crystallization communications

Expression, purification, crystallization and preliminary X-ray diffraction analysis of α-11 giardin from *Giardia lamblia*

Puja Pathuri,[a]‡ Emily Tam Nguyen[a]‡ and Hartmut Luecke[a,b,c]*

[a]Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697, USA, [b]Department of Physiology and Biophysics, University of California, Irvine, CA 92697, USA, and [c]Department of Information and Computer Sciences, University of California, Irvine, CA 92697, USA

‡ These authors contributed equally.

Correspondence e-mail: hudel@uci.edu

α-11 Giardin, a protein from the annexin superfamily, is a 35.0 kDa protein from the intestinal protozoan parasite *Giardia lamblia* which triggers a form of diarrhea called giardiasis. Here, the cloning, expression, purification and the crystallization of α-11 giardin under two different conditions and in two different space groups is reported. Crystals from the first condition diffracted to 1.1 Å and belong to a primitive orthorhombic space group, while crystals from the second condition, which included calcium in the crystallization solution, diffracted to 2.93 Å and belong to a primitive monoclinic space group. Determination of the detailed atomic structure of α-11 giardin will provide a better insight into its biological function and might establish whether this class of proteins is a potential drug target against giardiasis.

1. Introduction

Giardia lamblia is a protozoan parasite that causes intestinal illness in humans and other mammals (Palm *et al.*, 2003). It is a water-borne pathogen that occurs worldwide, infecting an estimated 250 million people. The life cycle of *Giardia* switches between cyst and trophozoite stages. Infection occurs when *Giardia* is ingested in its cyst form, which is dormant and water-resistant. Symptoms of the illness arise when cysts pass through the stomach and convert to trophozoites in a process called excystation (Svard *et al.*, 2003). During the trophozoite stage, the parasite is released into the small intestine, where it attaches and divides and completes its life cycle when the trophozoites encyst and are excreted in the feces as infective cysts. A unique characteristic of *Giardia* is that it triggers a wide range of symptoms, from asymptomatic to chronic infection, which are accompanied by acute diarrhea (giardiasis), intestinal malabsorption and weight loss (Roxstrom-Lindquist *et al.*, 2006).

Once inside the body, *Giardia* inhabits the upper small intestine, specifically the duodenum and jejunum, an extremely toxic environment with few living microbes. Because the cells of the intestine are continuously produced and replaced, *Giardia* has to constantly migrate and re-attach to the intestinal epithelium in order to prevent being swept away (Roxstrom-Lindquist *et al.*, 2006). Movement of *Giardia* within the intestine is aided by flagella, while the ventral disk, a cytoskeletal component, helps this parasite attach to the epithelial surface.

