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

Purification, crystallization and preliminary crystallographic analysis of the minor pilin FctB from *Streptococcus pyogenes*

Christian Linke,[a,b] Paul G. Young,[a,b] Hae Joo Kang,[a,b] Thomas Proft[b,c] and Edward N. Baker[a,c]*

[a]School of Biological Sciences, University of Auckland, Auckland, New Zealand, [b]Maurice Wilkins Centre for Molecular Biodiscovery, Auckland, New Zealand, and [c]School of Medical Sciences, University of Auckland, Auckland, New Zealand

Correspondence e-mail: ted.baker@auckland.ac.nz

Received 6 November 2009
Accepted 2 December 2009

The minor pilin FctB is an integral part of the pilus assembly expressed by *Streptococcus pyogenes*. Since it is located at the cell wall, it can be hypothesized that it functions as a cell-wall anchor for the streptococcal pilus. In order to elucidate its structure, the genes for FctB from the *S. pyogenes* strains 90/306S and SF370 were cloned for overexpression in *Escherichia coli*. FctB from strain 90/306S was crystallized by the sitting-drop vapour-diffusion method using sodium citrate as a precipitant. The hexagonal FctB crystals belonged to space group $P6_1$ or $P6_5$, with unit-cell parameters $a = b = 95.15$, $c = 100.25$ Å, and diffracted to 2.9 Å resolution.

1. Introduction

The Gram-positive pathogen *Streptococcus pyogenes* (group A streptococcus; GAS) is a major cause of throat and skin diseases such as pharyngitis, tonsillitis and erysipelas. It also leads to life-threatening conditions such as necrotizing fasciitis and streptococcal toxic shock syndrome (Cunningham, 2000). Recently, it was recognized that *S. pyogenes* possesses pili, which are thin hair-like appendages (Mora *et al.*, 2005). These pili were shown to be essential for the adhesion of *S. pyogenes* to human tonsil epithelia (Abbot *et al.*, 2007).

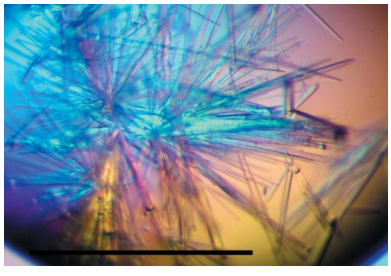
The streptococcal pilus is built from three proteins: a so-called major pilin and two minor pilins (Mora *et al.*, 2005). The major pilin is polymerized into the main shaft of the pilus by covalent linkage of adjacent molecules (Mora *et al.*, 2005; Kang *et al.*, 2007), while the minor pilin Cpa forms the tip of the pilus (Quigley *et al.*, 2009). The role of the second minor pilin, FctB (Spy0130 in *S. pyogenes* strain SF370), is not clear. It is covalently incorporated into the pilus structure but, in contrast to the major pilin and Cpa, it is localized at the cell wall (Mora *et al.*, 2005). Therefore, we hypothesize that FctB (Spy0130) could be the cell-wall anchor of the *S. pyogenes* pilus, analogous to the cell-wall anchor SpaB of *Corynebacterium diphtheriae* (Mandlik *et al.*, 2008).

In order to further characterize the role and structure of this minor pilin, we have cloned and purified it from two different group A streptococcus strains: FctB from *S. pyogenes* strain 90/306S and Spy0130 from strain SF370, which share 30% sequence identity. A full-length construct of FctB yielded hexagonal crystals that diffracted to 2.9 Å resolution. The structure of FctB should provide new insights into the structure and assembly of the pilus from *S. pyogenes*.

2. Methods

2.1. Cloning, expression and purification of FctB and Spy0130

Genomic DNA from *S. pyogenes* serotype M5 strain 90/306S (isolated from a patient in New Zealand) and serotype M1 strain SF370 was purified according to existing protocols (Proft *et al.*, 2000). Briefly, *S. pyogenes* was grown overnight in 20 ml brain heart infusion medium (BHI; Difco) at 210 K without agitation. Sedimented cells were washed in 1 ml lysis buffer (10 m*M* Tris–HCl pH 8, 50 m*M* EDTA) and resuspended in 400 µl lysis buffer containing 4 U mutanolysin (Sigma–Aldrich), 40 µg RNaseA and 0.5 mg lysozyme. After incubating the samples for 1 h at 210 K with gentle agitation, 20 µg proteinase K was added and incubated for an additional 30 min.



