Preliminary crystallographic studies of a *Schistosoma mansoni* antigen (Sm21.7) dynein light-chain (DLC) domain

M. A. F. Costa, F. T. G. Rodrigues, B. C. A. Chagas, C. M. F. Rezende, A. M. Goes and R. A. P. Nagem*

Departamento de Bioquímica e Imunologia,
Instituto de Ciências Biológicas, Universidade
Federal de Minas Gerais, 31270-901 Belo
Horizonte-MG, Brazil

Correspondence e-mail: nagem@icb.ufmg.br

Received 20 February 2014
Accepted 24 April 2014

Schistosomiasis is an inflammatory chronic disease that represents a major health problem in tropical and subtropical countries. The drug of choice for treatment, praziquantel, is effective in killing adult worms but fails to kill immature forms and prevent reinfection. One prominent antigen candidate for an anti-schistosomiasis vaccine is the protein Sm21.7 (184 amino-acid residues) from *Schistosoma mansoni*, a tegumental protein capable of reducing the worm burden in a murine immunization model. In the present work, the Sm21.7 gene was cloned and expressed in *Escherichia coli* and the full-length protein was purified to homogeneity. Crystals of recombinant Sm21.7 suitable for X-ray diffraction were obtained using PEG monomethyl ether 2000 as a precipitant. X-ray diffraction images of a native crystal (at 2.05 Å resolution) and a quick-cryosoaked NaI derivative (at 1.95 Å resolution) were collected on the W01B-MX2 beamline at the Laboratório Nacional de Luz Síncrotron (LNLS, Brazilian Synchrotron Light Laboratory/MCT). Both crystals belonged to the hexagonal space group $P6_122$, with similar unit-cell parameters $a = b = 108.5$, $c = 55.8$ Å. SIRAS-derived phases were used to generate the first electron-density map, from which a partial three-dimensional model of Sm21.7 (from Gln89 to Asn184) was automatically constructed. Analysis of dissolved crystals by SDS-PAGE confirmed that the protein was cleaved in the crystallization drop and only the Sm21.7 C-terminal domain was crystallized. The structure of the Sm21.7 C-terminal domain will help in the localization of the epitopes responsible for its protective immune responses, constituting important progress in the development of an anti-schistosomiasis vaccine.

1. Introduction

Schistosomiasis is an inflammatory chronic disease caused by worms of the *Schistosoma* genus that represents a major health problem in tropical and subtropical countries (Gryseels *et al.*, 2006). The main schistosomes infecting humans are *S. haematobium* in Africa, *S. mansoni* in Africa and South America, *S. intercalatum* in West Africa and *S. japonicum* and *S. mekongi* in Asia. The former causes fibrosis, structuring and calcification of the urinary tract, while the others cause chronic liver and intestinal fibrosis (Ross *et al.*, 2013; Thétiot-Laurent *et al.*, 2013).

Estimates indicate that over 200 million individuals worldwide are currently infected (King, 2009), with more than 250 000 deaths per year (van der Werf *et al.*, 2003). Moreover, the impact of the severe and debilitating effects of schistosomiasis account for the loss of 4.5 million disability-adjusted life years (DALY; Steinmann *et al.*, 2006).

Current schistosomiasis control strategies are predominantly based on the treatment of infected individuals with safe and effective drugs (Gryseels *et al.*, 2006). The drug of choice, praziquantel, is effective in killing adult worms but fails to kill immature forms and to prevent reinfection (Gray *et al.*, 2010). Besides, there is evidence supporting the development of drug-resistant parasites (Fallon & Doenhoff, 1994; Ismail *et al.*, 1996). These problems limit the success of control strategies based on chemotherapy only. Therefore, a protective anti-

schistosomiasis vaccine would be valuable to the current disease-control program.

Recently, a significant effort has been made to develop a vaccine against schistosomiasis, and several vaccine candidates have been identified. One prominent vaccine candidate is the protein Sm21.7 (184 residues; accession No. XP_002569898) from *S. mansoni*. It was identified by screening a sporocyst cDNA expression library using irradiated cercaria-vaccinated rabbit serum. Sequence analysis demonstrated that the antigen is composed of an N-terminal EF-hand-like motif (Arg17–Gly67) and an assigned C-terminal dynein light-chain (DLC) domain (Lys81–Thr181). Surprisingly, ^{45}Ca -binding experiments with recombinant (GST-fusion protein) and parasite-derived antigen showed no evidence of calcium binding, suggesting the presence of a nonfunctional EF-hand motif (Francis & Bickle, 1992).

