Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Expression, crystallization and preliminary structural analysis of the ectoplasmic region of apical membrane antigen 1 from *Plasmodium vivax*, a malaria-vaccine candidate

Brigitte Vulliez-Le Normand,[a] Juan Carlos Pizarro,[a] Marie-Laure Chesne-Seck,[a] Clemens H. M. Kocken,[b] Bart Faber,[b] Alan W. Thomas[b] and Graham A. Bentley[a]*

[a]Unité d'Immunologie Structurale, CNRS URA 2185, Institut Pasteur, 25 Rue du Dr Roux, 75724 Paris, France, and [b]Department of Parasitology, Biomedical Primate Research Centre, 2280 GH Rijswijk, The Netherlands

Correspondence e-mail: bentley@pasteur.fr

Received 28 June 2004
Accepted 27 August 2004

Apical membrane antigen 1 (AMA1), a type 1 transmembrane protein present in the microneme organelles of *Plasmodium*, is a leading malaria-vaccine candidate. The ectoplasmic region of AMA1 from *P. vivax* has been expressed in *Pichia pastoris* and crystallized in two different forms: an orthorhombic form (space group $P2_12_12_1$, unit-cell parameters $a = 54.1$, $b = 76.1$, $c = 103.9$ Å) and a monoclinic form (space group $C2$, unit-cell parameters $a = 150.0$, $b = 53.8$, $c = 60.3$ Å, $\beta = 113.2^\circ$). Native data have been collected to 2.0 Å resolution for the orthorhombic form and 1.8 Å for the monoclinic form. A platinum derivative was prepared for the orthorhombic and monoclinic crystals using K_2PtCl_4 and data were collected at several wavelengths to obtain phases by the MAD technique. A partial model has been built from the electron-density maps of both forms and refinement is in progress.

1. Introduction

Apical membrane antigen 1 (AMA1) is a type 1 transmembrane protein originally identified in *Plasmodium knowlesi* (Deans *et al.*, 1982, 1984) and subsequently found in a wide range of plasmodial species (Waters *et al.*, 1990) and other members of the phylum Apicomplexa, such as *Toxoplasma* (Donahue *et al.*, 2000; Hehl *et al.*, 2000) and *Babesia* (Gaffar *et al.*, 2004). The protein is present in the microneme organelles of the developing merozoite of the malaria parasite prior to schizont rupture (Healer *et al.*, 2002; Bannister *et al.*, 2003). At or around the time of schizont rupture, it relocalizes to the merozoite surface (Narum & Thomas, 1994). Antibodies raised to AMA1 can inhibit erythrocyte invasion (Thomas *et al.*, 1984; Kocken *et al.*, 1998; Hodder *et al.*, 2001), strongly suggesting that this surface antigen plays a critical although as yet uncharacterized role during the blood stage of the parasite's life cycle. Indeed, vaccination with recombinant AMA1 confers protective immunity to malaria in animal model systems and the antigen from *P. falciparum* is a leading vaccine candidate that is currently undergoing clinical trials.

AMA1 exhibits a primary structure that is well conserved between different species of *Plasmodium*. The gene encodes a signal sequence and prosequence in the N-terminal region of the protein, followed in succession by an ectoplasmic region, a transmembrane region and a cytoplasmic domain. The ectoplasmic region carries a characteristic signature of 16 cysteine residues that form a defined set of intramolecular disulfide bridges in the native protein (Hodder *et al.*, 1996). The pattern of disulfide bonding has suggested the division of the primary structure into three discrete regions termed domains I, II and III, respectively. Sequence identity within the ectoplasmic region between *P. falciparum* and *P. vivax*, the two dominant human malaria parasites, is 59%. Unlike many other plasmodial surface antigens, AMA1 has no sequence repeats; nevertheless, it is highly polymorphic, with much of the variability residing in domain I (Thomas *et al.*, 1990; Marshall *et al.*, 1996; Polley & Conway, 2001). Since the ectoplasmic region of AMA1 is a promising malaria-vaccine candidate, it is important to relate the location of protecting epitopes and the distribution of polymorphic sites to the three-dimensional structure of the molecule. Accordingly, we have undertaken a structural study of the ectoplasmic region of AMA1 from *P. vivax* (PvAMA1). In this preliminary report, we describe the expression, crystallization and phase determination by the multiwavelength anomalous diffraction (MAD) method using a platinum derivative.