Cells were lysed by the addition of sarcosyl [to a final concentration of 2%(*v*/*v*)] and the DNA was isolated by phenol/chloroform extraction followed by ethanol precipitation.

The gene for the mature cell-wall-anchored domain of FctB, extending from the signal peptidase site predicted by *SignalP* (Emanuelsson *et al.*, 2007) to the sortase cleavage site, was PCR-amplified from *S. pyogenes* strain 90/306S genomic DNA with Pfx polymerase (Invitrogen). As there were no sequence data available for strain 90/306S, primers were designed based on the genomic DNA sequence of *S. pyogenes* strain Manfredo (locus *spyM50106a*; NCBI GeneID 4964555; Holden *et al.*, 2007). Thus, the gene-specific primers used were M5_spy130_F1, 5′-AAA **GGC GCC** AAA GAC AGC ACT GTT CAA ACT AGC-3′, and M5_spy130_R1b, 5′-AAA **GAA TTC** TTA AGC TAA TGG TAA TGG GGT TTT AGG-3′ (5′ *Nar*I and 3′ *Eco*RI restriction-nuclease recognition sites are shown in bold). The amplified fragment was digested with *Nar*I and *Eco*RI and cloned into the respective sites in the multiple cloning site of vector pProEX HTa (Invitrogen). A truncated version of the extracellular domain of FctB (FctB$^{\Delta}$), lacking a 17-residue proline-rich C-terminal sequence, was cloned in the same way using the primers M5_spy0130_F1 and M5_spy0130_R4, 5′-AAA **GAA TTC** TTA TTT TAC TAA CCG TTT GGG CTT AAA AG-3′. The resulting plasmids (pProEX_HTa_FctB and pProEx_HTa_FctB$^{\Delta}$) were used to transform *Escherichia coli* (DH5α) cells and the nucleotide sequences were verified. The nucleotide sequence for FctB was deposited in the NCBI GenBank with accession number GU250526.

For FctB and FctB$^{\Delta}$ protein expression, *E. coli* BL21 (DE3) pRIL cells (Stratagene) were transformed with pProEX_HTa_FctB or pProEx_HTa_FctB$^{\Delta}$ and were grown in Luria–Bertani (LB) medium supplemented with the required antibiotics at 310 K until the OD_{600} reached 1.0. The cultures were induced with 0.2 m*M* IPTG at 291 K for 16 h and the cells were harvested by centrifugation. Cell pellets were resuspended in lysis buffer [50 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl, 2%(*v*/*v*) glycerol, 5 m*M* imidazole, 0.3 m*M* NaN_3] containing Complete Protease Inhibitor Cocktail Mini EDTA-free tablets (Roche), snap-frozen and stored at 253 K.

The recombinant proteins were purified from frozen cells, which were thawed in lysis buffer with Complete Protease Inhibitor Cocktail Mini EDTA-free tablets (Roche) and 2 µg ml^{-1} DNase I and lysed using a cell disruptor at 117 MPa (Constant Systems). Insoluble matter was sedimented by centrifugation (55 000*g*, 277 K, 45 min). The soluble phase was loaded onto a HiTrap Chelating 5 ml column (GE Healthcare) charged with $NiCl_2$. Bound protein was washed with wash buffer (lysis buffer + 20 m*M* imidazole) and eluted in a gradient with elution buffer (lysis buffer + 500 m*M* imidazole). FctB or FctB$^{\Delta}$ protein was dialysed overnight against TBS (50 m*M* Tris–HCl pH 7.4, 137 m*M* NaCl, 2.7 m*M* KCl, 0.3 m*M* NaN_3) and the His_6 tag was concomitantly removed using recombinant TEV protease. Undigested protein and TEV protease were removed by a second round of IMAC. In a final step, FctB or FctB$^{\Delta}$ protein was subjected to size-exclusion chromatography on a Superdex S75 10/300 column (GE Healthcare) in crystallization buffer (10 m*M* Tris–HCl pH 7.4, 137 m*M* NaCl, 2.7 m*M* KCl, 0.3 m*M* NaN_3). FctB and FctB$^{\Delta}$ eluted in a single peak, were 99% pure based on SDS–PAGE and were monodisperse as indicated by dynamic light-scattering (DLS) data. The final FctB and FctB$^{\Delta}$ constructs contained residues 1–139 and 1–122 of FctB (NCBI accession No. GU250526), respectively, preceded by an N-terminal Gly-Ala extension derived from the vector. The identity of these proteins was confirmed by MALDI–TOF mass spectrometry by the University of Auckland Centre for Genomics and Proteomics (FctB observed, 15 449.1 $\pm$ 0.5 Da; expected, 15 449.4 Da; FctB$^{\Delta}$ observed, 13 583.0 $\pm$ 0.5 Da; expected, 13 583.2 Da).