Since cell motility is an important aspect of survival, there is a strong connection between the cytoskeleton and *Giardia* virulence (Elmendorf *et al.*, 2003). Previous studies have identified the association of cytoskeletal proteins called α-giardins to the *Giardia* cytoskeleton. At least three of these proteins are components of the ventral disk (Ghosh *et al.*, 2001). α-Giardins, a multi-gene family consisting of 21 members, may be active participants in the excystation of *Giardia*. The members of the α-giardin family share a common ancestry with human annexins (Morgan & Fernandez, 1995). Although α-11 giardin shows high sequence homology to annexins, the sequence identity is low (Fig. 1). The sequence identity shared between α-11 giardin and the annexins (annexins A1, A2, A4, A5, A6
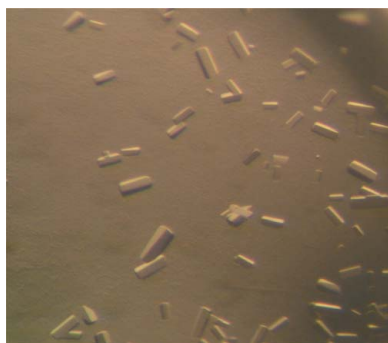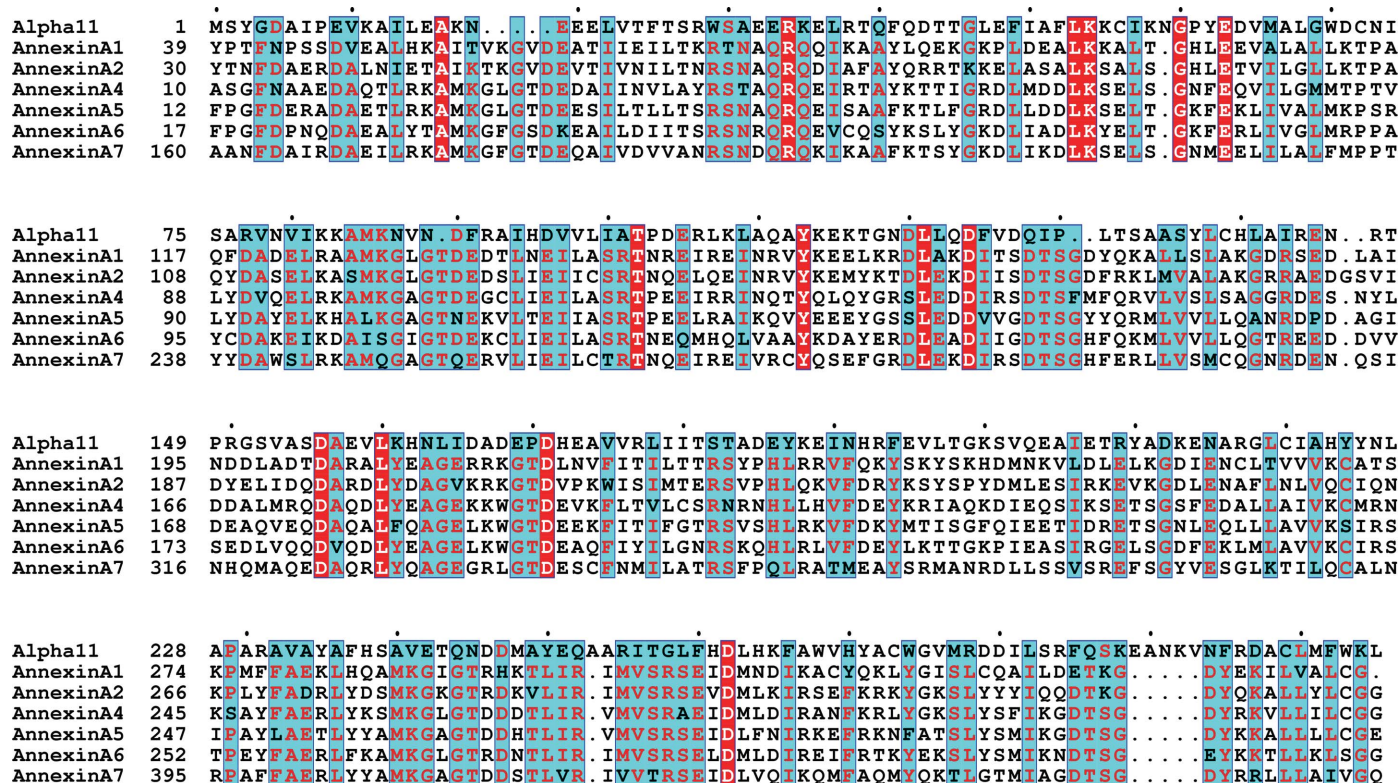
Figure 1
Amino-acid sequence alignment of α-11 giardin (alpha11) and annexins A1, A2, A4, A5, A6 and A7. Residues denoted as white letters on a red background are strictly conserved. Residues denoted in red on a cyan background are homologous, while residues denoted in black on a cyan background are non-homologous in this group. The alignment was performed using the program *ClustalW* (Thompson *et al.*, 1994) and the figure was prepared using *ESPript* (Gouet *et al.*, 1999).

