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

Crystallization and crystallographic analysis of the apo form of the orange protein (ORP) from *Desulfovibrio gigas*

Shabir Najmudin,*‡ Cecília Bonifácio, Américo G. Duarte, Sofia R. Pualeta, Isabel Moura, José J. G. Moura and Maria J. Romão

REQUIMTE, Centro de Química Fina e Biotecnologia, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

‡ Present address: CIISA – Faculdade de Medicina Veterinária, Universidade Técnica de Lisboa, Avenida da Universidade Técnica, 1300-477 Lisboa, Portugal.

Correspondence e-mail: shabir@dq.fct.unl.pt

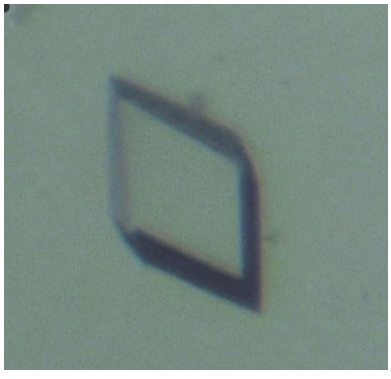
Received 8 May 2009
Accepted 17 June 2009

The orange-coloured protein (ORP) from *Desulfovibrio gigas* is a 12 kDa protein that contains a novel mixed-metal sulfide cluster of the type [S_2MoS_2-CuS_2MoS_2]. Diffracting crystals of the apo form of ORP have been obtained. Data have been collected for the apo form of ORP to 2.25 Å resolution in-house and to beyond 2.0 Å resolution at ESRF, Grenoble. The crystals belonged to a trigonal space group, with unit-cell parameters a = 43, b = 43, c = 106 Å.

1. Introduction

The orange-coloured protein (ORP) from *Desulfovibrio gigas* is a 12 kDa protein that contains a noncovalently attached novel mixed-metal sulfide cluster of the type [$S_2MoS_2CuS_2MoS_2$] (Bursakov *et al.*, 2004; George *et al.*, 2000). It was obtained from a culture grown in a lactate–sulfate medium under anaerobic conditions. It has been proposed to be involved in a metabolic pathway that is activated when the growth conditions are switched from syntrophic to sulfate-reducing (Pauleta *et al.*, 2007). A structural homology search classifies ORP as a member of the nitrogenase accessory factor-like superfamily. The structures of three members of this family are known and display a ribonuclease H-like fold. Two NMR solution structures, one from *Thermotoga maritima* (TM1290; PDB code 1rdu; Etezady-Esfarjani *et al.*, 2004) and the other from *Methanobacterium thermoautotrophicum* (MTH1175; PDB code 1eo1; Cort *et al.*, 2000), have higher sequence identity to ORP but have unidentified function. Both NMR (PDB code 1t3v) and X-ray (PDB code 1o13) structures of a homologue from *T. maritima* (TM1816) are also known. TM1816 is thought to be an NifB protein that is involved in FeMo-co biosynthesis (Columbus *et al.*, 2005).

We have heterologously expressed and purified the apo form of ORP (apo-ORP). Diffracting crystals of apo-ORP have been obtained and data have been collected for apo-ORP to 2.25 Å resolution in-house and to beyond 2.0 Å resolution at ID23-EH1, ESRF, Grenoble. Here, we report our preliminary findings.



2. Materials and methods

2.1. Protein isolation and purification

The *D. gigas* orange protein was heterologously obtained in the apo form as previously described in Pauleta *et al.* (2007), with several modifications. *Escherichia coli* BL21 (DE3) carrying the plasmid pET-ADORP (Pauleta *et al.*, 2007) was grown in 2×YT medium (16 g l^{-1} tryptone, 10 g l^{-1} yeast extract and 5 g l^{-1} NaCl) containing 100 µg ml^{-1} ampicillin. Growth was carried out at 310 K in an orbital shaker at 230 rev min^{-1} until the $OD_{600\,nm}$ reached 0.6. Protein expression was induced by the addition of 0.5 m*M* IPTG and growth was maintained under the same conditions for 16 h. The cells were harvested by centrifugation at 8000*g* and resuspended in 10 m*M* Tris–HCl pH 7.6 with a yield of 8.8 g of cells (wet weight) per litre of culture. Cells were disrupted using a French press and cell debris was removed by low-speed centrifugation (8000*g*). The soluble extract

was separated from the membrane extract by ultracentrifugation at 138 000*g*.