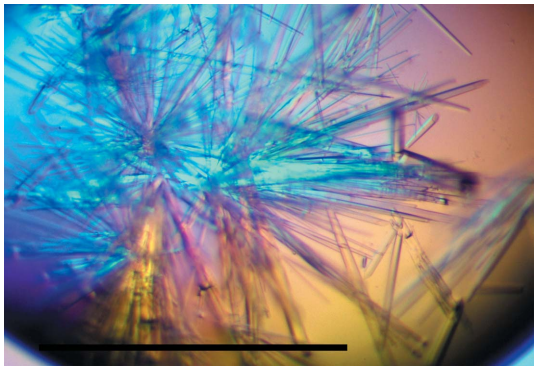


Figure 1
Crystals of FctB. The scale bar indicates 500 µm.

The gene for the mature cell-wall-anchored domain of Spy0130 (locus *spy_0130*; NCBI GeneID 900462), extending from the signal peptidase site predicted by *SignalP* (Emanuelsson *et al.*, 2007) but without the sortase cleavage site, was amplified by PCR from *S. pyogenes* serotype M1 (strain SF370) genomic DNA (Ferretti *et al.*, 2001) using the primers Spy0130_forw, 5′-**GGA TCC** GAG AAC CTC ACT GCA AGC-3′, and Spy0130_rev, 5′-**GAA TTC** CTA TTT GCT TGT CTT TAG C-3′ (5′ *Bam*HI and 3′ *Eco*RI restriction nuclease recognition sites are shown in bold). The amplified product was cloned into the *Eco*RI and *Bam*HI restriction sites of the pGEX-3c expression vector (Proft *et al.*, 2001). The resulting vector pGEX-3c_Spy0130 was used to transform *E. coli* BL21 (DE3) pRP (Stratagene).

For expression of GST-tagged Spy0130, cells were grown in ZYP-5052 autoinduction medium (Studier, 2005) for an initial 4 h at 310 K followed by 20 h at 301 K. Harvested cells were resuspended in a buffer containing 25 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl (buffer *A*) and lysed using a cell disruptor at 117 MPa (Constant Systems). The lysate was loaded onto a Glutathione Sepharose 4B (GE Healthcare) column that had been equilibrated at 277 K with buffer *A* containing 0.5 m*M* EDTA. Spy0130 was cleaved from the GST tag bound to the column by incubating overnight with buffer *A* containing recombinant picornavirus 3C protease and 2 m*M* DTT. The untagged Spy0130 was eluted from the column by passage of buffer *A*. The Spy0130 protein was further purified by size-exclusion chromatography using a Superdex S200 10/300 gel-filtration column (GE Healthcare) in buffer *A*. The final product contained amino acids 27–179 of Spy0130 (NCBI accession No. NP_268519.1) preceded by four additional N-terminal residues that remained after removal of the GST tag (Gly-Pro-Gly-Ser). The purified Spy0130 protein was monomeric and monodisperse in solution, as determined by DLS. While our work was in progress, biophysical studies of another construct of Spy0130 comprising residues 27–185 were published (Solovyova *et al.*, 2009).

