Available online at www.sciencedirect.com
SCIENCE DIRECT
Biochimica et Biophysica Acta 1764 (2006) 324–326
BIOCHIMICA ET BIOPHYSICA ACTA BBA
http://www.elsevier.com/locate/bba

Short crystallization paper

Protein preparation, crystallization and preliminary X-ray crystallographic analysis of Smu.1475c from caries pathogen *Streptococcus mutans*

Yan-Feng Zhou[a], Wei Mi[a], Lanfen Li[a], Xiaoyan Zhang[a], Yu-He Liang[a]
Xiao-Dong Su[a,*], Shicheng Wei[b,*]

[a] *National Laboratory of Protein Engineering and Plant Genetic Engineering, College of Life Sciences, Peking University, Beijing 100871, China*
[b] *Peking University School of Stomatology, Beijing 100081, China*

Received 2 August 2005; received in revised form 14 November 2005; accepted 15 November 2005
Available online 28 December 2005

Abstract

The gene *smu.1475c* encodes a putative protein of 211 residues in *Streptococcus mutans*, a primary pathogen for human dental caries. In this work, *smu.1475c* was cloned into pET28a and expressed in good amount from the *E. coli* strain BL21 (DE3). Smu.1475c protein was purified to homogeneity in a two-step procedure of Ni^{2+} chelating and size exclusion chromatography. Crystals were obtained by hanging-drop vapor-diffusion method and diffracted to 2.7 Å resolution. The crystal belongs to orthorhombic space group $P2_12_12_1$ with cell dimension of $a=68.3$ Å, $b=105.9$ Å, $c=136.2$ Å. The asymmetric unit is expected to contain four molecules with solvent content of 49.4%.

Keywords: *Streptococcus mutans*; Dental caries; Smu.1475c; Protein crystallography

The Gram positive bacterium *Streptococcus mutans* is the principal etiological agent in human dental caries [1] and a primary pathogen of infective endocarditis as well [2]. The complete genome sequences of *S. mutans* UA159, a serotype c strain, have been published in 2002. It contains 1,963 ORFs (open reading frames), about 63% of which have been assigned putative functions [3]. Studies on biofilm formation and cariogenesis caused by *S. mutans* so far were mainly focused on virulence factors such as the cell-surface fibrillar proteins, which mediate adherence to the tooth surface [4], the glucosyltransferase enzymes, which synthesize adhesive glucans and allow microbial accumulation on teeth [5], as well as factors related to acidogenicity and acid tolerance of the bacteria. These protein factors make the bacteria to form sustained biofilm on tooth surfaces and survive in acidic conditions promoting the formation of dental caries [6]. In order to study the cariogenesis mechanism and to discover potential drug or vaccine targets, we have initiated a structural genomics project on *S. mutans* and selected the 37% function unknown genes as our primary targets [7]. The *smu.1475c* gene encodes a hypothetical protein of 211 residues, with a molecular mass of 24.6 kDa. Sequence search showed that most proteins with high sequence identity to Smu.1475c are from *Streptococcus* genus (Fig. 1). Smu.1475c has no sequence related structurally known homolog in the PDB (Protein Data Bank). Determination of the crystal structure of Smu.1475c will help us to identify its exact biological function and to provide clues for further studies of this family of proteins.

To construct expression plasmid, target DNA fragment was amplified by polymerase chain reaction (PCR) [8] from the *S. mutans* genomic DNA, using primers SMU1475c-F (5′-cgcggatccgtgaattatcttgatttagccct-3′) and SMU1475c-R (5′-ccgctcgagttaatcttctttgaggtaaagca-3′), which contains *Bam*HI and *Xho*I restriction site, respectively. The PCR-amplified fragment was digested by *Bam*HI and *Xho*I and then ligated to the expression vector pET28a (Novagen) by conventional cloning method. The final vector was verified by DNA sequencing. The constructed plasmid was transformed into *Escherichia coli* strain BL21 (DE3) cells for expression experiments. The *E. coli* BL21 (DE3) strain containing pET28a-

* Corresponding authors. X.-D. Su is to be contacted at tel.: +86 10 62759743; fax: +86 10 62765669. S. Wei, tel.: +86 10 62179977 2308; fax: +86 10 62173402.
E-mail addresses: su-xd@pku.edu.cn (X.-D. Su), kqsc-wei@bjmu.edu.cn (S. Wei).