2. Experimental

2.1. Cloning and expression of PvAMA1

Cloning of the ectoplasmic region of PvAMA1 from the Sal I strain (GenBank entry AF063138), comprising residues 43–487 (numbering commencing from the first residue of the signal sequence), was performed essentially as reported for a similar construction of the same antigen (Kocken *et al.*, 1999). The *Pichia pastoris* pPICZα/KM71H expression kit (Invitrogen, Leek, The Netherlands) was used to prepare recombinant clones expressing secreted soluble product carrying a c-*myc*

Table 1
Crystal parameters and data-reduction statistics.

Platinum data sets from the orthorhombic and monoclinic forms were each measured from a single crystal. Values in parentheses refer to the highest resolution shell.

Crystal	Native	Pt (inflection)	Pt (remote)	Native	Pt (peak)	Pt (inflection)	Pt (remote)
Crystal form	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$C2$	$C2$	$C2$	$C2$
Unit-cell parameters							
a (Å)	54.10	53.70	53.76	150.03	151.18	151.74	151.76
b (Å)	76.10	74.84	74.89	53.83	53.83	53.94	53.98
c (Å)	103.93	104.36	104.31	60.29	61.01	61.19	61.21
β (°)				113.2	111.2	111.2	111.2
Z	4			4			
V_M (Å^3 Da^{-1})	2.0			2.2			
Solvent volume (%)	38			44			
Beamline	ID14-4	BM14	BM14	ID14-2	ID14-4	ID14-4	ID14-4
Detector	ADSC Q4	MAR CCD	MAR CCD	ADSC Q4	ADSC Q4	ADSC Q4	ADSC Q4
Wavelength (Å)	1.07197	1.07215	0.88875	0.933	1.07197	1.07234	0.97627
Resolution (Å)	30.0–2.0	30.0–3.1	30.0–3.0	30.0–1.80	30.0–2.60	30.0–2.60	30.0–2.60
Outer shell (Å)	2.07–2.00	3.21–3.10	3.11–3.00	1.90–1.80	2.66–2.60	2.66–2.60	2.66–2.60
Total reflections	147056	37792	55062	152093	39803	40791	46389
Unique reflections	26146	14653†	14420†	40893	23701†	23492†	25350†
Completeness (%)	88.7 (53.4)	98.9 (97.3)‡	98.0 (89.4)‡	99.2 (97.3)	90.1 (56.5)‡	88.0 (58.1)‡	94.6 (89.3)‡
$\langle I\rangle/\langle\sigma(I)\rangle$	23.1 (2.8)	13.6 (2.0)	14.0 (1.6)	10.5 (2.0)	12.6 (3.2)	16.5 (5.3)	13.0 (3.1)
R_{merge}§	0.055 (0.324)	0.083 (0.788)	0.099 (0.672)	0.063 (0.736)	0.066 (0.228)	0.051 (0.139)	0.069 (0.299)

† Each member of a Friedel pair is counted as a separate reflection. ‡ Completeness is given for Friedel pairs § $R_{merge} = \sum|I_h - \langle I_h\rangle|/\sum I_h$, where $\langle I_h\rangle$ is the average over symmetry-related reflections.

marker and a hexahistidine tag at the C-terminal end of the recombinant protein. Potential N-glycosylation sites were mutated as follows: Ser178→Asn, Asn226→Asp and Asn441→Gln. Residue replacements for Ser178 and Asn226 were chosen from the *P. falciparum* homologue, while the replacement for Asn441 is from the *P. chabaudi* sequence (Kocken *et al.*, 1999). Expression from the wild-type *PvAMA1* gene in *P. pastoris*, incidentally, did not present the same problems as did the *P. falciparum* homologue (Kocken *et al.*, 2002), largely because the codon usage of the former is better adapted to this expression system; the AT content of *PvAMA1* is 56%, which is close to that of *P. pastoris* (54%), whereas that of *PfAMA1* is 70%.