and A7) ranges from approximately 10% to 26% (14 strictly conserved residues), with annexin A4 having the most identical residues (26% sequence identity) and annexin A1 having the least identical residues (9.5% sequence identity). Annexins are a family of cytosolic proteins that bind to phospholipids and cellular membranes in a Ca^{2+}-dependent manner and are implicated in numerous membrane-related processes, including membrane aggregation and fusion, vesicle formation and organization of cytoskeletal contacts (Gerke & Moss, 2002). Since annexins and α-giardins are evolutionarily related, it is possible that the α-giardin family may be involved in biological functions similar to those of annexins. Previous studies have shown that α-11 giardin exhibits high levels of mRNA transcription in trophozoites, encysting cells and cysts, when compared with other members of the α-giardin family. However, constitutive expression of this protein is found to be lethal to *Giardia* (Weiland *et al.*, 2005). Therefore, tight regulation of the transcription of this protein is necessary to ensure survival. Very little is known about α-11 giardin, even though it has been shown to be one of the four members (α-4, α-7.1 and α-8 giardins) of the α-giardin family that was lethal to the parasite in localization experiments in which the *Giardia* proteins, expressed under control of their own promoter, were transfected into trophozoites (Weiland *et al.*, 2005).

In this report, we describe the cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of α-11 giardin. X-ray diffraction analysis is the first step in determining the three-dimensional structure of this 35.0 kDa protein, a representative of a family of proteins for which no X-ray structures have been reported. A high-resolution X-ray structure is hoped to aid in the

discovery of drug leads that will target and disrupt the cytoskeleton of *G. lamblia*.

2. Materials and methods

2.1. Cloning and expression

The α-11 giardin gene (NCBI accession No. AY781330) was amplified from *G. lamblia* WB-C6 genomic DNA using the following primers: primer 1, 5′-CC**GGATCC**TACGGAGACGCTATCCCTG-3′, and primer 2, 5′-CC**GTCGAC**TCATTTAGCAAGCTTCCAGAAC-3′. The amplified gene containing 5′-*Bam*HI and 3′-*Sal*I restriction sites (indicated in bold in the primers shown above) was digested and ligated into the pGEX-6P3 expression vector (Amersham Pharmacia, Piscataway, NJ, USA), which was cut with the same restriction enzymes. The pGEX-6P3 expression vector introduces an N-terminal glutathione-*S*-transferase (GST) tag and a PreScission protease-recognition sequence located between the GST tag and the α-11 giardin recombinant protein. As a result, four additional residues (H_2N-Gly-Pro-Leu-Gly) derived from the PreScission protease-cleavage site and the *Bam*HI restriction site remained at the N-terminus of the 306-amino-acid α-11 giardin after removal of the GST tag.

The recombinant plasmid pGEX-6P3-α-11 was transformed into Rosetta (DE3) *Escherichia coli* cells (Novagen, San Diego, CA, USA). Cells grown from an overnight pre-culture incubated at 310 K were grown in 1 l Luria–Bertani (LB) broth medium containing 50 µg ml^{-1} ampicillin at 310 K and 170 rev min^{-1}. The culture was

induced with 1 m*M* isopropyl β-D-thiogalactopyranoside (IPTG) at an OD$_{600}$ of 0.5 for 3 h at 310 K and 100 rev min^{-1}. Cells were harvested by centrifugation at 6000*g* for 15 min at 277 K and were then frozen and stored at 253 K until lysis.