Sm21.7 was tested as an antigen in recombinant vaccines (Ahmed *et al.*, 2001) and DNA vaccines (Ahmed *et al.*, 2006), conferring a significant level of protection in mice against *S. mansoni* challenge infection (Ahmed *et al.*, 2001, 2006). Northern blot analysis demonstrated that mRNA for Sm21.7 is present in the sporocyst, schistosomula and adult stages (Francis & Bickle, 1992), and RT-PCR showed that transcription occurred in eggs, cercaria and adult worms (Fitzsimmons *et al.*, 2007). The antigen has been observed in schistosomula (Curwen *et al.*, 2004), eggs, cercariae (Fitzsimmons *et al.*, 2007) and adult antigen preparation (Curwen *et al.*, 2004; Fitzsimmons *et al.*, 2007).

Recently, our group has demonstrated that an adult worm protein preparation (fraction PIII) containing a tegument protein named P22 was able to engender protective and immunomodulatory effects against murine schistosomiasis. A cDNA clone encoding P22 was isolated from an *S. mansoni* adult worm cDNA library using anti-PIII rabbit serum and showed 100% identity to Sm21.7 (Rezende *et al.*, 2011).

Mice immunized with recombinant Sm21.7 (rP22; C-terminally 6×His-tagged fusion protein) exhibited a 51 and 22% decrease in adult worm burden and in hepatic eggs, respectively. Additionally, a recombinant Sm21.7 vaccine produced a reduction of 60% in liver granuloma size and a 71% reduction in fibrosis in mice, suggesting that Sm21.7 could contribute to downmodulate granulomatous hypersensitivity to *S. mansoni* eggs (Rezende *et al.*, 2011).

Even though relevant biological information about Sm21.7 has been obtained in recent years, there is a lack of structural data for both of its domains. The primary structure of the Sm21.7 amino-terminal domain (EF-hand-like motif) shares 52% similarity with calmodulin, but does not bind calcium as reported previously (Francis & Bickle, 1992). The carboxy-terminal domain (dynein light-chain domain) shares only 24% identity (*E*-value of 0.29) with a fragment of a *Kluyveromyces lactis* Hsv2p protein for which the three-dimensional structure was recently determined (Baskaran *et al.*, 2012; Krick *et al.*, 2012). Apart from this, tertiary structures for DLC domains are available for proteins sharing only 19% identity to Sm21.7 (Romes *et al.*, 2012; Gallego *et al.*, 2013; Rao *et al.*, 2013).

In order to better characterize Sm21.7, compare its structure with the most similar proteins available and localize the epitopes responsible for its protective immune responses, we aimed to clone, express and purify the full-length antigen. Crystallization conditions were also determined, but despite all efforts, X-ray diffraction results showed that only the C-terminal domain of this protein was crystallized. Owing to the importance of schistosomiasis and the promising results of a vaccine composed of recombinant Sm21.7, we aimed to determine the initial phases for the crystalline structure of the Sm21.7 carboxy-terminal domain by the heavy-atom method.

2. Experimental methods

2.1. Cloning, expression and purification

The gene sequence encoding full-length Sm21.7 (Gene ID 8347700) was amplified by PCR from a previously constructed plasmid (pETDEST42-rP22; Rezende *et al.*, 2011). The primers Sm21.7Fw (5'-CAT **ATG** GAT AGT CCA ATG GAA AAA TTT ATT C-3') and Sm21.7Rv (5'-C **TCG AGT** CTA GTT ACT TGG TGT ACG CCA AG-3'), containing *Nde*I and *Xho*I restriction sites (bold sequences), respectively, were used to amplify the Sm21.7 coding sequence. The PCR-amplified fragment was inserted into pGEM-T vector (Promega) and the recombinant plasmids pGEM-Sm21.7 were transformed into *Escherichia coli* DH5α cells. pGEM-Sm21.7 plasmids were then double-digested with *Nde*I and *Xho*I and the released inserts were ligated into pET-28a(TEV) vector (Carneiro *et al.*, 2006) previously digested with the same enzymes. The recombinant plasmids pET-28a(TEV)-Sm21.7 were transformed into *E. coli* BL21(DE3) and selected on kanamycin plates.

