Tiago M. Bandejas,^{a,b} Micael C. Freitas,^{a,b} Dennis Petrasch,^c Arnulf Kletzin^c and Carlos Frazão^{b*}

^aIBET – Instituto de Biologia Experimental e Tecnológica, Apartado 12, 2780-901 Oeiras, Portugal, ^bITQB–UNL, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Avenida da República, 2780-157 Oeiras, Portugal, and ^cMicrobiology – Sulfur Biochemistry and Microbial Bioenergetics, Technische Universität Darmstadt, Schnittspahnstrasse 10, 64287 Darmstadt, Germany

Correspondence e-mail: frazao@itqb.unl.pt

Received 13 February 2013

Accepted 26 March 2013

SAD phasing towards structure determination of a thermostable Rieske ferredoxin with a novel stabilizing disulfide bridge

Rieske proteins and Rieske ferredoxins are ubiquitous electron-transfer metalloproteins that are characterized by a [2Fe–2S] cluster coordinated by pairs of cysteine and histidine residues. The thermoacidophilic archaeon *Acidianus ambivalens* contains a Rieske ferredoxin termed RFd2, which has an hitherto unknown additional region of 40–44 residues at the C-terminus with a Cx₃C motif that introduces a novel disulfide bond within the Rieske fold. RFd2 was crystallized with the aim of determining its three-dimensional structure in order to understand the contribution of this as yet unique disulfide bridge to the function and stability of RFd2. RFd2 crystals were successively improved, increasing their diffraction to 1.9 Å resolution. Molecular replacement did not solve the RFd2 structure, but a highly multiple in-house diffraction data set collected at the Cu K α edge led to solution of the phase problem.

1. Introduction

Rieske proteins and Rieske ferredoxins (RFds) are small (<130-amino-acid) iron–sulfur proteins that act as electron-transfer components. They contain the Rieske centre, a [2Fe–2S] cluster in which one of the Fe atoms is coordinated by the N^{δ1} atoms of two histidine side chains and the other by two cysteine S^γ ligands (Schmidt & Shaw, 2001). The Fe–S cluster is embedded in a small domain (usually about 100 amino acids) with an all- β fold. The cysteine and histidine residues are easily recognized by their characteristic signature: CxH-x_{15–18}-CxxH (Fig. 1). The Rieske proteins are membrane-bound components of, for example, cytochrome *bc*₁ complexes, which contain the Rieske domain fused to a transmembrane helix. At pH 7 they have reduction potentials of +150–350 mV (or greater; Brown *et al.*, 2008). In contrast, RFds are soluble proteins consisting of just the Rieske domain and are characterized by fairly low reduction potentials, typically of around –150 mV (Brown *et al.*, 2008; Schmidt & Shaw, 2001). All Rieske clusters demonstrate pH-dependency of their midpoint reduction potential, which changes according to the protonation of the cluster-coordinating histidine residues (Brown *et al.*, 2008; Zu *et al.*, 2003).

In addition, respiratory Rieske proteins may also include one, or sometimes two, disulfide bridges (Bönisch *et al.*, 2002). The disulfide bridge that is conserved in the Rieske proteins of membrane-bound electron-transport complexes is localized in the vicinity of the cluster, modulating its reduction potential (Zu *et al.*, 2002). Leggate & Hirst (2005) showed that it might have a protective function against hydrolysis in the oxidized state of high-potential Rieske clusters. The disulfide bridge, which is formed by conserved cysteine residues in the vicinity of the cluster ligands, is present in the archaeal respiratory Rieske proteins SoxF and SoxL and in sulredoxins (soluble RFds from archaea; Iwasaki *et al.*, 1995), but not in other RFds. The SoxF proteins from *Sulfolobus acidocaldarius* and other Crenarchaeota are the only Rieske proteins with a second disulfide bridge within the N-terminal part of the protein (Bönisch *et al.*, 2002).

