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

Cloning, overexpression, purification, crystallization, and preliminary X-ray studies of SP_0149, the substrate binding protein of an ABC transporter from *Streptococcus pneumoniae*

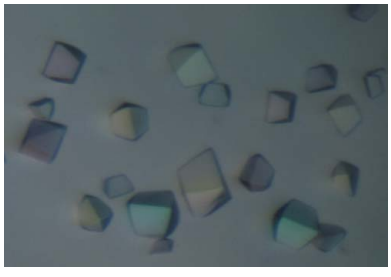
Jaya Bhushan,[a] Rajan Vyas,[b] Tripti Sharma,[a]‡ Devinder Sehgal[a] and Bichitra Kumar Biswal[b]*

[a]Molecular Immunology Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi, 110067, India, and [b]Protein Crystallography Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi, 110067, India

‡ Present address: Max-Planck Institute of Molecular Plant Physiology, 14476 Potsdam-Golm, Germany, and the University of Potsdam, Institute of Biochemistry and Biology, Heisenberg Group (BPMPB), Haus 20, D-14476 Potsdam-Golm, Germany.

Correspondence e-mail: bbiswal@nii.res.in

Received 8 March 2011
Accepted 15 May 2011



A truncated (29 residues from the N-terminus) and N-terminal His-tagged form of SP_0149 from pneumococcal strain ATCC BAA-334 was overexpressed and purified to homogeneity using affinity and gel-filtration chromatography. Diffraction quality crystals were grown at 293 K using the hanging-drop vapour-diffusion technique. X-ray diffraction data were collected to 2.3 Å resolution from a single-crystal that belonged to the orthorhombic space group $P2_12_12_1$ with the unit-cell parameters a = 54.56, b = 75.61, c = 75.52 Å. The calculated values of the Matthews coefficient assuming one molecule (with calculated molecular weight of 30 400 Da) in the crystal asymmetric unit and the corresponding solvent content were 2.56 Å^3 Da^{-1} and 52.0%, respectively.

1. Introduction

The human pathogen *Streptococcus pneumoniae* (also referred to as pneumococci) is the most common bacterial cause of pneumonia, and *S. pneumoniae* infection can also lead to septicemia and meningitis (Lynch & Zhanel, 2010). Infants, the elderly and immunocompromized individuals are at the highest risk of getting pneumococcal infections. Global estimates suggest that *S. pneumoniae* causes 11% of all deaths in children less than 5 years of age (O'Brien *et al.*, 2009). Approximately 800 000 children die each year from pneumococcal disease and more than 90% of these deaths occur in developing countries (Johnson *et al.*, 2010). The existing polysaccharide-based pneumococcal vaccines have limitations. The shortcomings of the 23-valent pneumococcal polysaccharide-based vaccine include its limited serotype coverage and poor immunogenicity in high risk groups. Current pneumococcal glycoconjugate vaccines effectively prevent most invasive disease caused by vaccine serotype strains but are very expensive. The replacement of vaccine serotypes by other non-vaccine serotypes and the rapid emergence of antibiotic resistant strains have compounded the medical treatment of pneumococcal infections (Lynch & Zhanel, 2010; McIntosh & Reinert, 2011). Thus, there is an urgent need to develop new approaches to prevent and treat pneumococcal diseases.

While the polysaccharide capsule is considered to be one of the key virulence factors of pneumococci, many surface-associated and secreted proteins are being studied for their role in pneumococcal physiology, pathogenesis and protective immunity (Bergmann & Hammerschmidt, 2006). Pneumococcal ATP-binding cassette (ABC) transporters constitute a major class of surface-exposed molecules that are being implicated in these processes. In general, ABC transporters have a major impact on pneumococcal physiology, and their disruption can have strong deleterious effects on virulence (Basavanna *et al.*, 2009). ABC transporters play an important role in bacteria by importing various nutrients required for survival and exporting molecules toxic to the cell, among other functions (Davidson *et al.*, 2008). The association of some ABC transporters with survival and/or virulence in the host, their locations in the pneumococcal cell wall and the fact that these importers are not found in humans make them ideal targets for development of novel anti-pneumococcal drugs and immunotherapy (Garmory & Titball, 2004). However, the successful development of post-infection therapies that target ABC transporters is likely to be dependent on the knowledge of their physiological role, structure and molecular

mode of action. Thus, determining the three-dimensional structure of ABC transporters in general and their substrate binding proteins in particular, can provide insights into the mechanism of their action at the molecular level (Oldham *et al.*, 2008; Locher, 2009; Rees *et al.*, 2009) and may help in the rational design of inhibitors for these transporters. In this paper, we report the molecular cloning, overexpression, purification and preliminary X-ray diffraction analysis of SP_0149 (GenBank accession number: NP_344691), the substrate-binding component of a probable L-methionine ABC transporter from *S. pneumoniae*.