Large-scale expression cultures of the recombinant 6×His-Sm21.7 (rSm21.7) were carried out by inoculating 20 ml of an overnight rSm21.7 culture into 1 l Luria–Bertani broth medium containing 50 µg ml⁻¹ kanamycin. The culture was incubated at 310.15 K and 200 rev min⁻¹ until a cell density of between 0.4 and 0.8 (OD_{600 nm}) was achieved. Expression of the recombinant protein was then induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM followed by incubation of the culture for 4 h at 310.15 K and 200 rev min⁻¹.

Cells were harvested by centrifugation at 7000g for 15 min at 277.15 K and resuspended in lysis buffer [50 mM Tris–HCl pH 7.5, 1% (w/v) sucrose, 1% (v/v) glycerol, 1% (v/v) Tween 20, 1 mM phenylmethylsulfonyl fluoride (PMSF)]. The lysis process was performed by three passes through an EmulsiFlex-C3 homogenizer (Avestin). The lysate was clarified by centrifugation at 10 000g for 1 h at 277.15 K. In order to purify the rSm21.7, imidazole was added to the supernatant to a final concentration of 30 mM and the mixture was loaded onto a 5 ml HisTrap HP Ni²⁺-affinity column (GE Healthcare) connected to an ÄKTAprime plus system (GE Healthcare). The affinity column had previously been equilibrated with a binding buffer solution (20 mM sodium phosphate pH 7.5, 500 mM NaCl, 30 mM imidazole, 1 mM PMSF). After washing the column with four volumes of the binding buffer, the rSm21.7 antigen was eluted using a gradient of imidazole. For this purpose, an elution buffer solution (20 mM sodium phosphate pH 7.5, 500 mM NaCl, 500 mM imidazole, 1 mM PMSF) was automatically mixed with the binding buffer solution within the system during the chromatographic run. The fractions containing rSm21.7 were analyzed by SDS–PAGE, pooled, concentrated to 5 ml using a sample concentrator (Vivaspin 20, MWCO 10 kDa; GE Healthcare) and loaded onto a size-exclusion column (HiLoad 16/60 Superdex 200 prep grade; GE Healthcare). The column was previously equilibrated with a buffer solution consisting of 50 mM sodium phosphate pH 7.5, 50 mM NaCl, 1 mM PMSF. The same ÄKTAprime plus system (GE Healthcare) was used. Eluted peak fractions were pooled and the homogeneity of the purified samples was verified by SDS–PAGE and mass spectrometry.

Following size-exclusion chromatography, crystallization tests were carried out with freshly concentrated protein at approximately 20 mg ml⁻¹ (Vivaspin 20, MWCO 10 kDa; GE Healthcare). The concentration of purified rSm21.7 was estimated by measuring the absorbance at 280 nm and employing its theoretical molar extinction coefficient as calculated using the *ProtParam* software (Gasteiger *et al.*, 2005) available on the ExPASy Proteomics Server (<http://www.expasy.org>; Artimo *et al.*, 2012).

2.2. Crystallization, data collection and processing

All crystallization trials were accomplished at 291.15 K using the hanging-drop vapour-diffusion method in 24-well plates.

Initial screening was performed using in-house-prepared solutions based on the Crystal Screen and Crystal Screen 2 formulations (Hampton Research). Drops consisting of 1 μ l protein solution (20 mg ml⁻¹ rSm21.7 in 50 mM sodium phosphate buffer pH 7.5, 50 mM NaCl, 1 mM PMSF) and an equal volume of crystallization solution were equilibrated against 1 ml reservoir solution (the same crystallization solution) in a manual setup.

Of the 98 conditions tested two hits were identified: (i) 200 mM ammonium sulfate, 100 mM sodium acetate trihydrate pH 4.6, 30% (w/v) PEG monomethyl ether 2000 and (ii) 200 mM ammonium sulfate, 100 mM sodium cacodylate trihydrate pH 6.5, 30% (w/v) PEG 8000. These two promising conditions were optimized by varying the pH and the precipitant and salt concentrations. Moreover, in each new condition, drops of different volumes and protein:precipitant solution ratios (2 μ l:2 μ l, 1 μ l:2 μ l and 2 μ l:1 μ l) were tested, totalling 288 drops.

