

Delft University of Technology

### Structure of the flavocytochrome C sulfide dehydrogenase associated with the copperbinding protein CopC from the haloalkaliphilic sulfuroxidizing bacterium thioalkalivibrio paradoxus ArH 1

Osipov, Eugeny M.; Lilina, Anastasia V.; Tsallagov, Stanislav I.; Safonova, Tatyana N.; Sorokin, Dimitry Y.; Tikhonova, Tamara V.; Popova, Vladimir O.

### DOI

### 10.1107/S2059798318005648

#### **Publication date** 2018

**Document Version** Final published version

#### Published in

Acta Crystallographica Section D: Structural Biology

**Citation (APA)** Osipov, E. M., Lilina, A. V., Tsallagov, S. I., Safonova, T. N., Sorokin, D. Y., Tikhonova, T. V., & Popova, V. O. (2018). Structure of the flavocytochrome C sulfide dehydrogenase associated with the copper-binding protein CopC from the haloalkaliphilic sulfuroxidizing bacterium thioalkalivibrio paradoxus ArH 1. Acta Crystallographica Section D: Structural Biology, 74(7), 632-642. https://doi.org/10.1107/S2059798318005648

#### Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

#### Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

#### Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

### Green Open Access added to TU Delft Institutional Repository

### 'You share, we take care!' - Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.



Received 5 January 2018 Accepted 10 April 2018

Edited by Z. S. Derewenda, University of Virginia, USA

**Keywords:** flavocytochrome *c* sulfide dehydrogenase; protein–protein complex; X-ray structure; copper-binding protein; CopC; *Thioalkalivibrio paradoxus*.

**PDB reference**: complex between flavocytochrome *c* sulfide dehydrogenase and the copper-binding protein CopC from *T. paradoxus*, 5n1t

**Supporting information**: this article has supporting information at journals.iucr.org/d

Structure of the flavocytochrome c sulfide dehydrogenase associated with the copper-binding protein CopC from the haloalkaliphilic sulfuroxidizing bacterium *Thioalkalivibrio paradoxus* ARh 1

# Eugeny M. Osipov,<sup>a</sup>\* Anastasia V. Lilina,<sup>a</sup> Stanislav I. Tsallagov,<sup>a</sup> Tatyana N. Safonova,<sup>a</sup> Dimitry Y. Sorokin,<sup>a,b</sup> Tamara V. Tikhonova<sup>a</sup> and Vladimir O. Popov<sup>a,c</sup>

<sup>a</sup>Research Center of Biotechnology of the Russian Academy of Sciences, 33 Leninsky Avenue, Building 2, Moscow 119071, Russian Federation, <sup>b</sup>Department of Biotechnology, Delft University of Technology, Delft, The Netherlands, and <sup>c</sup>Kurchatov Complex of NBICS-Technologies, National Research Centre 'Kurchatov Institute', Moscow, Russian Federation. \*Correspondence e-mail: e.m.osipov@gmail.com

Flavocytochrome c sulfide dehydrogenase from Thioalkalivibrio paradoxus (TpFCC) is a heterodimeric protein consisting of flavin- and monohaem cbinding subunits. TpFCC was co-purified and co-crystallized with the dimeric copper-binding protein TpCopC. The structure of the TpFCC-(TpCopC)<sub>2</sub> complex was determined by X-ray diffraction at 2.6 Å resolution. The flavinbinding subunit of TpFCC is structurally similar to those determined previously, and the structure of the haem-binding subunit is similar to that of the N-terminal domain of dihaem FCCs. According to classification based on amino-acid sequence, TpCopC belongs to a high-affinity CopC subfamily characterized by the presence of a conserved His1-Xxx-His3 motif at the N-terminus. Apparently, a unique  $\alpha$ -helix which is present in each monomer of TpCopC at the interface with TpFCC plays a key role in complex formation. The structure of the copperbinding site in  $T_P$ CopC is similar to those in other known CopC structures. His3 is not involved in binding to the copper ion and is 6-7 Å away from this ion. Therefore, the His1-Xxx-His3 motif cannot be considered to be a key factor in the high affinity of CopC for copper(II) ions. It is suggested that the TpFCC- $(TpCopC)_2$  heterotetramer may be a component of a large periplasmic complex that is responsible for thiocyanate metabolism.