Purification of the recombinant apoprotein was performed in two chromatographic steps at 277 K. The soluble extract was concentrated in a Dia-Flo apparatus over a YM3 membrane and loaded onto a Superdex 75 column (Amersham; 26 × 700 mm) equilibrated with 300 m*M* Tris–HCl pH 7.6. The fractions containing apo-ORP were pooled, concentrated and desalted using a dialysis membrane. In the second chromatographic step, anion-exchange chromatography, the sample was loaded onto a DEAE-Sepharose fast-flow column (Amersham; 26 × 140 mm) equilibrated with 10 m*M* Tris–HCl pH 7.6. The adsorbed proteins were eluted with a linear gradient between 10 and 500 m*M* Tris–HCl pH 7.6. Apo-ORP was eluted at 125 m*M* Tris–HCl pH 7.6.

The purity of apo-ORP was analysed in all the chromatographic steps by 16% acrylamide–bisacrylamide Tris–tricine SDS–PAGE (Fig. 1). The fractions containing pure apo-ORP were pooled, concentrated, buffer-exchanged to 0.1 *M* Tris–HCl pH 7.6 and stored at 193 K. The apo-ORP concentration was determined using the BCA kit (Sigma) with BSA (bovine serum albumin) as a standard. ORP is a 117-amino-acid residue peptide (Uniprot accession No. Q5DUA6) with no cysteinyl and three methionyl residues. Our apo-ORP has an extra three amino-acid residues, Ala-Ser-His, at the N-terminus (a cloning artifact), giving it a molecular weight of 12 kDa.

2.2. Crystallization

Crystallization conditions were screened by both the sitting-drop and the hanging-drop vapour-phase diffusion methods using an in-house modified version of a simplified screen (Gao *et al.*, 2005). Drops consisting of 1 µl apo-ORP at 15 mg ml^{-1} and 1 µl from the 1 ml reservoir solution were prepared at 277 or 293 K. Protein crystals were obtained within a few days using the following conditions: 0.1 *M* sodium acetate, 0.2 *M* NaH_2PO_4 pH 4.5, 15–30% PEG 3K or PEG 8K at 277 K and 0.1 *M* sodium acetate and 1 *M* $MgSO_4$ at 293 K (Fig. 2). The crystals were flash-cooled in liquid nitrogen after soaking in a cryoprotectant solution [the crystallization mother liquor containing 30%(*v*/*v*) glycerol] for a few seconds. The crystals obtained at 277 K using the hanging-drop method gave better diffraction.

2.3. Data collection and processing

Initial data sets were collected on a home source using Cu $K\alpha$ X-ray radiation (Enraf–Nonius rotating-anode generator operated at 5 kW and 100 mA equipped with a MAR Research imaging plate). Further native data sets (at λ = 0.9793 and 2.0664 Å) were collected on beamline ID23-EH1 at the ESRF (Grenoble, France) using a Quantum 4 charge-coupled device detector (ADSC) with the crystal cooled to 100 K using a Cryostream (Oxford Cryosystems Ltd). All data sets were processed using the programs *MOSFLM* (Leslie, 1992) and *SCALA* (Kabsch, 1978) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The crystals belonged to the trigonal space group $P3_121$ or $P3_221$ as indicated by *POINTLESS* (Evans, 2006). Plots of the acentric and centric moments and the cumulative intensity distribution from the *SCALA* output indicated that the crystal form was indeed $P3_121$ or $P3_221$ and not the lower symmetry twin parental *P*32. The anisotropy analysis plot calculated by *TRUNCATE* (French & Wilson, 1978) showed the data to be isotropic. The Matthews coefficient (V_M = 2.08 Å^3 Da^{-1}) indicated the presence of one molecule in the asymmetric unit and a solvent content of 40% (Matthews, 1968). The crystals were of a delicate nature and suffered radiation damage on exposure both in-house and at ID23-EH1, curtailing the amount of data collected per crystal. At the home source the crystals diffracted to 2.25 Å resolution. Several complete and highly redundant data sets were collected for apo-ORP crystals grown in crystallization buffer with 10–30% glycerol as a cryoprotectant. A total of 90–120° of data were collected with 1° oscillations before the crystal suffered icing problems and/or radiation damage. Statistics for the best data are given in Table 1. Attempts to solve the structure by molecular replacement using an ensemble of