X-ray diffraction data sets were collected on the W01B-MX2 beamline (Guimarães *et al.*, 2009) at the Laboratório Nacional de Luz Síncrotron (LNLS, Brazilian Synchrotron Light Laboratory/MCT; Campinas-SP, Brazil). Before data acquisition of a native sample, a single rSm21.7 crystal was transferred for a few seconds to a 10 μ l drop of a cryogenic solution formed by the addition of 1 μ l ethylene glycol to 9 μ l reservoir solution. The crystal was then flash-cooled to 100 K in a nitrogen stream and used for data acquisition. An iodide derivative was obtained using the quick-cryosoaking approach (Dauter *et al.*, 2000; Nagem *et al.*, 2001, 2003). Essentially, a native crystal was soaked for 30 s in the same cryogenic solution additionally containing sodium iodide (at a final concentration of 0.4 M) and then flash-cooled in the nitrogen stream. In both cases, X-ray diffraction images were recorded on a Rayonix MarMosaic 225 CCD detector and were processed with *HKL-2000* (Otwinowski & Minor, 1997).

The phase problem was solved by using the SIRAS (single isomorphous replacement with anomalous scattering) method within *PHENIX* (Adams *et al.*, 2010). Iodide sites were obtained using the anomalous scattering contribution present in the derivative data set, but both isomorphous and anomalous contributions were used for phase determination. SIRAS-derived phases were used to generate the first electron-density map, enabling the construction of a preliminary model of rSm21.7. Restrained refinement and visual inspection of the model were performed with *PHENIX* and *Coot* (Emsley *et al.*, 2010), respectively.

3. Results and discussion

The 555 bp gene encoding rSm21.7 was successfully amplified and cloned into the expression vector pET-28a(TEV). This vector adds an N-terminal 6 $\times$ His tag and a TEV protease cleavage site to the protein (MGHHHHHHENLYFQGH—target protein; Carneiro *et al.*, 2006).

After expression in *E. coli* BL21(DE3) and cell lysis (Fig. 1*a*), rSm21.7 was purified by Ni²⁺-affinity and size-exclusion chromatography (Fig. 1*b*). In the last purification step, the protein eluted with an apparent molecular weight of 33 kDa, making the correct assignment of its oligomeric state in the conditions tested impossible. The purified protein showed a single band of approximately 25 kDa on SDS-PAGE (Fig. 1*a*), which is in good agreement with the calculated (23 688.8 Da) and mass-spectrometry-derived (23 564.4 Da; data not shown) molecular weight of rSm21.7. The final

yield of purified protein was approximately 30 mg per litre of cell culture, allowing the preparation of approximately 1.5 ml of a 20.0 mg ml⁻¹ protein sample for crystallization assays per expression batch.

Owing to the fact that purified rSm21.7 is susceptible to degradation with time, a protease inhibitor (PMSF) was added to the lysis and purification buffers. Recombinant Sm21.7 samples also show signs of precipitation when stored at 277.15 K or thawed. Consequently, all experiments were performed with freshly purified protein.

Following the initial crystallization screening, hits were optimized by testing 288 new crystallization conditions with varying pH values (in the range 4.0–5.5 using sodium acetate buffer or 5.5–7.4 using sodium cacodylate buffer) and precipitant and salt concentrations [PEG 8000 or PEG monomethyl ether 2000 from 20 to 40% (w/v) and

Figure 1 Quality assessment of rSm21.7 expression and purification. (*a*) SDS-PAGE of rSm21.7 samples during protein expression and purification. Lane 1, molecular-weight marker (labelled in kDa); lane 2, non-induced cells; lane 3, 4 h IPTG-induced cells; lane 4, soluble fraction after cell lysis; lane 5, insoluble fraction after cell lysis; lane 6, sample obtained after Ni²⁺-affinity chromatography; lane 7, sample obtained after size-exclusion chromatography; lane 8, sample obtained from crystals used for structure determination. (*b*) Size-exclusion chromatography profile of rSm21.7. The protein eluted at a volume close to 94 ml, with an apparent molecular weight of 33 kDa, consistent with either a monomer or a dimer in solution. The column was calibrated with Sigma Gel Filtration Molecular Weight Markers (catalogue No. MWGF200) comprising the following standard molecules: cytochrome *c* from horse heart (12.4 kDa), carbonic anhydrase from bovine erythrocytes (29 kDa), albumin from bovine serum (66 kDa), alcohol dehydrogenase from yeast (150 kDa) and β -amylase from sweet potato (200 kDa).

ammonium sulfate from 100 to 300 mM, respectively]. Different crystallization drop volumes and protein:precipitant solution ratios (2 μ l:2 μ l, 1 μ l:2 μ l and 2 μ l:1 μ l) were also tested.