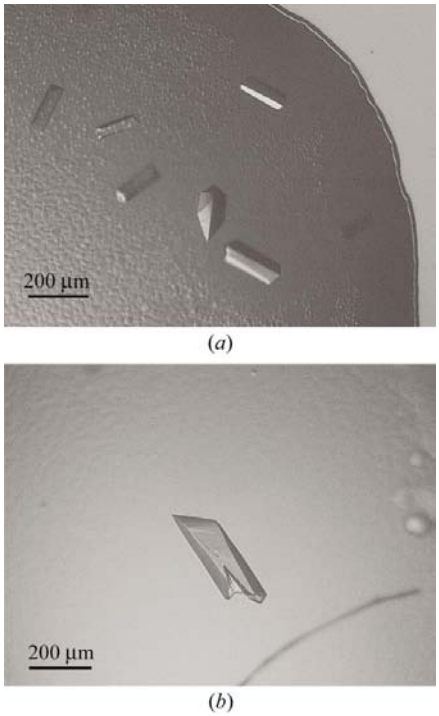

Figure 1
Crystals of PvAMA1. (*a*) Orthorhombic form; (*b*) monoclinic form. A scale is shown in each case.

A 20 ml preculture of recombinant *P. pastoris* was cultured in BMGY (1% yeast extract, 2% peptone, 1.34% yeast nitrogen base, 1% glycerol, 0.4 mg of biotin at a concentration of 0.4 mg ml^{-1}, 0.1 *M* potassium phosphate, pH 6.0) for 65 h at 303 K. 1 l BMMY (BMGY with glycerol replaced by 0.5% methanol) was inoculated with this culture and then grown for 24 h at 303 K with vigorous shaking. Cells were subsequently harvested by centrifugation at 3000*g* for 5 min and then resuspended in 200 ml of BMMY and grown as an induction culture at 303 K for 72 h. Methanol was added to a final concentration of 0.5% every 24 h. Cells were harvested by centrifugation and the supernatant was filtered using 0.22 µm Millex units.

2.2. Protein purification

2.2.1. Procedure I. The protein was purified by a metallo-affinity procedure as follows, making use of the C-terminal His tag. Culture supernatant was directly applied in batch mode to ProBond resin (Invitrogen, 4 ml slurry per 50 ml supernatant in a Falcon tube) previously equilibrated in 20 m*M* sodium phosphate, 500 m*M* NaCl pH 7.4 and incubated for 3 h at room temperature. The resin was then separated from the supernatant by centrifugation and treated with wash buffer (20 m*M* sodium phosphate, 500 m*M* NaCl pH 6.0) in successive volumes of 25 ml until the absorbance of the supernatant was close to the background. The protein was then eluted with the same wash buffer to which 500 m*M* imidazole had been added.

2.2.2. Procedure II. Since procedure I did not give the expected yields of PvAMA1 from the total recombinant protein, purification on an anion-exchange column was also used. The supernatant was precipitated with 70% ammonium sulfate and the resulting pellet was redissolved in 10 m*M* Tris–HCl buffer pH 8.0 and applied onto an Econo-Pac 10DG column (Biorad) either with or without a preliminary passage over a metallo-affinity column as described in procedure I. The desalted protein was then purified on a MonoQ column (Pharmacia) using an NaCl gradient in 20 m*M* Tris–HCl buffer pH 8.0.

The protein fraction not absorbed by the metallo-affinity resin did not react either with an anti-His-tag monoclonal antibody or an anti-c-*myc* monoclonal antibody in a Western blot, showing that C-terminal end of the expressed protein was labile towards proteolysis. Indeed, the molecular mass of the purified product was 54 150 ± 50 from procedure I and 51 600 ± 50 from procedure II, as determined by SELDI–TOF MS (surface-enhanced laser-desorption ionization time-of-flight mass spectrometry) retentate chromatography. The former value agrees with the expected molecular mass of the full-length product (54 175), whereas the latter corresponds well to that expected for the recombinant product cleaved at the junction of the AMA1 C-terminus and the c-*myc* marker (51 567). These results, moreover, confirm that there is no O-glycosylation. On average, the yield of total purified protein was 60 mg per litre of supernatant, with the full-length and truncated products obtained in about equal proportions.

