Recombinant production, crystallization and preliminary structural characterization of *Schistosoma japonicum* profilin

Nele Vervaeet,^a Juha Pekka Kallio,^a Susanne Meier,^a Emilia Salmivaara,^{a,b} Maike Eberhardt,^b Shuangmin Zhang,^c Xi Sun,^c Zhongdao Wu,^c Petri Kursula^{b,d,e} and Inari Kursula^{a,b*}

^aCentre for Structural Systems Biology, Helmholtz Centre for Infection Research and German Electron Synchrotron (DESY), Notkestrasse 85, Bldg. 25b, 22607 Hamburg, Germany, ^bDepartment of Biochemistry, University of Oulu, PO Box 3000, 90014 Oulu, Finland, ^cDepartment of Parasitology, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, People's Republic of China, ^dBiocenter Oulu, University of Oulu, PO Box 3000, 90014 Oulu, Finland, and ^eDepartment of Chemistry, University of Hamburg, DESY, Notkestrasse 85, Bldg. 25b, 22607 Hamburg, Germany

Correspondence e-mail:
inari.kursula@helmholtz-hzi.de

Received 4 September 2013

Accepted 24 September 2013

Helminthic parasites of the genus *Schistosoma* contain a tegumental membrane, which is of crucial importance for modulation of the host immune response and parasite survival. The actin cytoskeleton plays an important role in the function of the tegument. Profilins are among the most important proteins regulating actin dynamics. *Schistosoma japonicum* possesses one profilin-like protein, which has been characterized as a potential vaccine candidate. Notably, profilins are highly immunogenic molecules in many organisms. Here, the profilin from *S. japonicum* was expressed, purified and crystallized. A native data set to 1.91 Å resolution and a single-wavelength anomalous diffraction (SAD) data set to a resolution of 2.2 Å were collected. The crystals belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 31.82$, $b = 52.17$, $c = 59.79$ Å and $\alpha = 35.29$, $\beta = 52.15$, $\gamma = 59.82$ Å, respectively.

1. Introduction

Schistosomiasis, also known as bilharzia, is a devastating disease that poses a large health and socio-economic threat to developing countries in (sub)tropical regions. It is a chronic disease that affects many individuals with long-standing infections in poor rural areas (Engels *et al.*, 2002). The parasite is a zoonotic species that infects a wide range of organisms, including a large number of mammals (Gryseels *et al.*, 2006). The various species of *Schistosoma* have a complicated digenic life cycle, during which they use freshwater snails as an intermediate host and humans or other mammals as the definitive host (McManus & Loukas, 2008).

Although treatable with praziquantel, schistosomiasis is becoming an increasingly severe problem because of increasing drug resistance. Schistosomes can survive for extended times in the blood circulation of their host, despite its specific immune response (Pearce & MacDonald, 2002). The tegumental outer surface of the parasite is a unique double-membrane structure that covers the entire surface of the worm, forming the major contact interface between the parasite and its host. In addition to host recognition and modulation of the immune response, the tegument also plays an important role in parasite survival (Jones *et al.*, 2004). As the initial contact between the parasite and the host takes place *via* the tegument, its components are major antigens for the host immune system. Because of these crucial functions for the survival of the parasite, tegumental proteins are interesting targets for both drug and vaccine development.

The actin cytoskeleton plays an important role in the formation and the modulation of the tegument in helminthic parasites such as *S. japonicum*. Profilins are among the most important regulators of actin dynamics (Yarmola & Bubb, 2006). They are small monomeric (G) actin-binding proteins that also bind various filamentous (F) actin-binding regulatory proteins that contain a variable number of proline-rich sequence repeats and function in processes related to microfilament nucleation and elongation. In addition, profilins bind to membrane phospholipids and have been suggested to participate in membrane-trafficking and signalling events (Witke, 2004). *S. japonicum* possesses one profilin-like protein, which has been characterized as a potential vaccine candidate (Zhang *et al.*, 2008). Notably, profilins are highly immunogenic molecules in many organisms and are among the most prominent allergens in plants

© 2013 International Union of Crystallography
All rights reserved

Table 1

Recombinant-protein production information.