The best diffracting plate-like crystals with area dimensions of 0.5 $\times$ 0.1 mm were obtained after 30 d using the following optimized reservoir conditions: (i) 100 mM ammonium sulfate, 25%(w/v) PEG monomethyl ether 2000, 100 mM sodium acetate buffer pH 5.5 (1 μ l protein solution and 2 μ l reservoir solution) and (ii) 200 mM ammonium sulfate, 25%(w/v) PEG 8000, 100 mM sodium cacodylate buffer pH 5.5 (2 μ l protein solution and 2 μ l reservoir solution). Both conditions resulted in the formation of a protein precipitate from which crystals displaying the same macroscopic morphology grew (Fig. 2).

A single native rSm21.7 crystal grown in PEG monomethyl ether 2000 was submitted to an X-ray diffraction experiment and diffracted to 2.05 Å resolution (Fig. 3a). The crystal belonged to the hexagonal space group $P6_122$ as determined during phasing and model building (Table 1). The calculated Matthews coefficient (V_M ; Matthews, 1968) of 2.00 Å³ Da⁻¹ with a solvent content of 38.4% suggests the

presence of one full-length rSm21.7 molecule in the asymmetric unit of the crystal. However, all attempts to solve the rSm21.7 crystal structure by the molecular-replacement method using the EF-hand-like motif of *Paramecium tetraurelia* calmodulin (PDB entry 1clm; Rao *et al.*, 1993) and the dynein light-chain domain of *Kluyveromyces lactis* Hsv2p protein (PDB entries 4exv and 4av9; Baskaran *et al.*, 2012; Krick *et al.*, 2012) as search models were unsuccessful. As a consequence, the heavy-atom method was used for phasing and structure determination.

An iodide derivative was prepared by soaking a native crystal (grown in PEG 8000) in a cryogenic solution containing 0.4 M sodium iodide for 30 s. This crystal diffracted to 1.95 Å resolution (Fig. 3b) and was shown to be isomorphous to the native crystal (Table 1). An almost 6% anomalous difference observed in Friedel mates suggests systematic incorporation of iodide anions into the crystallographic structure.

Indeed, two iodide anion sites with high occupancies (1.00 and 0.95) were found in the crystallographic structure of the derivative using *SHELXD* (Sheldrick, 2010). These sites, together with both

(a)

(b)

Figure 2

Crystals of rSm21.7 from *S. mansoni*. (a) Crystals grown in 100 mM ammonium sulfate, 25%(w/v) PEG monomethyl ether 2000, 100 mM sodium acetate buffer pH 5.5. (b) Crystals obtained in 200 mM ammonium sulfate, 25%(w/v) PEG 8000, 100 mM sodium cacodylate buffer pH 5.5.

(a)

(b)

Figure 3

X-ray diffraction images of rSm21.7. Image obtained from (a) a native crystal and (b) an iodide-derivative crystal using a Rayonix MarMosaic 225 CCD detector. The brightness and/or contrast of the displayed images were adjusted by resolution shells to facilitate spot visualization.

Table 1

Data-collection and processing statistics for native and iodide-derivative rSm21.7 crystals.

Values in parentheses are for the lowest/highest resolution shells.