2.2. Purification

Cell pellets from 4 × 1 l of culture were suspended in lysis buffer [50 m*M* Tris–HCl pH 7.4, 1 m*M* EDTA, 100 m*M* NaCl, 1%(*v/v*) Nonidet P40 (NP40), 10%(*v/v*) glycerol, 1 m*M* dithiothreitol and 1 m*M* phenylmethylsulfonyl fluoride (PMSF)] containing a protease-inhibitor cocktail [20 µg ml^{-1} aprotinin, 2 µg ml^{-1} leupeptin, 2 µg ml^{-1} pepstatin, 0.5 m*M* benzamidine and 10 µ*M* *trans*-epoxy-succinyl-L-leucylamido(4-guanidino)butane (E-64)] and then lysed using a French press. Cell debris and denatured protein were removed by centrifugation at 40 000*g* for 1 h at 277 K. The super-natant was filtered through a 0.45 µm syringe filter (Fisher Scientific, Tustin, CA, USA) to remove any remaining fine cell debris before purification. The crude extract was loaded onto a pre-equilibrated [50 m*M* Tris–HCl pH 7.4, 1 m*M* EDTA, 100 m*M* NaCl, 1%(*v/v*) NP40, 10%(*v/v*) glycerol] 20 ml bed volume GSTrap Fast Flow column (Amersham Pharmacia, Piscataway, NJ, USA) at 277 K. The crude extract was incubated for 1.5 h on the column and then unbound protein was removed by washing with ten column volumes of 1 *M* NaCl. Additional impurities were removed by washing the
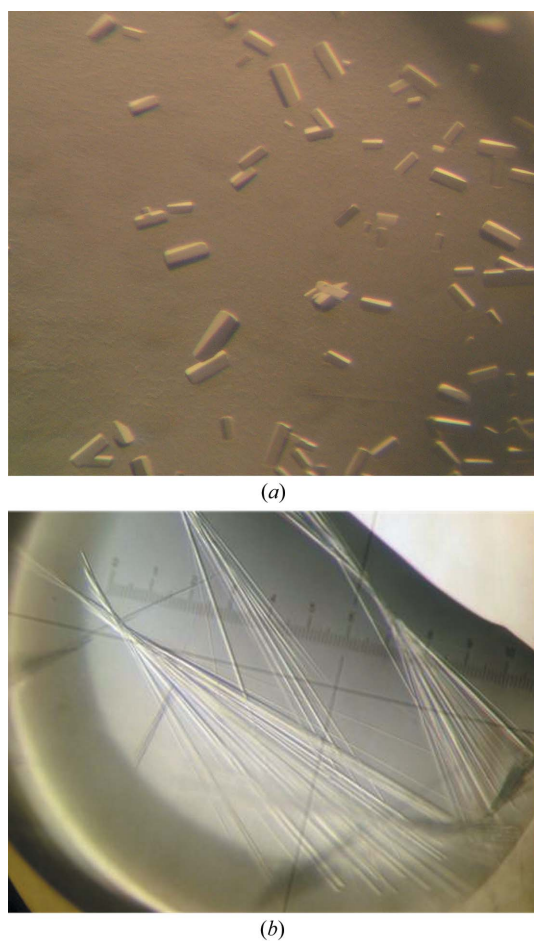

(*a*)

(*b*)

Figure 2
Crystals of α-11 giardin. Crystals grew in a primitive monoclinic space group in the presence of 50 m*M* Ca^{2+} to maximum dimensions of 20 × 20 × 50 µm in two weeks at 277 K (*a*) and in a primitive orthorhombic space group in the absence of calcium to maximum dimensions of 50 × 20 × 1000 µm in one week at 277 K (*b*).

column with ten column volumes of 50 m*M* Tris–HCl pH 8.0 and 150 m*M* NaCl. For on-column cleavage of the GST-fusion protein, the column was first equilibrated with protease-cleavage buffer (50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 1 m*M* EDTA and 1 m*M* dithio-threitol) and then incubated with 10 units of PreScission Protease (Amersham Pharmacia, Piscataway, NJ, USA) per milligram of fusion protein overnight at 277 K. The cleaved protein was eluted with protease-cleavage buffer and the flowthrough fractions containing α-11 giardin were collected. α-11 Giardin was separated from the PreScission Protease by applying the collected fractions onto a GSTrap Fast Flow column equilibrated with protease-cleavage buffer. In this step, the PreScission Protease remained bound to the column since it contains a GST tag, while the α-11 giardin eluted in the flowthrough. Fractions containing the purified α-11 giardin were pooled and concentrated to 11 mg ml^{-1} in 50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 1 m*M* EDTA and 1 m*M* dithiothreitol using a Centricon 10 kDa cutoff centrifugal filter (Millipore Corporation, Bedford, MA, USA) for crystallization. The purity of the sample during each purification step was confirmed using 12% SDS–PAGE stained with Coomassie Brilliant Blue.