The *SjPfn* sequence is indicated in bold and the 6×His tag and SUMO3 tag in italics.

Protein	<i>SjPfn</i>
Source organism	<i>S. japonicum</i>
DNA source	cDNA
Forward primer (vector)	5'-AGCACCACCACCACCACCAC
Reverse primer (vector)	5'-CCCTCCCGTCTGCTGCTGGA
Forward primer (insert)	5'-TGTTCAGCAGCAGACGGGAGGGATGAGC-GCTGATAGTTGGG
Reverse primer (insert)	5'-GTGGTGGTGGTGGTGGTCTTTAGTAACCC-ATTGCTCGTAATG
Expression vector	pETNKI-hisSUMO3-LIC (NKI)
Expression host	<i>E. coli</i> BL21 (DE3) RIPL
Complete amino-acid sequence of the construct produced	MGHHHHHHGGMSEKPKGVKTENDHINLKV-AGQDGSVVQFKIKRHTPLSKLMKAYCERQG-LSMRQIRFRFDGQPINETDTPAQLEMEDED-IDVFQQTGGMSADSWDNHCVTYVANNKC-LKNLCMTAIDGSHLGTSPDFRIPPELILQLK-SILDGGLDTSIFFMGEKYIVLQHDSSCLVSRG-GKKSILFYATGKICLVGQTVDDQNNCTQG-NFAISMRDHYERMGY

(Plattner *et al.*, 2008; Santos & Van Ree, 2011; Gadermaier *et al.*, 2013).

We have expressed and purified the recombinant *S. japonicum* profilin-like protein (hereafter referred to as *SjPfn*). The purified protein was folded, as analyzed by circular-dichroism (CD) spectroscopy, and was successfully crystallized together with iodide to obtain anomalous signal for experimental phasing.

2. Materials and methods

2.1. Macromolecule production

The *SjPfn* cDNA (Zhang *et al.*, 2008) was cloned into a prokaryotic expression vector using a sequence- and ligation-independent method (SLIC; Li & Elledge, 2012). The pETNKI-hisSUMO3-LIC vector (NKI Protein Facility, The Netherlands) and the coding region of *SjPfn* were amplified by the polymerase chain reaction (PCR) with the primers described in Table 1. *DpnI* was added to the PCR products to digest the template DNA, after which the products were purified from an agarose gel using a QIAquick gel-extraction kit (Qiagen). To create single-stranded overhangs, the resulting constructs were treated separately with T4 polymerase (New England BioLabs) at 296 K for 20 min. The reaction was terminated by adding a 1/20 volume of 500 mM ethylenediaminetetraacetic acid (EDTA) followed by inactivation of the T4 polymerase by incubation at 348 K for 20 min. The vector was mixed with the insert in a 1:5 molar ratio. The annealing reaction was performed at 338 K for 10 min and the reaction mixture was subsequently slowly cooled to room temperature (295 K) in a PCR machine to improve the annealing efficiency. The annealing product was transformed into heat-competent bacterial cells and plated onto kanamycin plates. The resulting constructs were verified by colony PCR followed by DNA sequencing.

Recombinant polyhistidine-SUMO3-*SjPfn* (His-SUMO3-*SjPfn*) was expressed in *Escherichia coli* BL21(DE3)-RIPL cells. Transformed cells were grown at 298 K in 1 l autoinduction medium (Studier, 2005) supplemented with 50 µg ml⁻¹ kanamycin. After 36 h, the cells were harvested by centrifugation at 277 K and 3000g for 30 min. The cell pellets were subsequently resuspended in lysis buffer consisting of 100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7, 150 mM NaCl, 5 mM β-mercaptoethanol (BME) complemented with 10 mM imidazole, 10 µg ml⁻¹ DNase I, 6.25 µM MgCl₂ and 1× cOmplete Mini Protease Inhibitor Cocktail tablet

(Roche). The cells were disrupted by sonication and the lysate was clarified by centrifugation.

