Teemu Haikarainen,^a Vuokko Loimaranta,^b Carlos Prieto-Lopez,^a Pradeep Battula,^a Jukka Finne^c and Anastassios C. Papageorgiou^{a*}

^aTurku Centre for Biotechnology, University of Turku and Åbo Akademi University, 20521 Turku, Finland, ^bDepartment of Medical Biochemistry and Genetics, University of Turku, Kiinamyllynkatu 10, 20520 Turku, Finland, and ^cDepartment of Biosciences, University of Helsinki, 00014 Helsinki, Finland

Correspondence e-mail:
tassos.papageorgiou@btk.fi

Received 11 March 2013
Accepted 9 April 2013

Expression, purification and crystallization of the C-terminal LRR domain of *Streptococcus pyogenes* protein 0843

Streptococcus pyogenes protein 0843 (Spy0843) is a recently identified protein with a potential adhesin function. Sequence analysis has shown that Spy0843 contains two leucine-rich repeat (LRR) domains that mediate interactions with the gp340 receptor. Here, the C-terminal LRR domain was overexpressed in *Escherichia coli*, purified and crystallized in the presence of 1.7–1.8 M ammonium sulfate pH 7.4 as precipitant. Data were collected from a single crystal to 1.59 Å resolution at 100 K at a synchrotron-radiation source. The crystal was found to belong to space group *I*₄, with unit-cell parameters *a* = *b* = 121.4, *c* = 51.5 Å and one molecule in the asymmetric unit. Elucidation of the crystal structure will provide insights into the interactions of Spy0843 with the gp340 receptor and a better understanding of the role of Spy0843 in streptococcal infections.

1. Introduction

Bacterial resistance to antibiotics has become a major threat to human health in recent years. As a result, new and more efficient approaches against bacterial diseases are urgently needed. In this regard, the use of anti-adhesive compounds has been proposed as an alternative strategy in antimicrobial therapy (Karlsson, 1998). The adhesion mechanisms in Gram-negative bacteria have been extensively investigated and are well understood. In contrast, little is known about adhesion mechanisms in Gram-positive bacteria, which include major pathogens such as *Streptococcus pyogenes* (Carapetis *et al.*, 2005) that causes pharyngitis, erysipelas, necrotizing fasciitis, rheumatic fever and glomerulonephritis.

S. pyogenes protein 0843 (Spy0843) is a recently identified protein (AAK33772; molecular weight ~110 kDa; Reid *et al.*, 2002) found in all *S. pyogenes* strains. The structure and the exact function of Spy0843 are unknown. It has been suggested that Spy0843 acts as an adhesin that mediates attachment to cell surfaces and induces invasion into cells. Antibodies against Spy0843 have been found to reduce *S. pyogenes* infection in mice (Reid *et al.*, 2002). Biological data have shown an interaction of Spy0843 with the gp340 receptor (salivary agglutinin), a member of the scavenger receptor cysteine-rich (SRCR) superfamily (Loimaranta *et al.*, 2005; Prakobphol *et al.*, 2000) found in secretions and at cell surfaces. Although some of the microbial interactions of gp340 have been characterized in detail, the basis of its wide ligand-recognition capacity has remained unclear. Sequence analysis with the SMART server (<http://smart.embl-heidelberg.de/>; Letunic *et al.*, 2012) has revealed that an ~550 amino-acid central part of Spy0843 contains two leucine-rich repeat (LRR) domains (Bella *et al.*, 2008). The specificity of the gp340–Spy0843 interactions has been shown to rely on the interactions of both LRRs in Spy0843 with a specific peptide sequence in gp340 (Loimaranta *et al.*, 2009), but the details are still unknown.