1570-9639/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2005.11.022

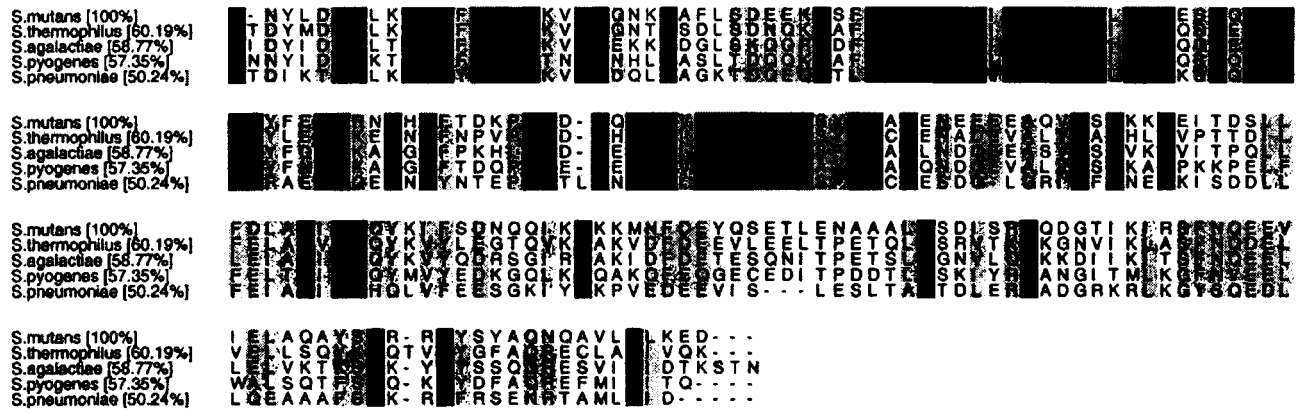

Fig. 1. Multiple sequence alignment of Smu.1475c homologous sequences. The alignment is performed using the program ClustalX [10]; the percentage identities with Smu.1475c are indicated after the organism names; fully and partially conserved residues are marked with blue and yellow background, respectively.

SMU1475c was grown overnight at 37 °C in 20 mL of Luria–Bertani (LB) medium containing kanamycin (50 μg/mL); the overnight culture was then added to 1 liter of LB medium containing 50 μg/mL kanamycin. When the Optical Density at wavelength 600 nm (OD_{600}) of the cell culture reached 0.6, the culture was moved to 30 °C and incubated for 5 min to cool the medium down, then isopropylthio-β-D-galactoside (IPTG) was added to the culture at a final concentration of 0.5 mM. After incubation at 30 °C for six hours, cells were harvested by centrifugation (6000 rpm, 10 min, Beckman, Rotor ID JA10) and stored at −20 °C for later use. For protein purification, the frozen pellet was allowed to thaw at room temperature, and the cells were suspended in 20 mL of buffer A (20 mM Tris/HCl pH 7.5, 500 mM NaCl), and disrupted by sonication. After centrifugation (18,000 rpm, 40 min, 4 °C, Beckman, Rotor ID JA25.50), the supernatant were collected and applied to a 5-mL Hitrap chelating column (Ni^{2+} charged, GE-Amersham Pharmacia Biotech, Sweden). Impurities and target protein were washed out by 5% and 30% buffer B (20 mM Tris/HCl pH 7.5, 500 mM NaCl, 500 mM imidazole) respectively. Then size exclusion chromatography (Hiload Superdex 75 XK16/60, GE-Amersham Pharmacia Biotech, Sweden) was chosen for the further purification, using buffer C (20 mM Tris/HCl pH 7.5, 200 mM NaCl). Protein purity was examined in each step by SDS-PAGE using Coomassie Blue staining.

Table 1
X-ray data collection and processing statistics

Resolution range (Å)	50–2.7 (2.82–2.7)[a]
Data completeness (%)	99.8 (99.7)
R_{sym} (%)[b]	9.85 (27.96)
Average *I/sig <I>*	5.37 (1.59)
Space group	$P2_12_12_1$
Unit cell parameters (Å)	a=68.3, b=105.9, c=136.2
No. of observed reflections	215,248
No. of unique reflections	27,936
Mol/Asym.	3–5
V_m ($Å^3$ Da^{-1})	3.10, 2.43 or 1.87
Solvent content (%)	40.4, 49.4 or 34.2

[a] Values in parentheses refer to the highest resolution shell.
[b] $R_{sym}=\sum|I_{obs}-I_{avg}|/\sum I_{obs}$.