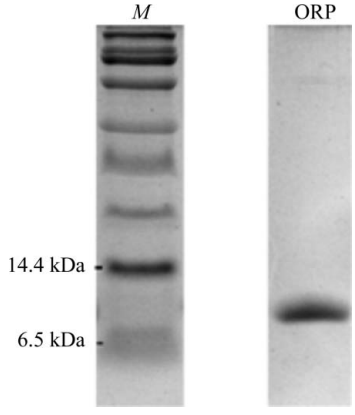
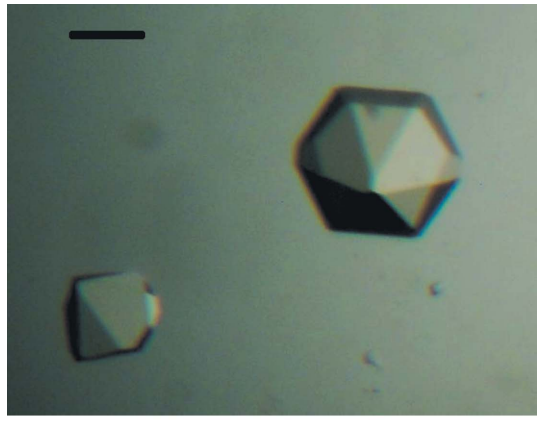
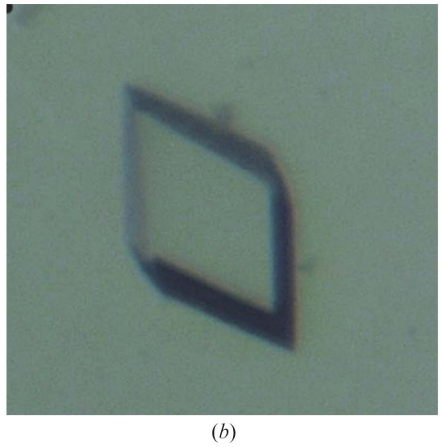


Figure 1
Tricine SDS–PAGE analysis of purified recombinant apo-ORP. Lane *M* contains standard markers (broad-range, Bio-Rad).



(*a*)



(*b*)

Figure 2
Crystals of apo-ORP obtained (*a*) at 277 K and (*b*) at 293 K by hanging-drop vapour diffusion in the presence of (*a*) 0.1 *M* sodium acetate, 0.2 *M* NaH_2PO_4 pH 4.5 and 8–25% PEG 3K and (*b*) 0.1 *M* sodium acetate and 1 *M* $MgSO_4$. The largest crystals are approximately 0.4 mm in their longest dimension (the scale bar is 0.1 mm in length).

Table 1
X-ray crystallography data-collection statistics.

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

Data set	In-house	High resolution	SAD
X-ray source	Cu $K\alpha$ rotating anode, in-house	ID23-EH1, ESRF	ID23-EH1, ESRF
Wavelength (Å)	1.5418	0.97934	2.0664
Space group	$P3_121$	$P3_121$	$P3_121$
Unit-cell parameters			
$a = b$ (Å)	42.65	43.08	42.94
c (Å)	105.14	106.08	105.55
Resolution limits (Å)	30.2–2.25	106.0–2.00	52.8–2.05
No. of observations	27434 (911/3326)	73707 (1657/4251)	62629 (2255/1298)
No. of unique observations	9680 (285/1381)	7446 (289/710)	6327 (272/314)
Multiplicity	2.8 (3.2/2.4)	9.9 (5.7/6.0)	9.9 (8.3/4.1)
Completeness (%)	95.4 (91.4/91.4)	90.7 (92.8/62.2)	84.1 (94.3/30.3)
$\langle I/\sigma(I)\rangle$	8.2 (22.0/3.6)	33.7 (31.1/16.9)	26.4 (29.6/7.6)
$R_{\rm merge}$†	0.201 (0.035/0.482)	0.050 (0.038/0.092)	0.069 (0.059/0.112)
$R_{\rm p.i.m.}$‡	0.201 (0.026/0.517)	0.017 (0.019/0.041)	0.023 (0.024/0.063)
Wilson B factor (Å^2)	26.8	22.1	30.8

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith measurement of reflection hkl and $\langle I(hkl)\rangle$ is the mean value of $I_i(hkl)$ for all i measurements. ‡ $R_{\rm p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$ and is a measure of the quality of the data after averaging the multiple measurements.

NMR structures [PDB codes 1rdu (Etezady-Esfarjani *et al.*, 2004) and 1eol (Cort *et al.*, 2000), with sequence identities of 35% and 31%, respectively, from a *PSI-BLAST* alignment (Altschul *et al.*, 1997)] were unsuccessful. These data were also not good enough to give a sufficient signal for the S atoms of Met for a SAD experiment at $\lambda = 1.5418$ Å. At the ID23-EH1 beamline at the ESRF a total of 210° of data for each crystal at $\lambda = 0.9793$ and 2.0664 Å were collected with 1° oscillations before the crystals started suffering serious radiation damage. These crystals were three to four months old by the time they were taken to the ESRF and this may have contributed to their fragile nature.