While searching for proteins with Rieske centres in the thermoacidophilic archaeon *Acidianus ambivalens*, we came upon a novel RFd subtype named RFd2 (Fig. 1). This protein has two prominent features: (i) a truncation at the N-terminus where the first strand of

the Rieske fold should reside and (ii) a new motif in the C-terminal part that includes a pair of cysteines: SKTPC₂₋₃C (Botelho *et al.*, 2010). Homologous proteins with this novel region were only found in the Crenarchaeota. Comparison with other RFds with known structure gave sequence similarity close to the conserved Cys/His motifs but not beyond.

The brightly coloured RFd2 exhibits characteristic absorption bands at 337, 456 and 570 nm and a shoulder at 490 nm. At 298 K and pH 7 the Fe–S cluster shows a high reduction potential compared with canonical RFds of +48 mV and a pK_a of 10.1, which shift to +77 mV and 8.9, respectively, upon reduction of the disulfide (Botelho *et al.*, 2010). Concomitantly, the melting temperature of 372 K decreases by 9 K, clearly suggesting a stabilizing structural role for the disulfide bridge. These novel features prompted us to unveil the structure of the RFd2 protein; here, we report its preliminary crystallographic characterization.

2. Experimental procedures and results

2.1. Sequence comparisons and definition of a novel protein region

Comparison of the *A. ambivalens* RFd2 protein sequence against the nonredundant protein database of GenBank with *PSI-BLAST* (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) resulted in numerous hits for the 62 N-terminal residues with known Rieske ferredoxins or domains, among which were several RFds with known three-dimensional structure, such as naphthalene dioxygenase from *Pseudomonas putida* (PDB entry 2qpz; 34% identity; Brown *et al.*, 2008). Separate *PSI-BLAST* searches with the remaining 44 residues including the disulfide-forming cysteines (Botelho *et al.*, 2010) only resulted in matches with 12 other crenarchaeal RFd2 homologues from the Sulfolobales order, including *Caldivirga maquilengensis* (Fig. 1), even when the search parameters were relaxed. These results show that the RFd2 family of proteins is composed of two regions which could form flexible building blocks to accommodate different but as yet unknown *in vivo* functions.

2.2. Gene expression and protein purification

The *rfd2* gene was expressed and the protein was purified as described previously (Botelho *et al.*, 2010); the protein was eluted as a monomer by gel filtration and judged to be homogeneous by SDS–PAGE and the typical UV–Vis absorption spectrum (Botelho *et al.*, 2010). The pink protein solution was concentrated to 9.5 mg ml^{−1} by ultrafiltration using a 10 kDa cutoff membrane (Amicon Ultra, Millipore), aliquoted, flash-frozen in liquid nitrogen and stored at 193 K. The protein was quantified using the method of Bradford (1976).

The iron content was quantified for the purified protein using the method of Fischer & Price (1964). The final RFd2 sample contained 2.4 ± 0.1 iron ions per molecule, which is in accordance with the presence of an intact [2Fe–2S] centre in the purified protein.

2.3. Crystallization and cryoprotection

Crystallization screening was carried out in round-bottom 96-well CrystalQuick plates (Greiner Bio-One) using nanolitre-scale drops produced by a Cartesian crystallization-robot dispensing system (Genomics Solutions); RFd2 solution (9.5 mg ml^{−1} in 70 mM Tris–HCl pH 7.7, 150 mM NaCl) was mixed with precipitant solutions from commercially available screening kits using 100:100, 100:200 and 200:100 nl volumes of protein and precipitant solution, respectively. The most promising crystallization conditions were obtained with ammonium sulfate, polyethylene glycol or trisodium citrate as precipitants, where pink/orange-coloured crystals could be observed (Fig. 2). These conditions were scaled up, but most yielded multiple and/or twinned crystals. Only conditions containing ammonium sulfate resulted in single crystals. After optimization, larger single crystals (Figs. 2b and 2c) were produced from a drop consisting of 1 µl volumes of both protein and precipitant solution (0.1 M Bis-Tris pH 6.5, 2.3 M ammonium sulfate). RFd2 crystals could only be cryopreserved by the stepwise addition of sodium malonate to a final concentration of 1.6 M.