2.3. Crystallization

Initial crystallization conditions were obtained by screening with commercially available crystallization kits from Nextal Biotechnol-ogies (Montreal, Quebec, Canada), Hampton Research (Laguna Niguel, CA, USA) and Emerald Biostructures (Bainbridge Island, WA, USA). Using the sitting-drop crystallization method, 2 µl protein solution was mixed with 2 µl precipitant solution over wells containing 200 µl precipitant solution in 96-well CompactClover Plates (Emerald Biostructures) at 277 K as well as 293 K. Preliminary crystallization conditions were obtained using two different precipi-tant solutions. Condition *A* yielded long thin needles using Nextal JCSG Suite solution H9: 200 m*M* lithium sulfate, 100 m*M* Bis-Tris pH 5.5 and 25%(*w/v*) PEG 3350 at 277 K. This condition was optimized to improve the quality of the crystals [200 m*M* lithium sulfate, 100 m*M* MES pH 6.0 and 25%(*w/v*) PEG 3350], which grew in one week to maximum dimensions of 50 × 20 × 1000 µm (Fig. 1). Condition *B* yielded small crystals in Index Screen solution No. 54 from Hampton Research: 50 m*M* calcium chloride, 100 m*M* Bis-Tris pH 6.5, 30%(*w/v*) PEG 550 MME at 277 K. After optimization, larger crystals grew in 50 m*M* calcium chloride, 100 m*M* Bis-Tris pH 6.5, 24%(*w/v*) PEG 550 MME in two weeks to maximum dimensions of 20 × 20 × 50 µm (Fig. 2).

2.4. Diffraction data collection

X-ray diffraction data were collected on beamline 9-1 at the Stanford Synchrotron Radiation Laboratory (SSRL) using remote robotic data collection at 0.979 Å. The detector was an Area Detector Systems Corporation (ADSC) Q315 detector. Single crystals from conditions *A* and *B* were soaked for a few seconds in their respective crystallization solutions containing 30% glycerol as a cryoprotectant before being flash-cooled in liquid nitrogen. For condition *A*, a high-resolution data set and a low-resolution data set were collected from one crystal to prevent overloading in the low-resolution shells. For the first data set, the crystal was exposed for 5 s per image and a total of 180 images with 1.0° oscillation angles were collected to 1.4 Å resolution with a crystal-to-detector distance of 180 mm. For the second data set, the crystal was exposed for 30 s and a total of 180 images with 1.0° oscillation angles were collected to 1.1 Å resolution using a crystal-to-detector distance of 110 mm. The two data sets were indexed and integrated separately and subsequently merged
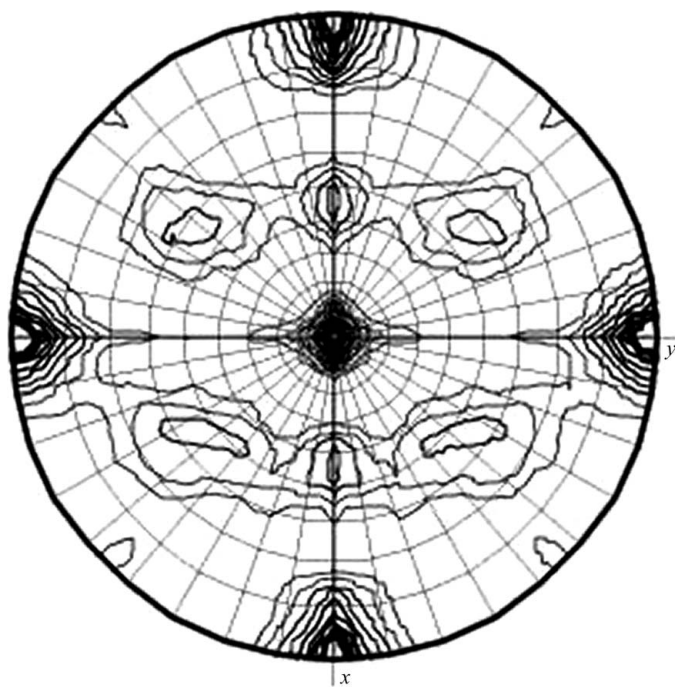
and scaled together using the program *d*TREK* (Pflugrath, 1999). For condition *B*, a total of 180 frames of 1° oscillation range were collected from a single crystal to 2.93 Å and were processed with the program *HKL*-2000 (Otwinowski & Minor, 1997). All data-collection statistics are summarized in Table 1.