2. Experimental methods

2.1. Pneumococcal strain and culture conditions

The pneumococcal strain ATCC BAA-334 (also referred to as TIGR4; serotype 4) was sourced from the American Type Culture Collection, USA and maintained as described previously (Rohatgi *et al.*, 2009).

2.2. Molecular cloning, heterologous expression and purification of recombinant SP_0149

Genomic DNA was isolated from ATCC BAA-334 using a commercially available kit following the manufacturer's instructions (Qiagen, Germany). A subfragment corresponding to amino acids 30–284 of the SP_0149 gene (GenBank accession number: NP_344691) was PCR amplified using genomic DNA, and GCAGACAATGCAACAACTATCAA and CCCCCC**AAGCTT**TTTACCAAACTGGTTGATCCAAAC as sense and antisense primer, respectively. A *Hin*dIII restriction site (bold) and an in-frame stop codon was engineered into the antisense primer. The signal sequence (amino acids 1–24) was excluded from the construct. The *Hin*dIII-digested PCR product was ligated into the *Stu*I–*Hin*dIII digested *Escherichia coli* expression vector pQE-30 Xa (Qiagen) using a quick ligation kit (New England BioLabs, USA). Ligation of SP_0149 into pQE-30 Xa results in the introduction of a six-histidine tag at the N-terminus of the recombinant protein. The ligated product was transformed in *E. coli* strain XL1 Blue. The recombinants were identified by restriction digestion and confirmed by nucleotide sequencing (Eurofins Genomics, India).

For expression purposes the recombinant construct was transformed in *E. coli* K-12 expression strain SG13009 (Qiagen) and was propagated in Luria–Bertani broth containing ampicillin and kanamycin at 100 and 25 µg ml^{-1}, respectively. Isopropyl β-D-1-thiogalactopyranoside (1 m*M*; Sigma, USA) was added to mid-logarithmic phase ($A_{600} \simeq 0.6$) *E. coli* culture to induce high-level expression of recombinant SP_0149 (rSP_0149) at 310 K for 2 h with aeration. Bacterial cells were pelleted by centrifuging at 6000*g* for 10 min, resuspended in lysis buffer (10 m*M* Tris, 300 m*M* NaCl, 10 m*M* imidazole, pH 8.0) and sonicated. The sonicate was centrifuged at 11 000*g* for 40 min at 277 K. rSP_0149 was purified from the supernatant using Ni-NTA affinity chromatography following the manufacturer's instructions (Sigma). The bound protein was eluted using a buffer containing 250 m*M* imidazole. The eluted fractions containing rSP_0149 protein were pooled and further purified using gel-filtration chromatography (GE Healthcare, USA). The purity of the rSP_0149 (in 10 m*M* Tris pH 8.0 and 50 m*M* NaCl) preparation was assessed by SDS–PAGE (Fig. 1).

2.3. Crystallization

Medium-throughput crystallization screening experiments were set up using commercially available screens: Crystal Screen and Crystal Screen 2 from Hampton (Hampton Research, USA) and JBScreen classic kits 1–10 covering 240 different conditions (Jena Bioscience, Germany). The crystallization experiments were performed in 96-well Intelli-plates using the sitting-drop vapour-diffusion technique and a Mosquito robot (TTP LABTECH, UK). Potential lead crystallization conditions showing growth of microcrystals were obtained from a few conditions: 10% polyethylene glycol (PEG) 8000, 100 m*M* MES sodium salt (pH 6.5) and 200 m*M* zinc acetate (JBScreen classic 4 condition D1); 30%(*w*/*v*) 2-propanol, 100 m*M* Tris–HCl pH 8.5 and 200 m*M* ammonium acetate (JBScreen classic 9 condition C3); 25%(*w*/*v*) *tert*-butanol, 100 m*M* Tris–HCl pH 8.5 and 100 m*M* $CaCl_2$ (JBScreen classic 9 condition C4). We expanded one such promising lead (JBScreen classic 4 condition D1) by varying the PEG 8000 concentration (while maintaining the concentration of the other components of the condition) from 5 to 20%, protein concentration from 20 to 40 mg ml^{-1} and the ratio of the protein to precipitant, using the hanging-drop vapour-diffusion technique. Diffraction quality crystals were obtained from the condition containing

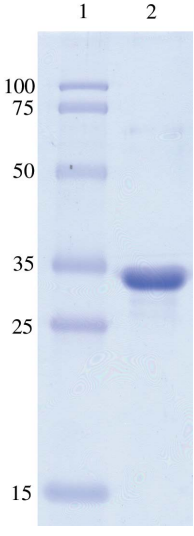


Figure 1
Purification of rSP_0149 protein. rSP_0149 was resolved by SDS–PAGE and visualized by Coomassie Brilliant Blue R-250 staining. Lane 1, the molecular mass marker (in kDa); Lane 2, purified rSP_0149.