In an effort to better understand the role and function of Spy0843, we have initiated structural studies on the entire protein and its individual LRR domains. Here, we describe the successful expression, purification, crystallization and data collection to high resolution of the C-terminal LRR domain of Spy0843 (Spy-2; residues 405–680 of Spy0843). Structural insights into Spy0843 could help in characterizing common features in bacteria–gp340 interactions by using the gp340–Spy0843 complex as a model system.

2. Materials and methods

2.1. Expression and purification

2.1.1. Construction of the expression plasmid. The genomic DNA from *S. pyogenes* NZ131 was purified as described by Segers *et al.* (1998). Based on the genomic sequence of strain NZ131 (NC_011375), a pair of specific primers (F405, 5'-ACTAAATGGT-TAGAGGAAGATTT-3', and R680, 5'-TTTTTT**GTCGAC**TACT-CTACTTATCTGGATCAACGATAAAAT-3') was designed to amplify the genomic DNA sequence encoding the C-terminal LRR domain of Spy0843. The PCR product was digested with *SalI* (the *SalI* recognition site is shown in bold) and subcloned into the *StuI*- and *SalI*-digested expression vector pQE 30Xa (Qiagen). The resulting recombinant plasmid was designated pQE-Spy2.

2.1.2. Transformation. The pQE-Spy2 was first transformed into *Escherichia coli* JM109. Transformants were selected on ampicillin agar and confirmed to be positive in colony PCR with the primers F405 and R680. Plasmid DNA was purified from one clone with a plasmid miniprep kit (Qiagen) and shown to have the correct PCR insert by restriction digestion analysis and DNA sequencing. For expression, the plasmid was transformed into *E. coli* M15 [pREP4] (Qiagen).

2.1.3. Purification of Spy-2. 2×10 ml cultures of Spy-2 in M15 cells were grown at 310 K overnight in LB medium supplemented with $25 \mu\text{g ml}^{-1}$ kanamycin and $100 \mu\text{g ml}^{-1}$ ampicillin. The culture was diluted to 500 ml (2×500 ml) of LB with the antibiotics and grown until the OD_{600} reached ~ 0.6 . Induction was carried out with 0.4 mM IPTG at 310 K for 5 h. The cells were collected by centrifugation (3000g, 20 min) and the pellets were stored at 253 K. For purification, the pellets were suspended in ~ 20 ml lysis buffer (50 mM sodium phosphate buffer pH 7.4, 0.5 M NaCl, 10% glycerol, 0.1% Triton X-100, 10 mM imidazole) supplemented with 1 mM PMSF, 5 mM MgCl_2 , 0.5 mg ml^{-1} lysozyme and 0.025 mg ml^{-1} DNase I. The solution was incubated on ice for 30 min, sonicated and centrifuged at 31 000g for 20 min. A 2 ml suspension of Ni-NTA resin was added to the cleared lysate and incubated with gentle rotation at 277 K for 45 min. The suspension was loaded onto a gravity-flow column and the flowthrough was collected. The column was washed

with 40 ml 50 mM sodium phosphate buffer pH 7.4, 0.5 M NaCl, 10% glycerol, 20 mM imidazole. The bound protein was eluted with 50 mM sodium phosphate buffer pH 7.4, 0.5 M NaCl, 250 mM imidazole. Fractions of 1 ml each were collected. The protein was concentrated to ~ 1 –1.5 ml and purified further by gel filtration (HiPrep Sephacryl S-200 column) with 10 mM Tris-HCl pH 7.5, 150 mM NaCl. Only the peak fractions were collected. The volume of each fraction was 2 ml. The purity of the recombinant Spy-2 was assessed by SDS-PAGE (Fig. 1).

2.2. Crystallization

Purified Spy-2 was concentrated by ultrafiltration using Amicon Ultra Centrifugal Filters (10 000 Da molecular-weight cut-off; Millipore) in a buffer consisting of 10 mM Tris-HCl pH 7.5, 150 mM NaCl. The protein concentration used in crystallization was adjusted to 10 mg ml^{-1} as measured by a Nanodrop 2000c spectrophotometer

Figure 2
A typical crystal of Spy-2. The dimensions of this crystal are approximately $0.35 \times 0.15 \times 0.15$ mm.