3. Results and discussion

A crystal from condition *A* which diffracted to 1.1 Å was found to belong to a primitive orthorhombic space group, with unit-cell parameters *a* = 45.95, *b* = 60.95, *c* = 119.91 Å. With a unit-cell volume of 3.4×10^5 Å^3 and assuming the presence of one monomer in the asymmetric unit, the Matthews coefficient V_M (Matthews, 1968) was calculated to be 2.4 Å^3 Da^{-1}, corresponding to a solvent content of 48.7%. A crystal from condition *B* was found to belong to a primitive monoclinic space group, with unit-cell parameters *a* = 129.30, *b* = 45.29, *c* = 131.90 Å, *β* = 116.39°. In this case, the Matthews coefficient was calculated to be 2.47 Å^3 Da^{-1} with a solvent content of 50.2%, assuming there to be four molecules in the asymmetric unit. The presence of four molecules in the asymmetric unit was confirmed by calculating a self-rotation function using the program *MOLREP* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994; Vagin & Teplyakov, 1997). The self-rotation function shows the presence of two twofold peaks in addition to the crystallographic twofold peak parallel to the *b* axis. The two noncrystallographic peaks were parallel to the *a* and *c* axes, indicating that the primitive monoclinic crystal form has pseudo-orthorhombic packing (Fig. 3). The crystals that grew in conditions *A* and *B* crystallized in two different space groups with different unit-cell parameters, suggesting that α-11 giardin might be binding the calcium present only in the crystallization solution of condition *B*. This may be significant since it has been shown that nearly all members of the annexin family bind

Table 1
Data-collection statistics for crystals from conditions *A* and *B*.

Values in parentheses are for the highest resolution shell.

	Condition *A*	Condition *B*
Space group	Primitive orthorhombic	Primitive monoclinic
Unit-cell parameters		
a (Å)	45.95	129.30
b (Å)	60.95	45.29
c (Å)	119.91	131.90
β (°)		116.39
Matthews coefficient (Å^3 Da^{-1})	2.40	2.47
No. of molecules in ASU	1	4
Solvent content (%)	48.7	50.2
Resolution (Å)	42.74–1.10 (1.14–1.10)	50.0–2.93 (3.01–2.93)
Total observations	1203126	730522
Unique reflections	133460	30834
Completeness (%)	97.2 (94.2)	99.6 (99.4)
R_{merge}† (%)	12.1 (48.2)	10.4 (49.3)
Average $I/\sigma(I)$	6.4 (2.0)	13.7 (2.6)
Mosaicity (°)	0.36	0.72
Data-processing program used	*d*TREK*	*HKL*-2000