2.2. Crystallization

Vapour-diffusion crystallization trials were carried out at 291 K on FctB, FctB$^{\Delta}$ and Spy0130 using a Cartesian nanolitre dispensing robot (Genomic Systems) and a locally compiled crystallization screen (Moreland *et al.*, 2005). The FctB and FctB$^{\Delta}$ proteins were concentrated to 27 mg ml^{-1} based on an absorption coefficient ε of 8940 M^{-1} cm^{-1} (Gasteiger *et al.*, 2005) and Spy0130 to 9, 20 and 40 mg ml^{-1} (ε of 11 920 M^{-1} cm^{-1}). Only FctB yielded crystals, which were obtained from 100 m*M* sodium cacodylate pH 6.5, 1 *M* sodium

citrate. The FctB crystals, which had a hexagonal rod-like shape, were optimized by screening around this condition in a hanging-drop vapour-diffusion format. The pH, sodium citrate concentration and drop size were varied. The crystals used for X-ray data collection (Fig. 1) were grown by mixing 1–2 µl protein solution with 1–2 µl precipitant (100 m*M* sodium cacodylate pH 6.1–6.5, 0.9–1 *M* sodium citrate) after 7–10 d at 291 K.

2.3. Collection of X-ray diffraction data

Crystals of FctB were transferred to cryoprotectant [100 m*M* sodium cacodylate pH 6.5, 1 *M* sodium citrate, 20%(*v*/*v*) glycerol] prior to flash-freezing in liquid nitrogen. A complete X-ray diffraction data set was collected in-house (Micromax-007HF, Rigaku; MAR345DTB, MAR Research) at 100 K. Diffraction images (Fig. 2) were integrated using *XDS* (Kabsch, 1993), reindexed using *POINTLESS* (Evans, 2006) and scaled using *SCALA* (Evans, 2006). Data-collection statistics are given in Table 1.

3. Results and discussion

Biophysical studies of Spy0130 have indicated that the molecule contains significantly disordered regions that were predicted to be detrimental to protein crystallization (Solovyova *et al.*, 2009). For this reason, we also expressed the equivalent protein, FctB, from a different strain of *S. pyogenes*. FctB has a proline-rich region at its C-terminus that is eight residues shorter compared with that of Spy0130, making it potentially better ordered. To maximize the chances of successful crystallization, we also prepared a truncated version lacking this proline-rich region. Surprisingly, it was the longer construct that crystallized.

X-ray diffraction data were collected to a resolution of 2.9 Å in-house from the FctB crystals. Based on the diffraction symmetry and systematic absences, the crystals were found to be hexagonal, with space group $P6_1$ or $P6_5$ and unit-cell parameters $a = b = 95.15$, $c = 100.25$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. The Matthews coefficient of 2.13 Å^3 Da^{-1} suggested the presence of four molecules per asymmetric unit and a solvent content of 42.4%. The *L* test and Britton analysis (Britton, 1972; Padilla & Yeates, 2003) performed with *phenix.xtriage* (Zwart *et al.*, 2005) indicated that twinning of the FctB crystals was highly unlikely.

We could find no suitable model for molecular replacement. Attempts to prepare crystals of selenomethionine-substituted FctB are currently under way.