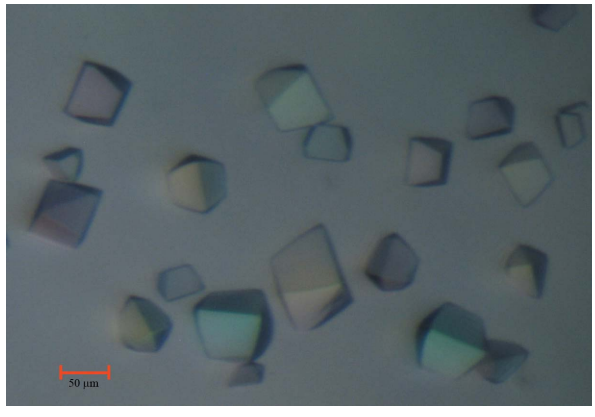


Figure 2
Crystals of rSP_0149 grown from the precipitant, 5% PEG 8000, 100 m*M* MES sodium salt (pH 6.5) and 200 m*M* zinc acetate. The scale bar represents 50 µm.

Table 1
Data collection statistics.

Values within parentheses refer to the highest resolution shell.

Unit-cell parameters (Å, °)	a = 54.56, b = 75.61, c = 75.52; $\alpha = \beta = \gamma = 90$
Space group	$P2_12_12_1$
Matthews coefficient (Å^3 Da^{-1})	2.56
Solvent content (%)	52.0
Temperature (K)	100
Detector	R-AXIS IV^{++}
Wavelength (Å)	1.5418
Resolution (Å)	50.00–2.30 (2.38–2.30)
Total number of reflections	57117 (5720)
Number of unique reflections	14495 (1430)
Multiplicity	3.9 (4.0)
$\langle I/\sigma(I)\rangle$	15.7 (3.5)
Completeness (%)	99.7 (99.9)
R_{sym} (%)†	10.4 (48.4)

† $R_{sym}(I) = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I(hkl)$ for n independent reflections and i observations of a given reflection. $\langle I(hkl)\rangle$ is the average intensity of the i observations.

rSP_0149 at 20 mg ml^{-1} in a buffer containing 10 m*M* Tris (pH 8.0) and 50 m*M* NaCl and the precipitant 5% PEG 8000, 100 m*M* MES sodium salt (pH 6.5) and 200 m*M* zinc acetate with 1:2.5 protein to precipitant ratio and 1 ml reservoir solution [5% PEG 8000, 100 m*M* MES sodium salt (pH 6.5) and 200 m*M* zinc acetate]. Crystals grew to an optimum size of approximately 50 × 45 × 30 µm in a week at 293 K (Fig. 2).

2.4. Intensity data collection and processing

A single crystal was mounted on a MicroMount (MiTeGen, USA) and rinsed in cryoprotectant solution [30%(*v*/*v*) glycerol in reservoir solution]. The crystal diffracted up to 2.3 Å resolution (Fig. 3). A native intensity data set was collected at 100 K with Cu Kα radiation using the in-house X-ray diffraction facility, a Rigaku FR-E+ SuperBright microfocus rotating anode dual-wavelength (Cu and Cr) X-ray generator mounted with R-AXIS IV^{++} detectors. The data set was indexed, integrated and scaled using *HKL*-2000 (Otwinowski & Minor, 1997). The detailed data-collection statistics are shown in the Table 1.