Figure 1
Purity of Spy-2 in 12% SDS-PAGE. Left lane, molecular-weight protein markers (labelled in kDa); right lane, purified Spy-2 visualized by Coomassie Brilliant Blue staining.

Figure 3
Representative diffraction image of the Spy-2 crystal used for data collection. The resolution rings are shown. A diffraction spot near the edge of the image is shown in the inset.

Table 1

Crystal parameters and data-processing statistics.

Values in parentheses are for the outermost resolution shell.

Space group	$I4_1$
Unit-cell parameters (Å)	$a = b = 121.4, c = 51.5$
Resolution range (Å)	20.0–1.59 (1.68–1.59)
Wavelength (Å)	0.8123
Temperature (K)	100
Total reflections	151837 (23142)
Unique reflections	49836 (7887)
Completeness (%)	97.8 (97.0)
Mosaic spread (°)	0.142
$R_{\text{merge}}^{\dagger}$ (%)	3.4 (26.1)
$R_{\text{meas}}^{\ddagger}$ (%)	4.1 (32.2)
$\langle I/\sigma(I) \rangle$	23.6 (4.3)

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the i th observation of reflection hkl and $\langle I(hkl) \rangle$ is the weighted average intensity for all observations i of reflection hkl . $\ddagger$ Redundancy-independent R value calculated as $R_{\text{meas}} = \sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $N(hkl)$ denotes the redundancy (Diederichs & Karplus, 1997).

(Thermo Scientific). Initial crystallization trials were performed with the Index (Hampton Research), PACT (Qiagen) and SaltRx (Hampton Research) crystal screens using the sitting-drop vapour-diffusion method at 289 K in 96-well Corning plates. The drops consisted of 0.75 μ l protein solution and an equal volume of precipitant solution. Crystals of Spy-2 appeared in several conditions of the Index screen. Initial characterization with an in-house X-ray source showed conditions No. 5 (0.1 M HEPES pH 7.5, 2 M ammonium sulfate) and No. 6 (0.1 M Tris-HCl pH 8.5, 2 M ammonium sulfate) as those with the most promising crystals. Optimization in a hanging-drop vapour-diffusion setup on 24-well Linbro plates improved the crystal size. The drops consisted of 2 μ l protein solution and 2 μ l reservoir solution. The best crystals were grown in 1.7–1.8 M ammonium sulfate, 0.1 M HEPES-NaOH pH 7.4 (Fig. 2).

2.3. Data collection

A single Spy-2 crystal was used for data collection after it was briefly soaked in a reservoir solution containing 20% (v/v) glycerol as a cryoprotectant. The crystal was subsequently placed in a stream of gaseous N_2 at 100 K. A total of 186 images (Fig. 3) were recorded on a MAR CCD (station X13 at EMBL-Hamburg c/o DESY). The program *BEST* (Bourenkov & Popov, 2010) was utilized for optimization of the data-collection strategy. The X-ray wavelength was 0.81230 Å, the crystal-to-detector distance was set to 141 mm and the rotation range was adjusted to 0.3°. All data were processed, integrated and scaled with *XDS* (Kabsch, 2010).

3. Results and discussion

3.1. Expression and purification

The expressed protein was purified using Ni-NTA affinity chromatography followed by a gel-filtration step. SDS-PAGE gave a major band with a molecular mass of about 33 kDa as expected from the amino-acid sequence of Spy-2 and the extended N-terminus (RGSHHHHHHSGSGSGIEGR) owing to the His₆-tag sequence and additional residues from the expression vector.