Figure 1

Sequence alignment of *A. ambivalens* RFd2 with close homologues from the Crenarchaeota showing GenBank identification numbers, the Fe–S cluster-coordinating residues (*) and disulfide-bridge-forming residues (linked by a line); upper block, the Fe–S-coordinating region, which showed high similarity to other RFds in *BLASTP* searches; lower block, the disulfide region, which showed no similarity to other proteins in relaxed *PSI-BLAST* searches.

2.4. X-ray data collection and processing

During the crystal optimization procedure, several diffraction data sets were measured in-house or at the ESRF and problems such as crystal disorder or twinning were successively overcome. The best diffraction data were obtained in-house at the Cu $K\alpha$ wavelength using a Bruker AXS MICROSTAR rotating-anode X-ray generator with Montel mirrors coupled to a Bruker AXS PROTEUM diffractometer and a PLATINUM135 CCD detector. Diffraction images were collected at 100 K under a nitrogen cryostream using fine-slicing frames (0.2° oscillations in two passes) with 20 s exposure over a total of 2444° of crystal rotation through 30 oscillation sweeps of the φ , χ or ω axes departing from different initial orientations in order to randomize systematic absorption errors and to improve the data accuracy and precision. Images were indexed with *PROTEUM* and spots were integrated with *SAINT*. The obtained intensities were scaled with *SADABS* and the data were merged and quality indicators were calculated with *XPREF* (Bruker AXS). No σ cutoff was applied during data collection. Crystallographic characterization and data statistics are presented in Table 1.

2.5. Solution of the phase problem

Crystal-content analysis assuming the presence of one RFd2 molecule in the asymmetric unit led to the estimation of 45% solvent content (Matthews, 1968). However, we did not succeed in solving the phase problem by molecular replacement (MR) using the programs *Phaser* (McCoy *et al.*, 2007), *MOLREP* (Vagin & Teplyakov, 2010) or *COMO* (Jogl *et al.*, 2001). As search structures, we tried either the PDB entry with the highest identity, the Rieske ferredoxin from the naphthalene 1,2-dioxygenase system of *P. putida* (PDB entry 2qpz; Brown *et al.*, 2008) with 34% identity, or an ensemble of four structures with identities of 30–34% where external variable regions and peptide termini were trimmed off: the oxygenase Rieske ferredoxins with PDB entries 2qpz (from *P. putida*; Brown *et al.*, 2008), 2de6 (from *Janthinobacterium*; Ashikawa *et al.*, 2006) and 2i7f (from *Sphingobium yanoikuyae* B1; Ferraro *et al.*, 2007) together with PDB entry 3d89 from *Mus musculus*, which is as yet without an assigned function but shows a resemblance to the Rieske ferredoxins of bacterial aromatic dioxygenases (Levin *et al.*, 2008). Additionally, we also had no success using the MR pipeline *MrBUMP* (Keegan &

Table 1

Diffraction data-collection and processing parameters.

Values in parentheses are for the highest resolution shell.

Radiation source	Bruker AXS MICROSTAR rotating anode
Wavelength (Å)	1.5418
Space group	$P6_322$
Unit-cell parameters (Å)	$a = b = 53.33$, $c = 135.30$
Resolution (Å)	38.14–1.90 (1.94–1.90)
No. of unique reflections	9729 (481)
Completeness (%)	99.9 (99.8)
Multiplicity	70.7 (21.2)
$\langle I/\sigma(I) \rangle$	52.4 (4.1)
$R_{\text{p.i.m.}}^\dagger$	0.006 (0.132)
$R_{\text{merge}}^\ddagger$	0.062 (0.608)
$\langle E^2 - 1 \rangle$	0.689
No. of molecules in asymmetric unit	1
V_M (Å ³ Da ^{−1})	2.24
Solvent content (%)	45.1
Wilson B (Å ²)	31.1

$^\dagger R_{\text{p.i.m.}} = \sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i$, where $N(hkl)$ is the data multiplicity, $I_i(hkl)$ is the observed intensity and $\langle I(hkl) \rangle$ is the average intensity of multiple observations of symmetry-related reflections. It is an indicator of the precision of the final merged and averaged data set (Weiss, 2001). $^\ddagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl) \rangle$ is the average intensity of multiple observations of symmetry-related reflections.