The fusion protein was bound to an Ni²⁺-charged Sepharose column (HisTrap FF, GE Healthcare) equilibrated with lysis buffer and washed extensively with lysis buffer without imidazole to remove nonspecifically bound proteins; the His-SUMO3-*SjPfn* was then eluted with 300 mM imidazole, 100 mM HEPES pH 7, 150 mM NaCl, 5 mM BME. The pooled fractions containing the fusion protein were concentrated to a volume of 2.5 ml and applied onto a PD-10 desalting column (GE Healthcare) to remove the imidazole in the buffer. The N-terminal 6×His-SUMO3 tag was cleaved using 1 µg ml⁻¹ recombinant SENP2 protease (Reverter & Lima, 2004; Marblestone *et al.*, 2006) for 2 h at room temperature (295 K). The cleaved proteins were passed through an Ni-NTA column to remove the SUMO3 tag and any uncleaved fusion protein, after which the *SjPfn* was concentrated. Final purification was performed by size-exclusion chromatography using a Superdex 75 10/300 GL column (GE Healthcare) equilibrated with 100 mM HEPES pH 7, 150 mM NaCl, 5 mM BME. Pure *SjPfn* was flash-frozen in liquid nitrogen and stored at 193 K. A CD spectrum of the purified cleaved *SjPfn* was measured from 260 to 185 nm at 293 K using an Applied Photophysics Chirascan Plus spectropolarimeter equipped with a TC125 thermal control unit (Quantum Northwest), a direct temperature probe and a 1 mm path-length quartz cuvette (Hellma). The spectrum was recorded at a concentration of 0.2 mg ml⁻¹ in 50 mM NaF, 20 mM sodium phosphate buffer pH 7, 1 mM tris(2-carboxyethyl)-phosphine (TCEP). The *DichroWeb* server (Lobley *et al.*, 2002) was used for secondary-structure determination using the *CDSSTR* algorithm (Compton & Johnson, 1986) and set3 optimized for 185–240 nm as a reference data set (Johnson, 1999).

2.2. Crystallization

Crystallization conditions for *SjPfn* were screened using the sitting-drop vapour-diffusion method in Swissci MRC 2 96-well plates. *SjPfn* was concentrated to 8 mg ml⁻¹ and crystallization experiments were set up manually by mixing 0.3 µl reservoir solution and 0.3 µl protein solution; the plates were incubated at 295 and 277 K. Commercially available screens [PEG/Ion, Crystal Screen Lite (Hampton Research, Aliso Viejo, USA), JSCG-plus, ProPLEX and Morpheus (Molecular Dimensions, Altamonte Springs, USA)] were used for initial screening. Data-collection quality crystals grew in 150 mM ammonium iodide, 25% (w/v) polyethylene glycol (PEG) 8000, 50 mM 2-(N-morpholino)ethanesulfonic acid (MES) pH 5.5 at room temperature (295 K).

2.3. Data collection and processing

Crystals of *SjPfn* were flash-cooled in liquid nitrogen after soaking them in a cryoprotectant solution consisting of 20% (v/v) glycerol in the reservoir solution. Preliminary X-ray diffraction tests and native data-set collection to 1.91 Å resolution were performed on EMBL beamline P14 at PETRA III/DESY, Hamburg, Germany. An additional single-wavelength anomalous diffraction (SAD) data set was collected on EMBL beamline P13 (PETRA III/DESY) using a wavelength of 2 Å in order to obtain anomalous signal from iodide and/or sulfur. The diffraction images were indexed, integrated and scaled using *XDS* (Kabsch, 2010) and *XDSi* (Kursula, 2004).

