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

Purification, crystallization and preliminary X-ray studies of the putative lysozyme SP0987 from *Streptococcus pneumoniae*

Siqiang Niu,[a,b] Miao Luo,[a,b] Ailong Huang,[a] Yibing Yin[b]*‡ and Deqiang Wang[a,b]*‡

[a]Key Laboratory of Molecular Biology of Infectious Diseases, Chongqing Medical University, YiXueYuanlu-1, Chongqing 400016, People's Republic of China, and [b]Department of Laboratory Medicine, Chongqing Medical University, YiXueYuanlu-1, Chongqing 400016, People's Republic of China

‡ The corresponding authors contributed equally to this work.

Correspondence e-mail: yibingyin@21cn.com, wangd@ustc.edu

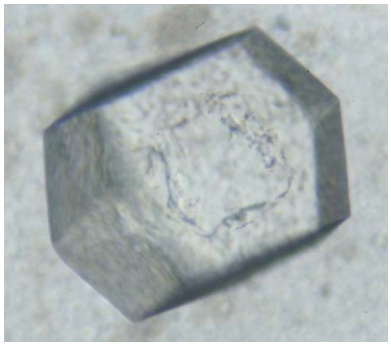
Received 27 November 2009
Accepted 4 January 2010

Streptococcus pneumoniae SP0987, which was identified as a hypothetical protein, has a very low sequence identity to other well characterized lysozyme structures. Since determination of three-dimensional structure is a powerful means of functional characterization, X-ray crystallography has been used to accomplish this task. Here, the expression, purification, crystallization and preliminary crystallographic analysis of SP0987 from *Streptococcus pneumoniae* TIGR4 are reported. The crystal belonged to space group $P2_12_12_1$ (with unit-cell parameters a = 36.46, b = 40.89, c = 147.44 Å) and diffracted to a resolution of 1.85 Å. The crystals are most likely to contain one molecule in the asymmetric unit, with a V_M value of 2.02 Å^3 Da^{-1}.

1. Introduction

Lysozymes cleave the β-1,4-glycosidic bond between *N*-acetylmuramic acid (NAM) and *N*-acetylglucosamine (NAG) in bacterial peptidoglycan (Nelson *et al.*, 2001). These enzymes are broadly expressed in all organisms. Based on the NAM–NAG bond cleaved by the structurally diverse set of enzymes, lysozymes can be classified into five families: GH22 (*e.g.* hen egg-white lysozyme), GH23 (*e.g.* goose egg-white lysozyme), GH24 (*e.g.* bacteriophage T4 lysozyme), GH73 (*e.g. Sphingomonas* flagellar protein) and the GH25 family exemplified by *Chalaropsis* lysozymes (Martinez-Fleites *et al.*, 2009). The GH25 lysozymes, which are composed of an eight-stranded β-barrel flanked by five α-helices, are structurally unrelated to the GH22–24 and GH73 lysozymes (Rau *et al.*, 2001). The characterized lysozymes from this family exhibit both β-1,4-*N*-acetylmuramidase and β-1,4-*N*,6-*O*-diacetylmuramidase activity (Fouche & Hash, 1978; Vollmer *et al.*, 2008).

SP0987 was identified as a hypothetical protein in the genomic sequence database for *Streptococcus pneumoniae* TIGR4. It consists of 266 amino-acid residues and has a predicted molecular weight of 30.0 kDa. When the amino-acid sequence of SP0987 was entered into the PFAM 23.0 database (July, 2008) of protein families, only glycosyl hydrolase family 25 (GH25) was found, suggesting that the protein may exhibit hydrolytic enzyme activity (Nelson *et al.*, 2001). To date, the structures of three members of this family of enzymes have been reported: PlyB (a lysin) from *Bacillus anthracis* (Porter *et al.*, 2007), MsrA from *S. pneumoniae* (Kim *et al.*, 2009) and serine hydrolase from *Saccharomyces cerevisiae* (Arndt *et al.*, 2009). The *S. pneumoniae* protein shares very low sequence identity (about 8–18%) with these three GH25-family members, implying that SP0987 may exhibit new structural and functional characteristics. In order to gain further insight into the function of SP0987, we have initiated the determination of its three-dimensional structure by X-ray crystallography. Here, we report the crystallization and preliminary X-ray study of SP0987 from *S. pneumoniae* TIGR4. Structure determination will be pursued using experimental phasing methods.