2.3. Crystallization and heavy-atom derivatives

Crystals were grown by vapour diffusion using the hanging-drop technique at 291 K. Initial crystallization conditions were screened by the sparse-matrix sampling approach; microcrystals were obtained in conditions 40 and 41 described by Jancarik

Table 2
Heavy-atom refinement.

Space group	Orthorhombic		Monoclinic		
Data set	Inflection	Remote	Peak	Inflection	Remote
f'	−20.9	−6.8	−22.8	−18.3	−7.6
f''	7.6	11.9	9.6	14.2	8.9
Resolution† (Å)		30.0–4.2			30.0–3.1
FOM (centric)		0.360			0.282
FOM (acentric)		0.554			0.431

† Resolution of heavy-atom refinement.

& Kim (1991). Optimization was then pursued by modification of the buffer components and the addition of various additives. The best crystals were obtained with a crystallization mixtire that typically consisted of 10–12%(*w*/*v*) PEG 3350, 100–200 m*M* imidazole pH 7.0, 5–10%(*v*/*v*) 2-propanol and 1%(*v*/*v*) DMF or 3%(*v*/*v*) *t*-butanol. The protein solution was diluted 40–60% with the crystallization buffer, giving a final concentration of 4–6 mg ml^{-1} in suspended drops of volume 2–3 µl. The crystals were monoclinic for metallo-affinity-purified protein (procedure I) and orthorhombic for MonoQ-purified protein (procedure II) (Fig. 1). The crystalline form obtained thus appears to be determined by the presence or absence of the C-terminal segment containing the c-*myc* and hexahistidine markers.

Orthorhombic and monoclinic crystals of the native protein were prepared for diffraction measurements by flash-freezing in liquid nitrogen after a brief transfer to a cryoprotectant buffer consisting of the original crystallization buffer in which the PEG 3350 concentration had been increased to 30%(*w*/*v*). A platinum derivative of both crystal forms was prepared by soaking crystals in a solution of potassium tetrachloroplatinate(II) (Aldrich) at a concentration of 2 mg ml^{-1} in a buffer of 30%(*w*/*v*) PEG 3350, 100 m*M* sodium cacodylate, 10%(*v*/*v*) 2-propanol pH 7.0 for 24 h. Crystals were mounted by flash-freezing in liquid nitrogen without back-washing, using the same solution as a cryoprotector.

2.4. Data collection and preliminary structure analysis

Data were collected at the ESRF, Grenoble, using beamlines BM14, ID14-2 and ID14-4. Crystals were maintained at 100 K in a stream of cold nitrogen gas during measurement, but they proved to be very sensitive to radiation damage, thus limiting the total exposure that could be tolerated in the beam. The orthorhombic crystals usually grew to no larger than about 50 × 50 × 200 µm, whereas the monoclinic form grew to maximum dimensions of about 100 × 100 × 400 µm. The larger monoclinic crystals, however, were prone to damage upon freezing and smaller sizes were therefore favoured for diffraction measurements. Diffraction intensities were integrated and scaled using the programs *DENZO* and *SCALEPACK*, respectively (Otwinowski & Minor, 1997). Details of measurements and data-processing statistics are given in Table 1.