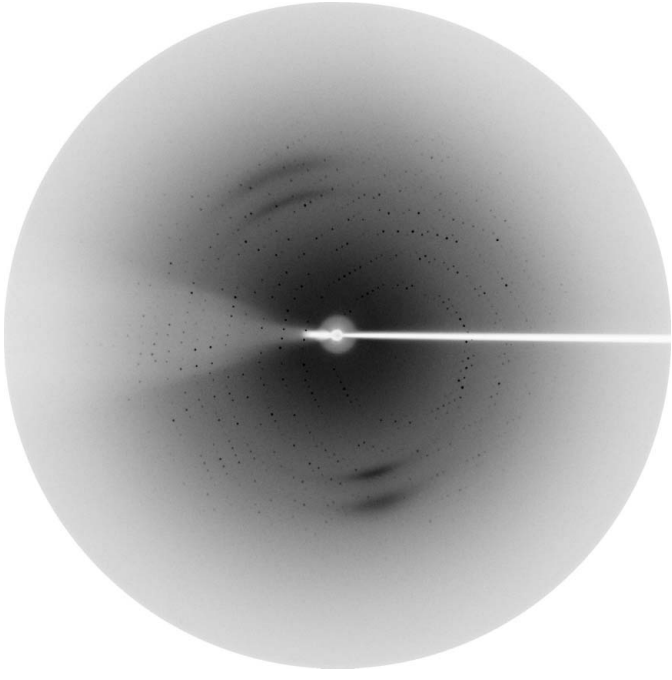


Figure 2
X-ray diffraction pattern of an FctB crystal. The resolution limit at the edge of the image is 2.2 Å.

Table 1
Data-collection statistics.

Values in parentheses are for the outermost resolution shell.

X-ray source	Cu $K\alpha$ anode
Wavelength (Å)	1.5418
Resolution range (Å)	31.15–2.90 (3.05–2.90)
Space group	$P6_1$ or $P6_5$
Unit-cell parameters (Å, °)	$a = b = 95.15$, $c = 100.25$, $\alpha = \beta = 90$, $\gamma = 120$
Total No. of observations	143360 (19699)
Unique reflections	11506 (1636)
Redundancy	12.5
Completeness (%)	99.7 (98.1)
Mean $I/\sigma(I)$	27.5 (7.6)
$R_{\rm merge}$† (%)	9.5 (36.5)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

CL was supported by a New Zealand International Doctoral Research Scholarship (Education New Zealand) and a University of Auckland NZIDRS Plus Scholarship. This work was supported by the Marsden Fund of New Zealand, the Health Research Council of New Zealand and the Tertiary Education Commission through funding of the Maurice Wilkins Centre for Molecular Biodiscovery.

References

Abbot, E. L., Smith, W. D., Siou, G. P. S., Chiriboga, C., Smith, R. J., Wilson, J. A., Hirst, B. H. & Kehoe, M. A. (2007). *Cell. Microbiol.* **9**, 1822–1833.
Britton, D. (1972). *Acta Cryst.* A**28**, 296–297.
Cunningham, M. W. (2000). *Clin. Microbiol. Rev.* **13**, 470–511.
Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. (2007). *Nature Protoc.* **2**, 953–971.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Ferretti, J. J. *et al.* (2001). *Proc. Natl Acad. Sci. USA*, **98**, 4658–4663.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. M. Walker, 571–607. Totowa: Humana Press.
Holden, M. T. G. *et al.* (2007). *J. Bacteriol.* **189**, 1473–1477.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Kang, H. J., Coulibaly, F., Clow, F., Proft, T. & Baker, E. N. (2007). *Science*, **318**, 1625–1628.
Mandlik, A., Das, A. & Ton-That, H. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 14147–14152.
Mora, M., Bensi, G., Capo, S., Falugi, F., Zingaretti, C., Manetti, A. G. O., Maggi, T., Taddei, A. R., Grandi, G. & Telford, J. L. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 15641–15646.
Moreland, N., Ashton, R., Baker, H. M., Ivanovic, I., Patterson, S., Arcus, V. L., Baker, E. N. & Lott, J. S. (2005). *Acta Cryst.* D**61**, 1378–1385.
Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* D**59**, 1124–1130.
Proft, T., Arcus, V. L., Handley, V., Baker, E. N. & Fraser, J. D. (2001). *J. Immunol.* **166**, 6711–6719.
Proft, T., Moffatt, S. L., Weller, K. D., Paterson, A., Martin, D. & Fraser, J. D. (2000). *J. Exp. Med.* **191**, 1765–1776.
Quigley, B. R., Zähner, D., Hatkoff, M., Thanassi, D. G. & Scott, J. R. (2009). *Mol. Microbiol.* **72**, 1379–1394.
Solovyova, A. S., Pointon, J. A., Race, P. R., Smith, W. D., Kehoe, M. A. & Banfield, M. J. (2009). *Eur. Biophys. J.*, doi:10.1007/s00249-009-0432-2.
Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
Zwart, P. H., Grosse-Kunstleve, R. W. & Adams, P. D. (2005). *CCP4 Newsl.* **43**, contribution 7.