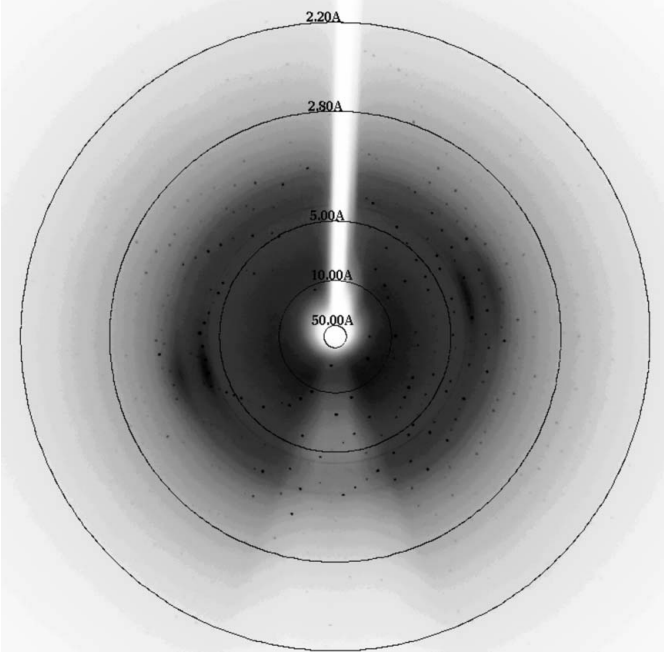


Figure 3
A diffraction image collected at 1° oscillation range from a single rSP_0149 crystal.

3. Results

rSP_0149 was purified to homogeneity by Ni-NTA affinity chromatography followed by gel-filtration chromatography (Fig. 1). The size of the protein that eluted from the gel-filtration column was consistent with it being a monomer of the truncated histidine-tagged SP_0149 construct. Optimization of a lead crystallization condition produced X-ray diffraction quality crystals (Fig. 2). A complete data set was collected up to 2.30 Å and was processed in orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 54.56, b = 75.61, c = 75.52 Å. Assuming the presence of one molecule of rSP_0149 (with calculated molecular weight of 30 400 Da) in the crystal asymmetric unit, the calculated values of the Matthews coefficient and the corresponding solvent content were 2.56 Å^3 Da^{-1} and 52.0% respectively (Matthews, 1968). Efforts are being made to solve the structure by the molecular replacement method with the programs *CCP*4 (Winn *et al.*, 2011) and *CNS* (Brünger *et al.*, 1998), using as a search model the structure of a similar protein from *Staphylococcus aureus* (Williams *et al.*, 2004; PDB code 1p99) which shares 29% amino-acid sequence identity with rSP_0149. Detailed results of these structural studies will be reported elsewhere.

The work reported in this paper was supported by an extramural research grant from the Department of Biotechnology (DBT), Government of India (to DS) and a start-up grant (to BKB) from the National Institute of Immunology, New Delhi, India. We thank DBT for funding the purchase of dual-wavelength X-ray generator. We thank Professor Avadesha Surolia for his constant encouragement for carrying out this work. We acknowledge Ravikant Pal for his help during data collection.

References

Basavanna, S., Khandavilli, S., Yuste, J., Cohen, J. M., Hosie, A. H., Webb, A. J., Thomas, G. H. & Brown, J. S. (2009). *Infect. Immun.* **77**, 3412–3423.
Bergmann, S. & Hammerschmidt, S. (2006). *Microbiology*, **152**, 295–303.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Davidson, A. L., Dassa, E., Orelle, C. & Chen, J. (2008). *Microbiol. Mol. Biol. Rev.* **72**, 317–364.
Garmory, H. S. & Titball, R. W. (2004). *Infect Immun.* **72**, 6757–6763.
Johnson, H. L., Deloria-Knoll, M., Levine, O. S., Stoszek, S. K., Freimanis Hance, L., Reithinger, R., Muenz, L. R. & O'Brien, K. L. (2010). *PLoS Med.* **7**, e1000348.
Locher, K. P. (2009). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 239–245.
Lynch, J. P., 3rd & Zhanel, G. G. (2010). *Curr. Opin. Pulm. Med.* **16**, 217–225.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McIntosh, E. D. & Reinert, R. R. (2011). *Expert Rev. Vaccines*, **10**, 109–129.
O'Brien, K. L., Wolfson, L. J., Watt, J. P., Henkle, E., Deloria-Knoll, M., McCall, N., Lee, E., Mulholland, K., Levine, O. S. & Cherian, T. (2009). *Lancet*, **374**, 893–902.
Oldham, M. L., Davidson, A. L. & Chen, J. (2008). *Curr. Opin. Struct. Biol.* **18**, 726–733.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rees, D. C., Johnson, E. & Lewinson, O. (2009). *Nat. Rev. Mol. Cell Biol.* **10**, 218–227.
Rohatgi, S., Dutta, D., Tahir, S. & Sehgal, D. (2009). *J. Immunol.* **182**, 5570–5585.
Williams, W. A., Zhang, R. G., Zhou, M., Joachimiak, G., Gornicki, P., Missiakas, D. & Joachimiak, A. (2004). *Biochemistry*, **43**, 16193–16202.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.