Since the orthorhombic crystals were obtained first, structure analysis began with this form. Diffraction data for the platinum derivative were measured at beamline BM14 at the ESRF, Grenoble using a single crystal. Measurements were made at two wavelengths, 1.0722 and 0.8887 Å, corresponding to the inflection and remote points of the fluorescence spectrum, respectively. Data could not be measured at more than two wavelengths from the same crystal because of sensitivity to radiation damage and the comparable size of the sample to the X-ray beam on BM14. There was also a large variability in the unit-cell parameters of cryofrozen crystals, which excluded the possibility of using data sets from different crystals in the phasing calculations. Analysis was therefore restricted to data from the best crystal of the platinum derivative. An initial heavy-atom search made with the program *SHELXD* (Schneider & Sheldrick, 2002) found two sites, which were subsequently refined using the program *SHARP* (de La Fortelle & Bricogne, 1997). Heavy-atom parameters were refined at 4.2 Å resolution, beyond which limit the correlation between the anomalous differences of the inflection and remote data sets fell below 0.3. Two additional platinum sites were subsequently found, giving an overall figure of merit (FOM) of 0.36. All four Pt atoms showed high temperature factors. Phases were then calculated to 3.0 Å using these heavy-atom parameters (FOM of 0.22) and were refined by density modification using the programs *SOLOMON* (Abrahams, 1997) and *DM* (Collaborative Computational Project, Number 4, 1994; Cowtan, 1994). In spite of the poor resolution, the resulting electron-density map showed clear indications of the polypeptide main chain in certain regions and a partial model could be built. This gave a preliminary polyalanine model that included 244 of the 467 residues comprising the recombinant protein, which were distributed over seven peptide segments.

At this stage, the monoclinic form of AMA1 was crystallized. A derivative was prepared with K_2PtCl_4 under the same conditions as for the orthorhombic form and diffraction intensities were measured at three wavelengths on ID14-4 at the ESRF: 1.07197 Å (peak), 1.07234 Å (inflection) and 0.97627 Å (remote). These three data sets could be measured from the same crystal because the smaller beam size on ID14-4 and the larger size of the monoclinic crystals allowed successive translations to previously unexposed regions of the sample for each wavelength.

The partial model from the orthorhombic form was then used as a search model in a molecular-replacement calculation to place the molecule in the monoclinic unit cell. The highest peak in the rotation function (peak height of 0.24, second highest peak at 0.22) gave a clear solution in the translation function (highest peak at 0.32, second highest peak at 0.25). An anomalous difference map from the peak data, phased by molecular-replacement solution, showed two very clear peaks that could be related *via* the search model to the two major platinum sites in the orthorhombic crystal form. The heavy-atom parameters were refined with *SHARP* using data to 3.1 Å (a limit based on the fall-off of correlation between anomalous differences from the three MAD data sets), leading to the inclusion of three additional platinum sites that could be related to peaks in the anomalous difference map. This gave an FOM of 0.28. Phases were then calculated to 2.8 Å with these platinum parameters (FOM of 0.26) and refined by density modification with *SOLOMON* and *DM*. As with the orthorhombic form, all Pt atoms had high temperature factors. Nonetheless, the quality of the phases was superior for the monoclinic data and the definition of the resulting electron density allowed easier interpretation. Phasing statistics for both crystal forms are presented in Table 2. The partial orthorhombic model, placed in the monoclinic unit cell by molecular replacement, followed the MAD-phased electron density of the latter crystal form very well. This showed that the heavy-atom phasing

results were coherent between the two crystal forms and were thus correct. Additional residues could be modelled in the monoclinic map as polyalanine and the electron density for side chains in the better defined regions eventually permitted the sequence to be assigned to most of the polypeptide segments comprising the partial model. At this stage, the model included seven separate peptide segments with a total of 268 residues. Native data sets of both crystal forms have also been measured to high resolution (see Table 1) and these are currently being used to refine the atomic model of AMA1 with the programs *ARP/wARP* (Morris *et al.*, 2002) and *REFMAC* (Murshudov *et al.*, 1997).

We thank the staff of the ESRF, Grenoble and Dr Martin Walsh for advice and assistance with diffraction measurements on the ID14 and BM14 beamlines. We also thank J.-P. Latge and J. Sarfati-Bert, Unité des Aspergillus, Institut Pasteur for precious advice and discussion. This work was financed by the European Commission (contract QLK2-CT-1999-01293), Centre National de la Recherche Scientifique, the Institut Pasteur and the Biomedical Primate Research Centre. Visits to the ESRF were financed in part by the European Commission (contract HPRI-CT-1999-00022: Access to Research Infrastructures).