© 2018 International Union of Crystallography

#### 1. Introduction

Flavocytochrome *c* sulfide dehydrogenase (FCC) is one of the central enzymes of oxidative sulfur metabolism in sulfuroxidizing bacteria (Fukumori & Yamanaka, 1979). This enzyme catalyzes the oxidation of sulfide and polysulfide ions to zero-valent sulfur accompanied by electron transfer to cytochrome *c*. The catalytically active form of the enzyme is a noncovalently bound heterodimer composed of a flavinbinding subunit (~45 kDa) and a monohaem (~10 kDa; Visser *et al.*, 1997; Kostanjevecki *et al.*, 2000) or dihaem (~25 kDa; Fukumori & Yamanaka, 1979; Chen *et al.*, 1994; Hirano *et al.*, 2012) cytochrome subunit. The flavin-binding subunit of FCC contains a covalently bound flavin adenine dinucleotide (FAD) molecule attached by a thioether bond.

The flavin-binding subunit of FCC has a similar fold to that of sulfide:quinone oxidoreductase (SQR; Shahak & Hauska, 2008). Like FCC, the latter enzyme oxidizes sulfide to elemental sulfur and contains an FAD molecule. However, unlike FCC, SQR is a membrane protein that lacks a cytochrome subunit and catalyzes electron transfer from FAD to the quinone pool in the cytoplasmic membrane. In SQR the FAD is bound noncovalently because this enzyme does not have a cysteine residue that is involved in a thioether bond to FAD as in FCC.

The structures of FCCs from two anoxygenic purple sulfur bacteria, Allochromatium vinosum (AvFCC; Chen et al., 1994) and Thermochromatium tepidum (TtFCC; Hirano et al., 2012), and the structures of three SQRs from the hyperthermophilic bacterium Aquifex aeolicus (Marcia et al., 2009), the thermoacidophilic sulfur-oxidizing y-proteobacterium Acidithiobacillus ferrooxidans (Cherney et al., 2010, 2012) and the hyperthermophilic acidophilic sulfur-oxidizing archaeon Acidianus ambivalens (Brito et al., 2009) have been determined by X-ray crystallography. A pair of active-site conserved redox-active cysteine residues in FCC and SQR located on the re side of the isoalloxazine ring of FAD are involved in the oxidation of sulfide ions accompanied by electron transfer to the FAD. The successive oxidation of several sulfide ions leads to elongation of the polysulfide chain between the catalytic cysteines. The following intermediates in this reaction were found by X-ray crystallography: a disulfur molecule S<sub>2</sub> between the catalytic cysteine residues in *Tt*FCC (Hirano et al., 2012) and an S<sub>8</sub> molecule in SQR from A. aeolicus (Marcia et al., 2009). Consecutive nucleophilic attacks of new sulfide molecules finally result in release of the polysulfide molecule for stereochemical reasons and restoration of the disulfide bridge in the active site (Brito et al., 2009). The electrons that are released upon sulfide oxidation are transferred from FAD to quinone in SQR or to haem c in the cytochrome subunit in FCC. In SQR, guinone is located in the hydrophobic pocket at a distance of 3 Å from FAD, resulting in direct electron transfer. In FCC, the distance between FAD and haem c is  $\sim 9$  Å. It is assumed that electron transfer from FAD to the haem is mediated by the tryptophan, threonine and tyrosine residues of the flavoprotein subunit located close to the haem group (Chen et al., 1994).

The Gram-negative sulfur-oxidizing  $\gamma$ -proteobacteria of the genus Thioalkalivibrio are haloalkaliphilic obligate autotrophs that naturally occur in alkaline (soda) lakes in Africa, Central Asia and North America (Sorokin et al., 2002, 2010, 2014; Grant & Sorokin, 2011). The microbial communities in these lakes have adapted to extreme conditions such as alkaline pH (up to pH 11) and high salinity (up to  $4 M \text{ Na}^+$ ). The oxidation of reduced sulfur compounds, such as thiosulfate, sulfide, polysulfide and molecular sulfur, is the main source of energy for the lithoautotrophic microbial communities in soda lakes (Sorokin et al., 2014). Ten out of 85 strains of Thioalkalivibrio isolated from soda lakes can grow with thiocyanate (SCN<sup>-</sup>) as a the sole energy source (Sorokin *et al.*, 2013). The genome of one of these strains, T. paradoxus ARh 1 (Berben et al., 2015), contains five copies of gene clusters that encode proteins homologous to FCC. The amino-acid sequence identity between the flavin-binding subunits encoded by the different copies of the FCC gene is up to 73%. Only one of these five copies corresponds to a dihaem variant,

while the other four are monohaem variants. The goal of this work is to study the structure of the monohaem FCC from *T. paradoxus* ARh 1 (*Tp*FCC; NCBI accession Nos. WP\_006748977 and WP\_006748978 for the cytochrome and flavinbinding subunits, respectively) synthesized during bacterial growth with thiocyanate in the presence of copper ions.