3. Results and discussion

Recombinant *SjPfn* (129 residues; molecular mass 14.4 kDa) was successfully overexpressed and purified using a two-step protocol

(Fig. 1*a*). The identity of the purified protein was verified by peptide mass fingerprinting using mass spectrometry at the Biocenter Oulu Proteomics Core Facility (data not shown). The folding and secondary-structure content of the purified SjPfn were analyzed using CD spectroscopy. The CD spectrum clearly indicates the protein to be folded (Fig. 1*b*). SjPfn was estimated to contain 18% α -helices and 31% β -sheets, which is roughly similar to homologous profilin crystal structures, which generally consist of 30–32% α -helices and 31–33% β -strands [PDB entries 1a0k (Thorn *et al.*, 1997), 4esp (Wang *et al.*, 2013) and 1cqa (Fedorov *et al.*, 1997)]. However, the helical content seems to be slightly lower than in its homologues in the crystalline form.

The initial crystallization hit was obtained from the PEG/Ion screen and consisted of 200 mM ammonium iodide, 20% (w/v) polyethylene glycol (PEG) 3350 pH 6.2. This condition was further optimized by screening different PEGs in the molecular-weight range 200–10 000 at different concentrations, combined with altering the ammonium iodide concentration and the pH. Single crystals with a longest dimension of approximately 100 μ m were obtained from the optimized condition (Table 2) in 5–7 d (Fig. 2*a*).

The diffraction power of the crystals varied greatly from no diffraction to diffraction beyond 2 Å resolution, and it was not

possible to predict the diffraction properties from the crystal morphology. After screening many crystals from slightly different

Table 2
Crystallization details for SjPfn.

Method	Sitting-drop vapour diffusion
Plate type	Swissci MRC 2 96-well plates
Temperature (K)	295
Protein concentration (mg ml ⁻¹)	8
Buffer composition of protein solution	100 mM HEPES pH 7, 150 mM NaCl, 5 mM BME
Composition of reservoir solution	150 mM ammonium iodide, 25% (w/v) PEG 8000, 50 mM MES pH 5.5
Volume and ratio of drop	0.3 μ l protein + 0.3 μ l reservoir
Volume of reservoir (μ l)	70

Figure 1
Purification and folding analysis of SjPfn. (a) Size-exclusion chromatogram and Coomassie-stained denaturing gel of the peak fractions containing SjPfn and eluting at 13–15 ml. The positions of molecular-mass markers are shown at the side of the gel and are labelled in kDa. (b) CD spectrum of SjPfn from 260 to 185 nm measured at 293 K.

Figure 2
Diffraction analysis of SjPfn. (a) A single crystal of SjPfn grown using 150 mM ammonium iodide, 25% (w/v) PEG 8000, 50 mM MES pH 5.5 at 295 K. (b) Self-rotation function at $\chi = 180^\circ$ calculated from the 2.2 Å resolution SjPfn data ($P2_12_12_1$) using *MOLREP*. No additional peaks indicating noncrystallographic symmetry were detected.

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	EMBL beamline P14 at PETRA III/DESY	EMBL beamline P13 at PETRA III/DESY
Wavelength (Å)	1.24	2.00
Temperature (K)	100	100
Detector	PILATUS 6M	PILATUS 6M
Crystal-to-detector distance (mm)	292	141
Rotation range per image (°)	0.1	0.5
No. of frames	1800	1500
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å, °)	$a = 31.82, b = 52.17,$ $c = 59.79,$ $\alpha = \beta = \gamma = 90$	$a = 35.29, b = 52.15,$ $c = 59.82,$ $\alpha = \beta = \gamma = 90$
Resolution range (Å)	50–1.91 (1.96–1.91)	39.3–2.20 (2.26–2.20)
Total No. of reflections	47684 (1451)	138834 (9560)
No. of unique reflections	14494 (729)	10071 (696)
Completeness† (%)	96.9 (66.2)	93.1 (86.4)
Multiplicity	3.3 (1.9)	13.7 (13.7)
$\langle I/\sigma(I) \rangle^\ddagger$	6.0 (1.3)	24.3 (10.8)
$R_{\text{meas}}^\S$ (%)	14.3 (75.5)	23.8 (8.0)
$CC_{1/2}^\P$ (%)	99.4 (72.2)	99.9 (99.5)
Overall B factor from Wilson plot (Å ²)	34.9	32.3
SigAno††	0.86 (0.77)	3.3 (1.7)

