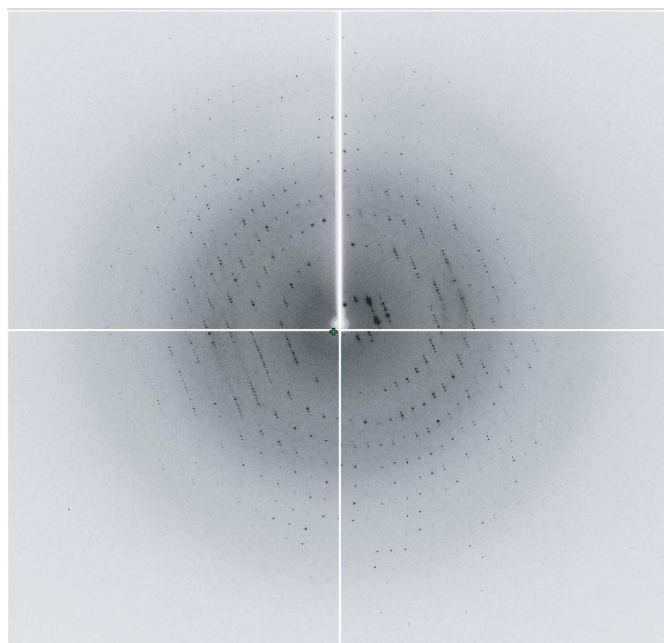


Figure 2
Diffraction image of a crystal of SMU.412c protein obtained using the condition 2.8 M sodium acetate pH 7.0 (Index Screen condition No. 24).

Table 1

Data-collection statistics for SMU.412c protein.

Values in parentheses are for the highest resolution shell.

Resolution (Å)	50–1.8 (1.9–1.8)
Completeness (%)	98.3 (100)
R_{merge} (%) [†]	7.1 (41.5)
Mean $I/\sigma(I)$	31.3 (5.1)
Space group	$P4_12_12$
Unit-cell parameters (Å)	$a = b = 53.3, c = 141.1$
No. of observed reflections	36286
No. of unique reflections	23313
Molecules per ASU	1
V_M (Å ³ Da ⁻¹)	3.2
Solvent content (%)	61.8

[†] $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the i th observation of reflection hkl and $\langle I(hkl) \rangle$ is the mean intensity of all observations of hkl .

cryoprotectant and maintained at 100 K in a stream of cold nitrogen during data collection. The crystal-to-detector distance was set to 330 mm and the wavelength was 1.000 Å. A total of 160 frames were collected with 1° φ oscillation per frame. Fig. 2 shows a typical diffraction pattern. Diffraction data were processed using the XDS program suite (Kabsch, 1993).

The crystals diffracted to a resolution of 1.8 Å (Fig. 2) and belonged to space group $P4_12_12$ as indicated by systematic absences. The unit-cell parameters were $a = b = 53.3, c = 141.1$ Å. Assuming the presence of one molecule per asymmetric unit, the V_M value is 3.2 Å³ Da⁻¹ (Matthews, 1968), corresponding to a solvent content of 61.8%. Data-collection statistics are listed in Table 1. The phases of the structure have been determined using the molecular-replacement method and refinement of the structure is in progress.

This work was supported by a grant from the National Natural Science Foundation of China (30530190). Peking University's 985 and 211 grants are greatly acknowledged. We are also grateful for the

support from the Photon Factory of KEK, Tsukuba, Japan for data collection (2008-G613).

References

- Ahel, I., Rass, U., El-Khamisy, S. F., Katyal, S., Clements, P. M., McKinnon, P. J., Caldecott, K. W. & West, S. C. (2006). *Nature (London)*, **443**, 713–716.
- Ajdić, D. *et al.* (2002). *Proc. Natl Acad. Sci. USA*, **99**, 14434–14439.
- Bleiweis, A. S., Oyston, P. C. & Brady, L. J. (1992). *Adv. Exp. Med. Biol.* **327**, 229–241.
- Bolotin, A. *et al.* (2004). *Nature Biotechnol.* **22**, 1554–1558.
- Brenner, C. (2002). *Biochemistry*, **41**, 9003–9014.
- Brenner, C., Bieganowski, P., Pace, H. C. & Huebner, K. (1999). *J. Cell. Physiol.* **181**, 179–187.
- Date, H. *et al.* (2001). *Nature Genet.* **29**, 184–188.
- Dumon, K. R., Ishii, H., Vecchione, A., Trapasso, F., Baldassarre, G., Chakrani, F., Druck, T., Rosato, E. F., Williams, N. N., Baffa, R., During, M. J., Huebner, K. & Croce, C. M. (2001). *Cancer Res.* **61**, 4827–4836.
- Huber, O. & Weiske, J. (2008). *Cell Cycle*, **7**, 1326–1331.
- Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
- Kijas, A. W., Harris, J. L., Harris, J. M. & Lavin, M. F. (2006). *J. Biol. Chem.* **281**, 13939–13948.
- Korsisaari, N. & Makela, T. P. (2000). *J. Biol. Chem.* **275**, 34837–34840.
- Liu, T. J. (1999). *The Microbiology Basis of Nonnasality Illness*. Beijing: People's Health Press.
- Loesche, W. J. (1986). *Microbiol. Rev.* **50**, 353–380.
- Matthews, B. W. (1968). *J. Biol. Chem.* **237**, 778–786.
- Quivey, R. G., Kuhnert, W. L. & Hahn, K. (2001). *Crit. Rev. Oral Biol. Med.* **12**, 301–314.
- Roz, L., Gramegna, M., Ishii, H., Croce, C. M. & Sozzi, G. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 3615–3620.
- Semba, S., Trapasso, F., Fabbri, M., McCorkell, K. A., Volinia, S., Druck, T., Iliopoulos, D., Pekarsky, Y., Ishii, H., Garrison, P. N., Barnes, L. D., Croce, C. M. & Huebner, K. (2006). *Oncogene*, **25**, 2860–2872.
- Seraphin, B. (1992). *DNA Seq.* **3**, 177–179.
- Ulman, R. F., Miller, S. J., Strampfer, M. J. & Cunha, B. A. (1988). *Heart Lung*, **17**, 209–212.
- Yamashita, Y., Bowen, W. H., Burne, R. A. & Kuramitsu, H. K. (1993). *Infect. Immun.* **61**, 3811–3817.