In all steps, FCC was co-purified and then co-crystallized with a protein of the CopC family (copper-resistance protein). CopC proteins can bind copper(I) and copper(II) ions with high affinity  $(10^{-7} \ge K_d \ge 10^{-13} M \text{ and } 10^{-10} \ge K_d \ge 10^{-17} M$ , respectively; Zhang et al., 2006; Wijekoon et al., 2015). The CopC gene is located either in the operon encoding the copper-resistance proteins CopABCDERS or, together with CopD, near the genes for multicopper oxidases (Lawton et al., 2016). It is hypothesized that in the latter case CopC acts as a copper chaperone facilitating the incorporation of copper ions into the active site of folded multicopper oxidases (Lawton et al., 2016). Based on an analysis of amino-acid sequences, the proteins of the CopC family have recently been classified into several subfamilies according to their ability to bind copper ions (Lawton et al., 2016): C<sub>1-1</sub> CopCs containing copper(I)and copper(II)-binding sites, C1-0 CopCs containing one copper(I)-binding site, C<sub>0-1</sub> CopCs containing one copper(II)binding site, C<sub>0-2</sub> CopCs containing one high-affinity copper(II)-binding site (high affinity means that the binding of copper ion is two orders stronger compared with the proteins of the  $C_{0-1}$  subfamily) and  $C_{0-0}$  CopCs, which do not bind copper ions. The structures of CopCs from Methylosinus trichosporium OB3b (C<sub>0-1</sub> CopC; Lawton et al., 2016), Escherichia coli (C1-1 CopC; Wernimont et al., 2003) and Pseudomonas syringae ( $C_{1-1}$  CopC; Zhang et al., 2006) have been determined. In the structures of CopC from M. trichosporium and P. syringae, the copper(II) ion is coordinated by the N-terminal amino group, His1  $N^{\delta},$  and His  $N^{\delta}$  and Asp  $O^{\delta}$ of the conserved His-Xxx-Asp motif. CopCs of the C<sub>0-2</sub> subfamily are additionally characterized by the presence of a His1-Xxx-His3 motif at the N-terminus (Lawton et al., 2016). The high affinity of  $C_{0-2}$  CopCs for copper(II) was confirmed in the study of CopC from Pseudomonas fluorescens SBW25 (Wijekoon *et al.*, 2015). The structure of  $C_{0-2}$  CopCs remains unknown. The higher affinity of C<sub>0-2</sub> CopCs for copper(II) has been attributed to the formation of a His3 N<sup> $\delta$ </sup>-copper(II) coordination bond instead of the Asp  $O^{\delta}$ -copper(II) bond (Wijekoon et al., 2015; Lawton et al., 2016).

An amino-acid sequence alignment and analysis of CopC proteins from different subfamilies showed that the CopC gene from *T. paradoxus* ARh 1, which is located in the operon that also contains the CopD gene, produces a  $C_{0-2}$  CopC. In this study, we determined the structure of FCC in complex with this  $C_{0-2}$  CopC [*Tp*FCC–(*Tp*CopC)<sub>2</sub>].

#### 2. Materials and methods

#### 2.1. Protein isolation and purification

*T. paradoxus* ARh 1 was grown in batch mode on a mineral sodium carbonate–sodium bicarbonate medium containing

0.6 *M* total Na<sup>+</sup> at pH 9.75 as described previously (Sorokin *et al.*, 2001) with sodium thiocyanate (10 m*M*) as an energy and nitrogen source. After the consumption of the thiocyanate, additional 10 m*M* thiocyanate was added. The Cu<sup>2+</sup> concentration in the growth medium was 30  $\mu$ g l<sup>-1</sup>.

After full utilization of the thiocyanate, the cells from a 10 l culture were collected by centrifugation at 6000 rev min<sup>-1</sup> for 20 min at 4°C, resuspended in approximately 100 ml of the medium, pelleted again at 6000 rev min<sup>-1</sup> and finally washed with 20 mM MOPS buffer pH 7.5 containing 0.6 M NaCl. The wet weight of the cells was 21 g.

To obtain the periplasmic fraction, the cells were incubated for 10 min in 42 ml lysis buffer (20 mM MOPS pH 7.5, 1 mg ml<sup>-1</sup> lysozyme, 500 mM sucrose) at room temperature. Milli-Q water (63 ml) and 1 mM PMSF were then added and incubation was continued at 4°C for 15 min. Spheroplasts and cell debris were separated from the periplasmic fraction by centrifugation at 8000 rev min<sup>-1</sup> for 30 min.