† The completeness is low in the highest resolution shell because these reflections were only recorded at the corners of the square-shaped detector. ‡ For the 1.91 Å resolution data set, the mean $I/\sigma(I)$ falls below 2.0 at 2.05 Å resolution. § R_{meas} is the redundancy-independent R factor as defined by Diederichs & Karplus (1997) and Weiss & Hilgenfeld (1997). $R_{\text{meas}} = \sum_{hkl} [N(hkl)/[N(hkl) - 1]]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$. ¶ $CC_{1/2}$ is defined as the correlation coefficient between two random half data sets, as described by Karplus & Diederichs (2012). †† SigAno is the mean anomalous difference in units of its estimated standard deviation as defined in XDS: SigAno = $[|F(+)| - |F(-)|]/\delta$, where $F(+)$ and $F(-)$ are structure-factor estimates obtained from the merged intensity observations in each parity class.

conditions, a native data set to 1.91 Å resolution could be recorded (Table 3). The crystal belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 31.82, b = 52.17, c = 59.79$ Å. This data set was initially used in phasing attempts using molecular replacement. Probably because of the small size and the rather featureless shape of the profilin monomer and its low sequence identity to the closest homologues (~20%), the molecular-replacement trials were not successful. Because SjPfn contains seven cysteines and five methionines and the crystallization condition contained iodide, the SAD method was also attempted using this data set. However, as the data-collection wavelength was not optimal, the phasing power was not sufficient to obtain a solution, although some heavy-atom sites could be found. Therefore, a second data set was collected to 2.2 Å resolution at a wavelength of 2 Å.

The crystal used for the latter data set also belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 35.29, b = 52.15, c = 59.82$ Å. Data processing showed a strong anomalous signal. The Matthews coefficients were calculated as 1.72 and 1.91 Å³ Da⁻¹ with solvent contents of 28.7 and 35.7% (Matthews, 1968) for the two data sets, respectively, suggesting that the asymmetric unit could accommodate only one SjPfn molecule. Data-collection and processing statistics are given in Table 3. The data quality was evaluated using *phenix.xtriage* (Adams *et al.*, 2010) and no indications of twinning or pseudo-translational symmetry were detected. A self-rotation function calculated using *MOLREP* (Vagin & Teplyakov, 2010) was also consistent with the presence of one molecule in the asymmetric unit (Fig. 2*b*). Structure determination using the SAD data is currently ongoing.

SjPfn has a fairly low sequence identity to canonical profilins, and from the sequence it is not clear whether it possesses the characteristic functional properties of profilins. Other parasites, such as the Apicomplexa, also possess highly divergent profilins which still bind

to actin, proline-rich sequence motifs and phosphoinositides (Kursula *et al.*, 2008; Kucera *et al.*, 2010). However, certain other profilins, such as mammalian profilin 4, do not bind to actin or proline-rich sequences despite their overall conserved fold (Behnen *et al.*, 2009). Structural and biochemical characterization of SjPfn will shed light on its properties and role in the parasite and will serve as a basis for evaluating the suitability of this protein for rational drug or vaccine design. The structure will also be an important addition to the pool of structural information on profilins from different organisms and will contribute to the understanding of the evolutionary relationships between the members of this ancient protein family.