	Native	Iodide derivative
Wavelength (Å)	1.459	1.459
Temperature (K)	100	100
Crystal-to-detector distance (mm)	120	105
Rotation range per image (°)	0.5	0.5
Exposure time per image (s)	16	6
Total rotation range (°)	130	150
Space group	<i>P</i> 6 ₁ 22	<i>P</i> 6 ₁ 22
Unit-cell parameters (Å)	<i>a</i> = <i>b</i> = 108.5, <i>c</i> = 55.8	<i>a</i> = <i>b</i> = 109.1, <i>c</i> = 56.4
Resolution range (Å)	50.00–2.05 (50.00–4.69/ 2.11–2.05)	50.00–1.95 (50.00–4.46/ 2.01–1.95)
No. of observations	169511	183507
No. of unique reflections	22353 (1899/1778)	27084 (2284/2092)
Data completeness (%)	96.7 (97.8/92.9)	98.9 (99.0/92.0)
$\langle I/\sigma(I) \rangle$	26.9 (47.5/2.7)	16.9 (52.4/2.1)
Multiplicity	7.6 (8.3/4.9)	6.8 (9.4/3.1)
$R_{\text{merge}}^{\dagger}$ (%)	6.8 (3.4/50.9)	10.7 (4.3/52.1)
Average mosaicity (°)	0.81	1.18
Matthews coefficient $\ddagger$ (Å ³ Da ^{−1})	4.20	4.30
Solvent content $\ddagger$ (%)	70.7	71.4
$R_{\text{anom}}^{\S}$ (%)	—	5.98 $\P$
$R_{\text{isom}}^{\dagger\dagger}$ (%)	—	21.6 $\P$
FOM $\ddagger\ddagger$	—	0.43 (0.69)

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl) \times 100$, where $I_i(hkl)$ is the *i*th intensity measurement of reflection *hkl* and $\langle I(hkl) \rangle$ is its average. $\ddagger$ Numbers refer to suggested values calculated by assuming the presence of 96 carboxy-terminal residues of rSm21.7 per asymmetric unit (Gln89–Asn184). $\S R_{\text{anom}} = \sum_{hkl} |I(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} I(hkl) \times 100$. $\P$ The signal was computed with reflections up to 3.0 Å resolution. $\dagger\dagger R_{\text{isom}} = \sum_{hkl} |F_{\text{deriv}}(hkl) - F_{\text{nat}}(hkl)| / \sum_{hkl} |F_{\text{nat}}(hkl)| \times 100$. $\ddagger\ddagger$ The values refer to FOM prior to and after (in parentheses) density modification computed for all reflections.

data sets (native and derivative), were used for SIRAS phasing within PHENIX. Density modification followed by automatic model building against the native data set resulted in a partial rSm21.7 model with an *R* value and *R*_{free} of 0.23 and 0.25, respectively. Surprisingly, this partial model contained 96 residues from the carboxy-terminal dynein light-chain domain of rSm21.7 (from Gln89 to Asn184), with a theoretical molecular weight of 11 277 Da. As confirmed by SDS–PAGE prepared with dissolved crystals (Fig. 1a), only the C-terminal domain of rSm21.7 was crystallized. The observed Matthews coefficient (*V*_M) of this crystal is 4.20 Å³ Da^{−1}, with a solvent content of 70.7% (Table 1). The cleavage of Sm21.7 is also inferred from two-dimensional gel electrophoresis/mass-spectrometry results (Curwen *et al.*, 2004), suggesting that a similar *in vivo* post-translation modification might occur. Visual model inspection/correction and automated maximum-likelihood refinement are currently under way. The detailed structure of the C-terminal dynein light-chain domain of Sm21.7 will be published elsewhere.

This work was supported by the Fundação de Desenvolvimento à Pesquisa do Estado de Minas Gerais (FAPEMIG; Project Rede-170/08) and the Laboratório Nacional de Luz Síncrotron (LNLS, Brazilian Synchrotron Light Laboratory/MCT; Project-9248/ROBOLAB-14983/MX1-14663). MAFC, FTGR and CMFR are recipients of scholarships from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). BCAC is the recipient of a scholarship from