3.2. Crystallization of Spy-2

Optimization of the crystallization condition produced crystals of suitable size for further crystallographic studies. The crystals diffracted to 2.8 Å in-house and a complete data set was collected to 1.59 Å on a synchrotron beamline. The overall completeness was 97.8% with an R_{merge} of 3.4%. The crystals belonged to the tetragonal space group $I4_1$, with unit-cell parameters $a = b = 121.4, c = 51.5$ Å. Final statistics are summarized in Table 1. Assuming 1 molecule of the protein in the asymmetric unit and a molecular mass of 33 kDa, the Matthews coefficient V_M (Matthews, 1968) is 2.82 Å³ Da⁻¹, corresponding to a solvent content of ~56.4%. A survey of the Protein Data Bank revealed no structures of homologous proteins suitable for molecular replacement. Consequently, structure determination of Spy-2 will be pursued by single-wavelength or multi-wavelength anomalous dispersion methods. For the latter, the three Met residues present in Spy-2 will be substituted by selenomethionine. The structure of Spy-2 will facilitate further studies on the interactions of the protein with the gp340 receptor. The high resolution of the current data would allow a better characterization of the LRR geometric features such as bending and twist. Currently, only two LRR structures are at higher resolution than Spy-2: the structure of *Listeria monocytogenes* internalin at 1.50 Å (Schubert *et al.*, 2002) and that of a *Bacteroides ovatus* hypothetical leucine-rich repeat protein at 1.19 Å resolution (Joint Center for Structural Genomics, unpublished work). Efforts to express and crystallize larger fragments of the Spy0843 protein are also in progress.

We thank the Academy of Finland (grant 121278 to ACP), Sigrid Juselius Foundation, Magnus Ehrnrooth Foundation, University of Turku Foundation and Biocenter Finland for financial and infrastructure support. Access to EMBL-Hamburg c/o DESY was provided by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement 226716 and BioStruct-X (grant agreement 283570).

References

- Bella, J., Hindle, K. L., McEwan, P. A. & Lovell, S. C. (2008). *Cell. Mol. Life Sci.* **65**, 2307–2333.
- Bourenkov, G. P. & Popov, A. N. (2010). *Acta Cryst.* **D66**, 409–419.
- Carapetis, J. R., Steer, A. C., Mulholland, E. K. & Weber, M. (2005). *Lancet Infect. Dis.* **5**, 685–694.
- Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Karlsson, K. A. (1998). *Mol. Microbiol.* **29**, 1–11.
- Letunic, I., Doerks, T. & Bork, P. (2012). *Nucleic Acids Res.* **40**, D302–D305.
- Loimaranta, V., Hytönen, J., Pulliainen, A. T., Sharma, A., Tenovuo, J., Strömberg, N. & Finne, J. (2009). *J. Biol. Chem.* **284**, 18614–18623.
- Loimaranta, V., Jakubovics, N. S., Hytönen, J., Finne, J., Jenkinson, H. F. & Strömberg, N. (2005). *Infect. Immun.* **73**, 2245–2252.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Prakobphol, A., Xu, F., Hoang, V. M., Larsson, T., Bergstrom, J., Johansson, I., Frängsmyr, L., Holmskov, U., Leffler, H., Nilsson, C., Borén, T., Wright, J. R., Strömberg, N. & Fisher, S. J. (2000). *J. Biol. Chem.* **275**, 39860–39866.
- Reid, S. D., Green, N. M., Sylva, G. L., Voyich, J. M., Stenseth, E. T., DeLeo, F. R., Palzkill, T., Low, D. E., Hill, H. R. & Musser, J. M. (2002). *J. Bacteriol.* **184**, 6316–6324.
- Schubert, W. D., Urbanke, C., Ziehme, T., Beier, V., Machner, M. P., Domann, E., Wehland, J., Chakraborty, T. & Heinz, D. W. (2002). *Cell*, **111**, 825–836.
- Segers, R. P., Kenter, T., de Haan, L. A. & Jacobs, A. A. (1998). *FEMS Microbiol. Lett.* **167**, 255–261.