2. Materials and methods

2.1. Cloning and expression

A fragment of *sp0987* (spanning residues 25–266) was amplified from *S. pneumoniae* TIGR4 genomic DNA using the following

primers: the sense primer 5′-GTA**GAATTC**CGCCCACTTAATGA CAGCCAAT-3′ and the antisense primer 5′-CGGGC**CTCGAG**TT-AAGGTTTTAAAAAAAGTTTTC-3′, which contain *Eco*RI and *Xho*I restriction sites (nucleotides shown in bold), respectively. The amplified fragment was digested using the *Eco*RI and *Xho*I endonucleases and then cloned into pET28a vector (previously digested with the same restriction enzymes) to generate pET28-*SP0987*-His6. The complete nucleotide sequence of the insert was confirmed by DNA-sequence analysis on an ABI 3100 DNA Sequencer (Applied Biosystems). The recombinant plasmid was then transformed into *Escherichia coli* BL21 (DE3) cells. These cells were grown (at 293 K) in 1.2 l LB medium containing 50 mg l^{-1} kanamycin. At an $OD_{600\,nm}$ of 0.4, protein expression was induced by the addition of 0.2 m*M* isopropyl *β*-D-1-thiogalactopyranoside (IPTG). 15 h later the cells were harvested by centrifugation for 10 min at 4000*g* and 277 K.

2.2. Purification

The cell pellet was resuspended in 50 ml lysis buffer (20 m*M* Tris–HCl pH 8.0 and 300 m*M* sodium chloride) and lysed by sonication on ice. All subsequent purification steps were performed at 277 K. The lysate was centrifuged at 15 000*g* for 30 min. The supernatant was applied onto a 10 ml Ni^{2+}–NTA affinity column (Qiagen) equilibrated with lysis buffer. Nonspecifically bound proteins were washed from the column using 200 ml lysis buffer containing 15 m*M* imidazole. The recombinant protein was then eluted from the column with 20 ml elution buffer containing 20 m*M* Tris–HCl pH 8.0, 300 m*M* sodium chloride and 400 m*M* imidazole. The protein was concentrated and buffer-exchanged to the final buffer (5 m*M* Tris–HCl pH 8.0 and 10 m*M* sodium chloride) using a 10 kDa cutoff Millipore Amicon concentrator. The final purified protein concentration was about 50 mg ml^{-1} and the purity was determined by SDS–PAGE to be about 95%.

2.3. Crystallization

Crystallization experiments were carried out using the hanging-drop and sitting-drop vapour-diffusion methods. The Crystal Screen I and II and PEG/Ion Screen kits (Hampton Research) provided the reservoir solutions. The crystallization droplet contained 1.0 µl His-tagged protein and 1.0 µl reservoir solution. The drops were equilibrated against 200 µl reservoir solution at 293 K. After 10 d, crystals of SP0987 were observed using reservoir conditions consisting of 20%(*w*/*v*) polyethylene glycol (PEG) 3350 and 0.2 *M* lithium citrate. Further optimization of this condition, varying the PEG 3350 concentration and the pH, gave good diffraction-quality crystals at 24–26% PEG 3350 and 0.2 *M* lithium citrate in the presence of 0.1 *M* bis-tris pH 7.31. Crystals appeared after 8 d equilibration against the crystallization solution and grew to full size (0.4 × 0.25 × 0.15 mm) in 10 d (Fig. 1).