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Ahmed, H. M., Romeih, M. H., Sherif, S. A., Fahim, F. A. & Saber, M. A. (2001). *Arab J. Biotechnol.* **4**, 229–249.
- Ahmed, H. M., Romeih, M. H., Shousha, T. S. A., Dabaa, E. E. & Saber, M. A. (2006). *J. Am. Sci.* **2**, 59–69.
- Artimo, P. *et al.* (2012). *Nucleic Acids Res.* **40**, W597–W603.
- Baskaran, S., Ragusa, M. J., Boura, E. & Hurley, J. H. (2012). *Mol. Cell*, **47**, 339–348.
- Carneiro, F. R. G., Silva, T. C. L., Alves, A. C., Haline-Vaz, T., Gozzo, F. C. & Zanchin, N. I. T. (2006). *Biochem. Biophys. Res. Commun.* **343**, 260–268.
- Curwen, R. S., Ashton, P. D., Johnston, D. A. & Wilson, R. A. (2004). *Mol. Biochem. Parasitol.* **138**, 57–66.
- Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* **D56**, 232–237.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Fallon, P. G. & Doenhoff, M. J. (1994). *Am. J. Trop. Med. Hyg.* **51**, 83–88.
- Fitzsimmons, C. M., McBeath, R., Joseph, S., Jones, F. M., Walter, K., Hoffmann, K. F., Kariuki, H. C., Mwatha, J. K., Kimani, G., Kabatereine, N. B., Vennervald, B. J., Ouma, J. H. & Dunne, D. W. (2007). *Int. Arch. Allergy Immunol.* **142**, 40–50.
- Francis, P. & Bickle, Q. (1992). *Mol. Biochem. Parasitol.* **50**, 215–224.
- Gallego, P., Velazquez-Campoy, A., Regué, L., Roig, J. & Reverter, D. (2013). *J. Biol. Chem.* **288**, 12283–12294.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. Walker, pp. 571–607. Totowa: Humana Press.
- Gray, D. J., McManus, D. P., Li, Y., Williams, G. M., Bergquist, R. & Ross, A. G. (2010). *Lancet Infect. Dis.* **10**, 733–736.
- Gryseels, B., Polman, K., Clerinx, J. & Kestens, L. (2006). *Lancet*, **368**, 1106–1118.
- Guimarães, B. G., Sanfelici, L., Neuenschwander, R. T., Rodrigues, F., Grizolli, W. C., Raulik, M. A., Piton, J. R., Meyer, B. C., Nascimento, A. S. & Polikarpov, I. (2009). *J. Synchrotron Rad.* **16**, 69–75.
- Ismail, M., Metwally, A., Farghaly, A., Bruce, J., Tao, L. F. & Bennett, J. L. (1996). *Am. J. Trop. Med. Hyg.* **55**, 214–218.
- King, C. H. (2009). *N. Engl. J. Med.* **360**, 106–109.
- Krick, R., Busse, R. A., Scacioc, A., Stephan, M., Janshoff, A., Thumm, M. & Kühnel, K. (2012). *Proc. Natl Acad. Sci. USA*, **109**, E2042–E2049.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Nagem, R. A. P., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst.* **D57**, 996–1002.
- Nagem, R. A. P., Polikarpov, I. & Dauter, Z. (2003). *Methods Enzymol.* **374**, 120–137.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Rao, L., Romes, E. M., Nicholas, M. P., Brenner, S., Tripathy, A., Gennerich, A. & Slep, K. C. (2013). *Mol. Biol. Cell*, **24**, 2362–2377.
- Rao, S. T., Wu, S., Satyshur, K. A., Ling, K.-Y., Kung, C. & Sundaralingam, M. (1993). *Protein Sci.* **2**, 436–447.
- Rezende, C. M. F., Silva, M. R., Santos, I. G. D., Silva, G. A. B., Gomes, D. A. & Goes, A. M. (2011). *Immunol. Lett.* **141**, 123–133.
- Romes, E. M., Tripathy, A. & Slep, K. C. (2012). *J. Biol. Chem.* **287**, 15862–15873.
- Ross, A. G. P., Olveda, R. M., Acosta, L., Harn, D. A., Chy, D., Li, Y., Gray, D. J., Gordon, C. A., McManus, D. P. & Williams, G. M. (2013). *Microbes Infect.* **15**, 858–865.
- Sheldrick, G. M. (2010). *Acta Cryst.* **D66**, 479–485.
- Steinmann, P., Keiser, J., Bos, R., Tanner, M. & Utzinger, J. (2006). *Lancet Infect. Dis.* **6**, 411–425.
- Thétiot-Laurent, S. A. L., Boissier, J., Robert, A. & Meunier, B. (2013). *Angew. Chem. Int. Ed.* **52**, 7936–7956.
- Werf, M. J. van der, de Vlas, S. J., Brooker, S., Looman, C. W., Nagelkerke, N. J. D., Habbema, J. D. F. & Engels, D. (2003). *Acta Trop.* **86**, 125–139.