The purified protein was concentrated to 20 mg/mL by ultrafiltration (Millipore Amincon). Initial crystallization screening was performed using the Hampton Kits, Crystal Screen, Crystal Screen 2 and Index (Hampton Research, CA, USA) by hanging-drop vapor-diffusion method performed at 16 °C. 1 μL protein solution was mixed with 1 μL reservoir solution and equilibrated against 500 μL reservoir solution. Micro crystals were observed from several conditions, and optimization was applied according to the condition Index 55. The crystals suitable for X-ray diffraction were obtained in an optimized condition containing 0.05 M $MgCl_2$, 0.1 M Tris–HCl (pH 7.5) and 25% v/v PEG MME 550. X-ray diffraction data were collected on a Bruker SMART 6000 CCD detector using Cu Kα radiation from a Bruker Nonius FR591 rotating anode generator operated at 45 kV and 100 mA. The crystal to detector distance was set to 6 cm. 1000 frames were collected with 0.2° φ oscillation per frame, and the exposing time was 100 s per frame. During data collection, the crystal was maintained at 100 K using nitrogen gas, the cryo-protecting buffer contained 0.05 M $MgCl_2$, 0.1 M Tris–HCl pH 7.5, 30% v/v PEG MME 550 and 15% v/v PEG 400. Diffraction data were processed using the Bruker on-line software PROTEUM suite. The crystal diffracted to 2.7 Å and belonged to space group $P2_12_12_1$, with cell dimension of a=68.3 Å, b=105.9 Å, c=136.2 Å. There could be 3 to 5 molecules per asymmetric unit, which will give V_m values of 3.10, 2.43 or 1.87 $Å^3$ Da^{-1} [9], corresponding to solvent contents of 40.4%, 49.4% or 34.2%. The crystallographic parameters and data collection statistics are listed in Table 1.

Acknowledgements

This work was supported by grant from National High Technology and Development Program of China (863 program 2002BA711A13). Peking University's 985 and 211 grants are also greatly acknowledged. XDS is a recipient of National

Science Fund for Distinguished Young Scholars of NSFC (30325012).

References

[1] W.J. Loesche, Role of *Streptococcus mutans* in human dental decay, Microbiol. Rev. 50 (1986) 353–380.
[2] R.F. Ullman, S.J. Miller, M.J. Strampfer, B.A. Cunha, *Streptococcus mutans* endocarditis: report of three cases and review of the literature, Heart Lung 17 (1988) 209–212.
[3] D. Ajdic', W.M. McShan, R.E. McLaughlin, G. Savic', J. Chang, M.B. Carson, C. Primeaux, R. Tian, S. Kenton, H. Jia, S. Lin, Y. Qian, S. Li, H. Zhu, F. Najar, H. Lai, J. White, B.A. Roe, J.J. Ferretti, Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen, Proc. Natl. Acad. Sci. 99 (2002) 14434–14439.
[4] A.S. Bleiweis, P.C. Oyston, L.J. Brady, Molecular, immunological and functional characterization of the major surface adhesin of *Streptococcus mutans*, Adv. Exp. Med. Biol. 327 (1992) 229–241.
[5] Y. Yamashita, W.H. Bowen, R.A. Burne, H.K. Kuramitsu, Role of the *Streptococcus mutans* gtf genes in caries induction in the specific-pathogen-free rat model, Infect. Immun. 61 (1993) 3811–3817.
[6] R.G. Quivey, W.L. Kuhnert, K. Hahn, Genetics of acid adaptation in oral streptococci, Crit. Rev. Oral Biol. Med. 12 (2001) 301–314.
[7] X.Y. Zhang, W. Mi, Y.F. Zhou, X.Y. Liu, L.F. Li, Y.H. Liang, S.C. Wei, X.D. Su, Protein preparation, crystallization and preliminary X-ray crystallographic analysis of Smu.260 from *Streptococcus mutans*—A cariogenic dental pathogen, Prog. Biochem. Biophys. 32 (2005) 217–220.
[8] R. Saiki, D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, H.A. Erlich, Primer-directed enzymatic amplification of DNA with thermostable DNA polymerase, Science 239 (1988) 487–491.
[9] K.A. Kantardjieff, B. Rupp, Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals, Protein Sci. 12 (2003) 1865–1871.
[10] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res. 25 (1997) 4876–4882.