2.4. Data collection and processing

X-ray diffraction data were collected from a single crystal using a Rigaku RU-200 rotating-anode X-ray generator equipped with Osmic focusing mirrors and a MAR Research image-plate system (diameter 345 mm). The crystals were mounted in a nylon-fibre loop and flash-cooled in a nitrogen-gas stream at 100 K without additional cryoprotectant. The crystal-to-detector distance was 150 mm. The oscillation range was 1° per image and a total of 137° of data were collected (see Fig. 2). The data were indexed and scaled with the *HKL*-2000 program package (Otwinowski & Minor, 1997). Table 1 lists the final data-collection and processing statistics.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell

Space group	$P2_12_12_1$
Unit-cell parameters (Å, °)	a = 36.46, b = 40.89, c = 147.44
Wavelength (Å)	1.5418
Resolution (Å)	50–1.85 (1.90–1.85)
No. of observed reflections	171963
No. of unique reflections	19741
Completeness (%)	95.1 (91.0)
$\langle I/\sigma(I)\rangle$	14.3 (8.6)
R_{merge}† (%)	2.9 (18.5)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith measurement of an equivalent reflection with indices hkl.

3. Results and discussion

Initial attempts to purify recombinant full-length SP0987 were not successful. The full-length protein was not expressed in BL21, B834 and Rosetta *E. coli* strains (data not shown). On the basis of bioinformatics predictions, the hydrophobic N-terminal residues (1–24) may fold into a transmembrane helix (Tusnády & Simon, 2001; Krogh *et al.*, 2001). Accordingly, we deleted these hydrophobic residues and constructed the plasmid pET28-*SP0987*-His6 spanning residues 25–266, which includes the GH25 muramidase domain (July, 2008). Fortunately, the fragment was expressed in both BL21 (DE3) and B834 (DE3) cells, yielding about 50 mg of highly purified protein per litre of bacterial culture.

A complete diffraction data set was collected to 1.85 Å resolution from a single crystal and the data-collection statistics are reported in Table 1. A total of 171 963 measured reflections in the resolution range 50–1.85 Å were merged to 19 741 unique reflections with an R_{merge} of 2.9%. Analysis of the diffraction intensities indicated that the most plausible space group was orthorhombic $P2_12_12_1$, with unit-cell parameters a = 36.46, b = 40.89, c = 147.44 Å. Based on the molecular weight of SP0987 residues 25–266 (27.15 kDa) and space group $P2_12_12_1$, it was assumed that the crystal contained three molecules per asymmetric unit. The assumption gives a V_M value of 2.02 $Å^3$ Da^{-1} and a solvent content of 39.27% (Matthews, 1968). Molecular replacement using *AMoRe* (Navaza, 2001), *Phaser* (McCoy *et al.*, 2005) and *MOLREP* (Vagin & Teplyakov, 2000) was carried out using the structure of the catalytic domain of PlyB ($PlyB_{cat}$) as a search model (PDB code 2nw0; Porter *et al.*, 2007).

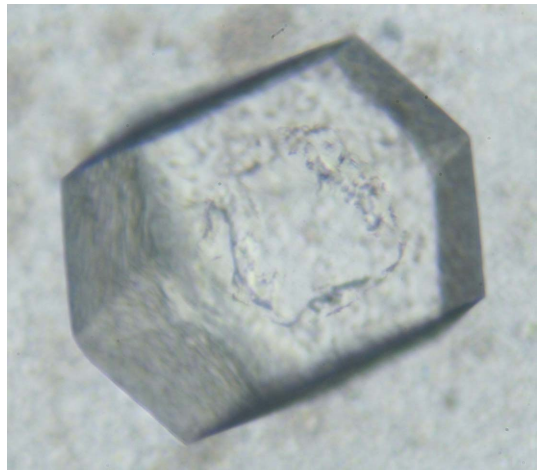


Figure 1
The crystal of SP0987 from *S. pneumoniae* TIGR4.