References

Abrahams, J. P. (1997). *Acta Cryst.* D**53**, 371–376.
Bannister, L. H., Hopkins, J. M., Dluzewski, A. R., Margos, G., Williams, I. T., Blackman, M. J., Kocken, C. H. M., Thomas, A. W. & Mitchell, G. H. (2003). *J. Cell Sci.* **116**, 3825–3834.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cowtan, K. (1994). *Jnt CCP4/ESF–EAMCB Newsl. Protein Crystallogr.* **31**, 34–38.
Deans, J. A., Alderson, T., Thomas, A. W., Mitchell, G. H., Lennox, E. S. & Cohen, S. (1982). *Clin. Exp. Immunol.* **49**, 297–309.
Deans, J. A., Thomas, A. W., Alderson, T. & Cohen, S. (1984). *Mol. Biochem. Parasitol.* **11**, 189–204.
Donahue, G. G., Carruthers, V. B., Gilk, S. D. & Ward, G. E. (2000). *Mol. Biochem. Parasitol.* **111**, 15–30.
Gaffar, F. R., Yatsuda, A. P., Franssen, F. F. J & de Vries, E. (2004). *Infect. Immun.* **72**, 2947–2955.
Healer, J., Crawford, S., Ralph, S., McFadden, G. & Cowman, A. F. (2002). *Infect. Immun.* **70**, 5751–5758.
Hehl, A. B., Lekutis, C., Grigg, M. E., Bradley, P. J., Dubremetz, J.-F., Ortega-Barria, E. & Boothroyd, J. C. (2000). *Infect. Immun.* **68**, 7078–7086.
Hodder, A. N., Crewther, P. E. & Anders, R. F. (2001). *Infect. Immun.* **69**, 3286–3294.
Hodder, A. N., Crewther, P. E., Matthew, M. L. S. M., Reid, G. E., Moritz, R. L., Simpson, R. J. & Anders, R. J. (1996). *J. Biol. Chem.* **271**, 29446–29452.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kocken, C. H. M., Dubbeld, M. A., van der Wel, A., Pronk, J. T., Waters, A. P., Langermans, J. A. M. & Thomas, A. W. (1999). *Infect. Immun.* **67**, 43–49.
Kocken, C. H. M., van der Wel, A. M., Dubbeld, M. A., Narum, D. L., van de Rijke, F. M., van Gemert, G. J., van der Linde, X., Bannister, L. H., Janse, C., Waters, A. P. & Thomas, A. W. (1998). *J. Biol. Chem.* **273**, 15119–15124.
Kocken, C. H. M., Withers-Martinez, C., Dubbeld, M. A., van der Wel, A., Hackett, F., Blackman, M. J. & Thomas, A. W. (2002). *Infect. Immun.* **70**, 4471–4476.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494
Marshall, V. M., Zhang, L.-X., Anders, R. F. & Coppel, R. L. (1996). *Mol. Biochem. Parasitol.* **77**, 109–113.
Morris, R. J., Perrakis, A. & Lamzin, V. S. (2002). *Acta Cryst.* D**58**, 968–975.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Narum, D. L. & Thomas, A. W. (1994). *Mol. Biochem. Parasitol.* **67**, 59–68.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Polley, S. D. & Conway, D. J. (2001). *Genetics*, **158**, 1505–1512.
Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.
Thomas, A. W., Deans, J. A., Mitchell, G. H., Alderson, T. & Cohen, S. (1984). *Mol. Biochem. Parasitol.* **13**, 187–199.
Thomas, A. W., Waters, A. P. & Carr, D. (1990). *Mol. Biochem. Parasitol.* **42**, 285–288.
Waters, A. P., Thomas, A. W., Deans, J. A., Mitchell, G. H., Hudson, D. E., Miller, L. H., McCutchan, T. F. & Cohen, S. (1990). *J. Biol. Chem.* **265**, 17974–17979.