† $R_{merge} = \sum_{hkl} \sum_j |I_j(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_j \langle I(hkl) \rangle$, where $I_j(hkl)$ and $\langle I(hkl) \rangle$ are the intensity of measurement *j* and the mean intensity for the reflection with indices *hkl*, respectively.

calcium with affinities ranging from low micromolar to millimolar. It has also been shown through various biochemical assays that annexins bind to phospholipid bilayers (Gerke *et al.*, 2005) and monolayers (Freites *et al.*, 2004) in a Ca^{2+}-dependent manner. Structure determination of α-11 giardin, the first member of the giardin family with reported well diffracting crystals, will provide us with a better understanding of the possible function of this protein in *G. lamblia* and of how similar this protein from the α-giardin family is to other annexin structures. A crystal structure based on the very high resolution diffraction achieved for crystals from condition *A* would also result in the highest resolution annexin structure to date.

The amino-acid sequence of α-11 giardin was compared with amino-acid sequences deposited in the Protein Data Bank (PDB) in order to obtain a search model for molecular replacement. The closest homologue in the PDB is bovine annexin A4 (PDB code 1ann), which is only 26% identical to the amino-acid sequence of α-11 giardin. Molecular-replacement trials were performed with the program *Phaser* (Storoni *et al.*, 2004) implemented in the *CCP*4 suite and using annexin A4 (either in its entirety or broken into domains) as the search model. Unfortunately, *Phaser* could not find any solution even when using different combinations of the four domains from annexin A4. Since molecular replacement was unsuccessful, we decided the best approach to determine the structure of α-11 giardin would be multi-wavelength anomalous dispersion (MAD) with selenomethionyl-derivatized crystals. The amino-acid sequence of α-11 giardin shows that there are five non-terminal methionines in 306 amino acids, with a molecular weight of 35.0 kDa, making it a reasonable candidate for MAD phasing using selenomethionine incorporation. To this end, we have successfully expressed and purified selenomethionine-labeled α-11 giardin. Furthermore, we were able to crystallize the selenomethionine-labeled protein using condition *A* described above. Experimental phasing and structure determination are in progress.

Figure 3
Self-rotation function for $\kappa = 180°$ calculated using data from the primitive monoclinic crystal form, which contains four molecules of α-11 giardin in the asymmetric unit. The resolution limits are 25–2.93 Å and the Patterson vector cutoff radius is 30.0 Å.

crystallization communications

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Elmendorf, H. G., Dawson, S. C. & McCaffery, J. M. (2003). *Int. J. Parasitol.* **33**, 3–28.

Freites, J. A., Ali, S., Dennin, M. B., Rosengarth, A. & Luecke, H. (2004). *Langmuir*, **20**, 11674–11683.

Gerke, V., Creutz, C. E. & Moss, S. E. (2005). *Nature Rev. Mol. Cell Biol.* **6**, 449–461.

Gerke, V. & Moss, S. E. (2002). *Physiol. Rev.* **82**, 331–371.

Ghosh, S., Frisardi, M., Rogers, R. & Samuelson, J. (2001). *Infect. Immun.* **69**, 7866–7872.

Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). *Bioinformatics*, **15**, 305–308.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Morgan, R. O. & Fernandez, M. P. (1995). *Mol. Biol. Evol.* **12**, 967–979.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Palm, J. E., Weiland, M. E., Griffiths, W. J., Ljungstrom, I. & Svard, S. G. (2003). *J. Infect. Dis.* **187**, 1849–1859.

Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.

Roxstrom-Lindquist, K., Palm, D., Reiner, D., Ringqvist, E. & Svard, S. G. (2006). *Trends Parasitol.* **22**, 26–31.

Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.

Svard, S. G., Hagblom, P. & Palm, J. E. (2003). *FEMS Microbiol. Lett.* **218**, 3–7.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Weiland, M. E., McArthur, A. G., Morrison, H. G., Sogin, M. L. & Svard, S. G. (2005). *Int. J. Parasitol.* **35**, 617–626.