We would like to thank the NKI Protein Facility, supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) (grant No. 175.010.2007.012), for providing the LIC vector. We acknowledge access to beam time and excellent technical support at the EMBL beamlines P13 and P14 at PETRA III/DESY, Hamburg. This work was financially supported by the Academy of Finland, the Emil Aaltonen Foundation, the Sigrid Jusélius Foundation, the German Ministry for Education and Research (BMBF) and the Hamburg Research and Science Foundation.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Behnen, M., Murk, K., Kursula, P., Cappallo-Obermann, H., Rothkegel, M., Kierszenbaum, A. L. & Kirchhoff, C. (2009). *BMC Cell Biol.* **10**, 34.
- Compton, L. A. & Johnson, W. C. (1986). *Anal. Biochem.* **155**, 155–167.
- Diederichs, K. & Karplus, P. A. (1997). *Nature (London)*, **4**, 269–275.
- Engels, D., Chitsulo, L., Montresor, A. & Savioli, L. (2002). *Acta Trop.* **82**, 139–146.
- Fedorov, A. A., Ball, T., Mahoney, N. M., Valenta, R. & Almo, S. C. (1997). *Structure*, **5**, 33–45.
- Gadermaier, G., Hauser, M. & Ferreira, F. (2013). *Methods*, doi:10.1016/j.jymeth.2013.06.014.
- Gryseels, B., Polman, K., Clerinx, J. & Kestens, L. (2006). *Lancet*, **368**, 1106–1118.
- Johnson, W. C. (1999). *Proteins*, **35**, 307–312.
- Jones, M. K., Gobert, G. N., Zhang, L., Sunderland, P. & McManus, D. P. (2004). *Bioessays*, **26**, 752–765.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
- Kucera, K., Koblansky, A. A., Saunders, L. P., Frederick, K. B., De La Cruz, E. M., Ghosh, S. & Modis, Y. (2010). *J. Mol. Biol.* **403**, 616–629.
- Kursula, I., Kursula, P., Ganter, M., Panjikar, S., Matuschewski, K. & Schüler, H. (2008). *Structure*, **16**, 1638–1648.
- Kursula, P. (2004). *J. Appl. Cryst.* **37**, 347–348.
- Li, M. Z. & Elledge, S. J. (2012). *Methods Mol. Biol.* **852**, 51–59.
- Lobley, A., Whitmore, L. & Wallace, B. A. (2002). *Bioinformatics*, **18**, 211–212.
- Marblestone, J. G., Edavettal, S. C., Lim, Y., Lim, P., Zuo, X. & Butt, T. R. (2006). *Protein Sci.* **15**, 182–189.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McManus, D. P. & Loukas, A. (2008). *Clin. Microbiol. Rev.* **21**, 225–242.
- Pearce, E. J. & MacDonald, A. S. (2002). *Nature Rev. Immunol.* **2**, 499–511.
- Plattner, F., Yarovsky, F., Romero, S., Didry, D., Carlier, M.-F., Sher, A. & Soldati-Favre, D. (2008). *Cell Host Microbe*, **3**, 77–87.
- Reverter, D. & Lima, C. D. (2004). *Structure*, **12**, 1519–1531.
- Santos, A. & Van Ree, R. (2011). *Int. Arch. Allergy Immunol.* **155**, 191–204.
- Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
- Thorn, K. S., Christensen, H. E., Shigeta, R., Huddler, D., Shalaby, L., Lindberg, U., Chua, N.-H. & Schutt, C. E. (1997). *Structure*, **5**, 19–32.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Wang, Y., Fu, T.-J., Howard, A., Kothary, M. H., McHugh, T. H. & Zhang, Y. (2013). *J. Agric. Food Chem.* **61**, 1573–1578.
- Weiss, M. S. & Hilgenfeld, R. (1997). *J. Appl. Cryst.* **30**, 203–205.
- Witke, W. (2004). *Trends Cell Biol.* **14**, 461–469.
- Yarmola, E. G. & Bubb, M. R. (2006). *Trends Biochem. Sci.* **31**, 197–205.
- Zhang, S.-M., Lv, Z.-Y., Zhou, H.-J., Zhang, L.-Y., Yang, L.-L., Yu, X., Zheng, H. & Wu, Z.-D. (2008). *Parasitol. Res.* **102**, 1367–1374.