ORP contains three methionines and the crystals were grown in 0.2 *M* NaH_2PO_4. Therefore, at $\lambda = 2.0664$ Å one would expect some anomalous signal for S and P atoms (at approximately 6 keV, $f'_S = 0.96$ and $f'_P = 0.75$) if they are sufficiently ordered. Feeding this data set into *SOLVE*/*RESOLVE* (Terwilliger, 2004) eight sites were found, but only 2% of the structure could be built. *SHELXD* and *SHELXE* (Sheldrick, 2008) were also used. 27 heavy-atom sites with an occupancy of more than 10% were located. The pseudo-free correlation coefficient, contrast and connectivity figures of merit given by *SHELXE* for the correct heavy-atom enantiomer ($P3_121$) were 48.0%, 0.348 and 0.783, respectively, as opposed to 39.8%, 0.343 and 0.782, respectively, for the wrong hand ($P3_221$). The best molecular-replacement solutions were consistently in $P3_221$ for the data collected in-house.

The program *BALBES* (Long *et al.*, 2008) was used to search for a solution for the data collected at $\lambda = 0.9793$ Å processed in both $P3_121$ and $P3_221$. The best solution was found using the X-ray structure of the NifB-like protein TM1816 from *T. maritima* (PDB code 1o13) from the JCSB project (Columbus *et al.*, 2005), with an R and $R_{\rm free}$ of 0.399 and 0.447, respectively, in space group $P3_121$. The sequence identity of ORP with PDB entry 1o13 is 25%. Automated model building using *ARP*/*wARP* (Cohen *et al.*, 2004) with the solution from *BALBES* and the X-ray data collected at $\lambda = 0.9793$ Å located 111 amino-acid residues with a sequence coverage of 91% and an estimated correctness of the model of 99%. Structure completion and analysis is ongoing.

This work was supported in part by Fundação para a Ciência e a Tecnologia (Lisbon, Portugal) through grant POCI/QUI/55350/2004 and the individual grants SFRH/BPD/20357/2004 (SN) and SRFH/BD/39009/2007 (AD). The authors would like to thank Drs Ana L. Carvalho and José Trincão for help with data collection and Alice S. Pereira for help with protein isolation and purification.

References

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
Bursakov, S. A., Gavel, O. Y., Di Rocco, G., Lampreia, J., Calvete, J., Pereira, A. S., Moura, J. J. & Moura, I. (2004). *J. Inorg. Biochem.* **98**, 833–840.
Cohen, S. X., Morris, R. J., Fernandez, F. J., Ben Jelloul, M., Kakaris, M., Parthasarathy, V., Lamzin, V. S., Kleywegt, G. J. & Perrakis, A. (2004). *Acta Cryst.* D**60**, 2222–2229.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Columbus, L., Peti, W., Etezady-Esfarjani, T., Herrmann, T. & Wüthrich, K. (2005). *Proteins*, **60**, 552–557.
Cort, J. R., Yee, A., Edwards, A. M., Arrowsmith, C. H. & Kennedy, M. A. (2000). *J. Struct. Funct. Genomics*, **1**, 15–25.
Etezady-Esfarjani, T., Herrmann, T., Peti, W., Klock, H. E., Lesley, S. A. & Wüthrich, K. (2004). *J. Biomol. NMR*, **29**, 403–406.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
French, S. & Wilson, K. (1978). *Acta Cryst.* A**34**, 517–525.
Gao, W., Li, S. & Bi, R. (2005). *Acta Cryst.* D**61**, 776–779.
George, G. N., Pickering, I. J., Yu, E. Y., Prince, R. C., Bursakov, S. A., Gavel, O. Y., Moura, I. & Moura, J. J. G. (2000). *J. Am. Chem. Soc.* **122**, 8321–8322.
Kabsch, W. (1978). *Acta Cryst.* A**34**, 827–828.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst.* D**64**, 125–132.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Pauleta, S. R., Duarte, A. G., Carepo, M. S., Pereira, A. S., Tavares, P., Moura, I. & Moura, J. J. G. (2007). *Biomol. NMR Assign.* **1**, 81–83.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Terwilliger, T. (2004). *J. Synchrotron Rad.* **11**, 49–52.