Winn, 2008) within the *CCP4* suite (Winn *et al.*, 2011) or with the *BALBES* server (Long *et al.*, 2008).

We then decided to use the anomalous signal inherent to the Fe and S atoms of RFd2 to determine the crystal structure. A diffraction data set was measured in-house with very high multiplicity and the phase problem was readily solved by single anomalous diffraction (SAD) using the *HKL2MAP* GUI (Pape & Schneider, 2004), which relies on (i) *SHELXC* to analyse the diffraction data, estimate anomalous difference factors and produce initial phase shifts for the reflections corresponding to the largest normalized anomalous differences, (ii) *SHELXD* to determine the anomalous heavy-atom substructure and (iii) *SHELXE* to produce iterative phase improvement by density modification and to construct a preliminary polyaniline model (Sheldrick, 2010). The obtained 1.9 Å resolution diffraction data contained anomalous signal to 2.3 Å resolution, where the correlation coefficient between the anomalous differences reached 30% (Schneider & Sheldrick, 2002). The SAD heavy-atom substructure consisted of two sites 2.9 Å apart and with occupancies

Figure 2

Successive improvement of the *A. ambivalens* RFd2 crystals. (a) Crystal A ($400 \times 150 \times 350 \mu\text{m}$) was grown in 1.6 M trisodium citrate pH 6.5, diffracted to 2.6 Å resolution and displayed twinning. (b) Crystal B ($600 \times 450 \times 30 \mu\text{m}$) grew from 0.1 M Bis-Tris pH 6.5, 1.6 M ammonium sulfate and gave diffraction data to 2.15 Å resolution on the ESRF microfocus beamline ID23-EH2. (c) Crystal C ($300 \times 150 \times 50 \mu\text{m}$) was grown under similar conditions to crystal B except for the higher precipitant concentration (2.3 M ammonium sulfate) and the use of a 1:1 protein:crystallization solution volume ratio. Using home-source X-ray radiation, the crystal diffracted to 1.9 Å resolution. Scale bars correspond to $\sim 0.3 \text{ mm}$.

Figure 3

SAD electron-density maps and polyaniline model of RFd2. (a) Electron density of the Rieske [2Fe–2S] cluster and surrounding polypeptide, showing the two Fe atoms (stars) 2.86 Å apart within 10σ ($21 \text{ e}^- \text{ \AA}^{-3}$) blobs and the coordinating S atoms as a 1.5σ ($1.6 \text{ e}^- \text{ \AA}^{-3}$) contoured mesh. (b) View of three β -strands at 1.5σ density containing the polyaniline model (atomic bonds are shown as sticks) showing the characteristic β -sheet hydrogen-bonding pattern. The images were produced with *Coot* (Emsley & Cowtan, 2004).

of 1.0, while the remaining peaks refined to occupancies below 0.21 and were considered to be spurious. The resulting anomalous dimetal centre was therefore compatible with a [2Fe–2S] Rieske cluster with a typical Fe–Fe distance of $\sim 2.7 \text{ \AA}$ (Bönisch *et al.*, 2002; Fig. 3a). The *SHELXE* autotraced model consisted of a single 99-residue polyaniline chain from the expected 114 residues of the recombinant RFd2 including the His₆ tag, with a map correlation coefficient of 43%. In addition to a diamond-shaped density which corresponds to the Rieske [2Fe–2S] centre, the experimental SAD electron-density maps also showed clear polypeptide secondary-structure features (Fig. 3b). Sequence assignment, model completion and structural refinement of RFd2 are in progress.