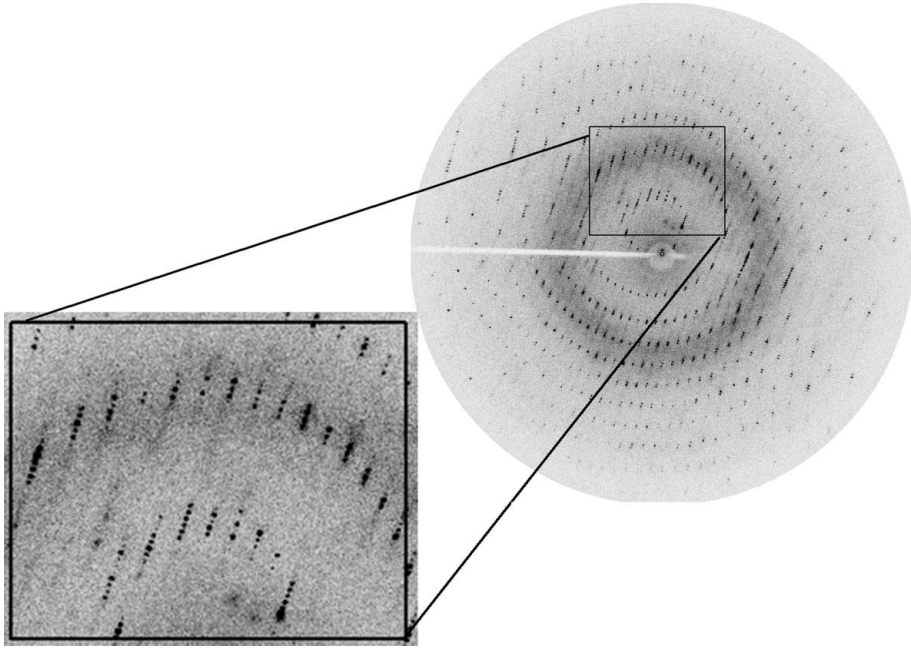


Figure 2
X-ray diffraction pattern of the SP0987 crystal from *S. pneumoniae* TIGR4 recorded using a MAR345 image plate. An enlarged image is shown on the left.

The homologous $PlyB_{cat}$ has about 25% sequence identity to the C-terminal domain of SP0987 from *S. pneumoniae* TIGR4. However, efforts to determine the structure using molecular replacement with $PlyB_{cat}$ have proven unsuccessful. We are now in the process of producing selenomethionylated SP0987 protein in order to use multi-wavelength anomalous dispersion methods for structure determination (the SP0987 protein contains four methionine residues).

The authors would like to acknowledge the X-ray facility of the Key Laboratory of Structural Biology, Chinese Academy of Science. We also would like to thank Professor Liwen Niu, Professor Maikun Teng and Dr Zhiqiang Zhu, School of Life Science, University of Science and Technology of China. Financial support for this project to DW was provided by research grants from the Chinese National Natural Science Foundation (grant Nos. 30600101 and 30770481) and the Natural Science Foundation of Chongqing (grant No. 2006BB5275). We also thank Dr D. K. Worthylake and Dr L. LeCour Jr (Department of Biochemistry and Molecular Biology, Louisiana State University) for carefully editing the manuscript.

References

Arndt, J. W. *et al.* (2005). *Proteins*, **58**, 755–758.
Fouche, P. B. & Hash, J. H. (1978). *J. Biol. Chem.* **253**, 6787–6793.
Kim, Y. K., Shin, Y. J., Lee, W. H., Kim, H.-Y. & Hwang, K. Y. (2009). *Mol. Microbiol.* **72**, 699–709.
Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. (2001). *J. Mol. Biol.* **305**, 567–580.
Martinez-Fleites, C., Korczynska, J. E., Davies, G. J., Cope, M. J., Turkenburg, J. P. & Taylor, E. J. (2009). *Carbohydr. Res.* **344**, 1753–1757.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Navaza, J. (2001). *Acta Cryst.* D**57**, 1367–1372.
Nelson, D., Loomis, L. & Fischetti, V. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 4107–4112.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Porter, C. J., Schuch, R., Pelzek, A. J., Buckle, A. M., McGowan, S., Wilce, M. C., Rossjohn, J., Russell, R., Nelson, D., Fischetti, V. A. & Whisstock, J. C. (2007). *J. Mol. Biol.* **366**, 540–550.
Rau, A., Hogg, T., Marquardt, R. & Hilgenfeld, R. (2001). *J. Biol. Chem.* **276**, 31994–31999.
Tusnády, G. E. & Simon, I. (2001). *Bioinformatics*, **17**, 849–950.
Vagin, A. & Teplyakov, A. (2000). *Acta Cryst.* D**56**, 1622–1624.
Vollmer, W., Joris, B., Charlier, P. & Foster, S. (2008). *FEMS Microbiol. Rev.* **32**, 259–286.