3. Concluding remarks

RFd2 represents a new Rieske ferredoxin subfamily as shown by sequence analysis (Botelho *et al.*, 2010; Fig. 1) and the inadequacy of available three-dimensional models of Rieske centres in attempts to determine the structure by molecular replacement. The process of crystallographic structure determination involved laborious crystal-improvement work. The $1.9 \text{ \AA}$ resolution in-house diffraction with 70-fold multiplicity contained anomalous signal with a correlation coefficient between the anomalous differences of above 30% to $2.3 \text{ \AA}$ resolution (Schneider & Sheldrick, 2002), which was sufficient to solve the crystal phase problem from scratch and to obtain an initial polyaniline model that accounted for 87% of the RFd2 residues. With the elucidation of the RFd2 structure, the relative localization of the new disulfide bridge relative to the Rieske centre will be clarified.

The authors would like to thank the staff of ID23 at the ESRF in Grenoble for the use of synchrotron radiation. TMB is supported by Programa Ciência (POPH/QREN). MCF is supported by an IBET

fellowship. This work was supported by Fundação para a Ciência e a Tecnologia through grant No. PEst-OE/EQB/LA0004/2011.

References

- Ashikawa, Y., Fujimoto, Z., Noguchi, H., Habe, H., Omori, T., Yamane, H. & Nojiri, H. (2006). *Structure*, **14**, 1779–1789.
- Bönisch, H., Schmidt, C. L., Schäfer, G. & Ladenstein, R. (2002). *J. Mol. Biol.* **319**, 791–805.
- Botelho, H. M., Leal, S. S., Veith, A., Prosiniecki, V., Bauer, C., Fröhlich, R., Kletzin, A. & Gomes, C. M. (2010). *J. Biol. Inorg. Chem.* **15**, 271–281.
- Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
- Brown, E. N., Friemann, R., Karlsson, A., Parales, J. V., Couture, M. M.-J., Eltis, L. D. & Ramaswamy, S. (2008). *J. Biol. Inorg. Chem.* **13**, 1301–1313.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst. D* **60**, 2126–2132.
- Ferraro, D. J., Brown, E. N., Yu, C.-L., Parales, R. E., Gibson, D. T. & Ramaswamy, S. (2007). *BMC Struct. Biol.* **7**, 10.
- Fischer, D. S. & Price, D. C. (1964). *Clin. Chem.* **10**, 21–31.
- Iwasaki, T., Isogai, Y., Iizuka, T. & Oshima, T. (1995). *J. Bacteriol.* **177**, 2576–2582.
- Jogl, G., Tao, X., Xu, Y. & Tong, L. (2001). *Acta Cryst. D* **57**, 1127–1134.
- Keegan, R. M. & Winn, M. D. (2008). *Acta Cryst. D* **64**, 119–124.
- Leggate, E. J. & Hirst, J. (2005). *Biochemistry*, **44**, 7048–7058.
- Levin, E. J., Elsen, N. L., Seder, K. D., McCoy, J. G., Fox, B. G. & Phillips, G. N. Jr (2008). *Acta Cryst. D* **64**, 933–940.
- 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.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Pape, T. & Schneider, T. R. (2004). *J. Appl. Cryst.* **37**, 843–844.
- Schmidt, C. L. & Shaw, L. (2001). *J. Bioenerg. Biomembr.* **33**, 9–26.
- Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst. D* **58**, 1772–1779.
- Sheldrick, G. M. (2010). *Acta Cryst. D* **66**, 479–485.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst. D* **66**, 22–25.
- Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.
- Winn, M. D. *et al.* (2011). *Acta Cryst. D* **67**, 235–242.
- Zu, Y., Couture, M. M.-J., Kolling, D. R., Crofts, A. R., Eltis, L. D., Fee, J. A. & Hirst, J. (2003). *Biochemistry*, **42**, 12400–12408.
- Zu, Y., Fee, J. A. & Hirst, J. (2002). *Biochemistry*, **41**, 14054–14065.