The periplasmic fraction, as a source of TpFCC, was immediately applied onto a DEAE-Sepharose FF column equilibrated with 50 mM MOPS buffer pH 7.5 and eluted with a gradient of 15–70% 1 M NaCl in 50 mM MOPS buffer pH 7.5. Fractions containing TpFCC were combined, diluted and applied onto a Mono Q 10/100 anion-exchange column (GE Healthcare, USA) equilibrated with 50 mM Tris buffer pH 7.5. The proteins were eluted with a gradient of 0–100% 1 M NaCl in 50 mM Tris buffer pH 7.5.

Fractions containing TpFCC were collected, pooled and applied onto a Superdex 200 10/300 gel-filtration column (GE Healthcare, USA) equilibrated with 50 mM Tris buffer pH 7.5, 200 mM NaCl. The buffer for isolation and subsequent crystallization was chosen based on the results of thermofluorescence analysis (see below). An ÄKTA FPLC system (GE Healthcare, USA) was used for the runs.

The molecular weight of TpFCC in solution was determined by size-exclusion chromatography on a Superdex 75 10/300 column equilibrated with 50 mM Tris buffer pH 7.5, 200 mMNaCl. The column was calibrated using molecular-mass standards for size-exclusion chromatography (GE Healthcare, USA).

Finally, the protein was concentrated to 40 mg ml<sup>-1</sup> and used for crystallization. The protein concentration was determined by the Bradford assay (Bradford, 1976). The purity of the fractions was analyzed by SDS–PAGE according to Laemmli (1970). Identification of proteins was carried out by MALDI-TOF MS using an Ultraflex III MALDI-TOF/TOF mass spectrometer (Bruker, Germany).

#### 2.2. Enzyme-activity assay

The enzyme activity of TpFCC was measured by spectrophotometry in a temperature-controlled cell at 25°C using a Cary 100 Bio spectrometer (Thermo Scientific Varian, USA). The reaction mixture (1 ml) was composed of 50 mM glycine buffer pH 9.5, 10  $\mu$ M sodium sulfide and 14 nM enzyme; 20  $\mu$ M horse heart cytochrome c (Sigma, USA) was used as an electron acceptor. The reaction was initiated with sulfide. The

#### Table 1

Data-collection and refinement statistics.

Values in parentheses are for the outer shell.

Data collection	
Space group	$P2_{1}2_{1}2_{1}$
Unit-cell parameters (Å)	a = 39.5, b = 138.4, c = 155.8
Resolution (Å)	29.76-2.60 (2.70-2.60)
$\langle I/\sigma(I) \rangle$	11.4 (2.1)
$CC_{1/2}$ (%)	99.7 (78.5)
Completeness (%)	99.7 (99.8)
Measured reflections	151142 (15674)
Unique reflections	27206 (2841)
Multiplicity	5.5
$R_{\text{meas}}$ † (%)	3.5
Refinement	
$R_{\rm cryst}$ (%)	18.6 (32.0)
$R_{\rm free}$ (%)	25.7 (38.6)
R.m.s.d. from ideal values	
Bond lengths (Å)	0.013
Bond angles (°)	1.9
Average B factors $(Å^2)$	
Protein	56.8
FAD and haem	46.0
Water	47.5
Copper	82.6
No. of Ramachandran plot outliers	2

†  $R_{\text{meas}}$  for the inner shell (29.76–7.86 Å).

reaction rate was calculated from the change in the absorbance of reduced cytochrome c at 550 nm ( $\varepsilon_{550} = 22500 M^{-1} \text{ cm}^{-1}$ ) in the presence and absence of the enzyme.

#### 2.3. Thermofluorescence analysis

The buffer for the final steps of isolation and crystallization was chosen by thermofluorescence analysis (Ericsson *et al.*, 2006) in 96-well PCR plates using a FRET-based MyIQ RT-PCR (Bio-Rad, USA). Each well contained 7.5 µl of 300× SYPRO Orange (Thermo Fisher, USA), 5 µl deionized water, 12.5 µl buffer solution and 12.5 µl protein solution at 1 mg ml<sup>-1</sup>. The negative control contained 12.5 µl water instead of the protein solution. The samples were heated from 20 to 90°C at a rate of 1°C min<sup>-1</sup> and the fluorescence yield was measured. To obtain the temperature midpoint for the protein-unfolding transition ( $T_m$ ), the Boltzmann model was used to fit the fluorescence imaging data (Ericsson *et al.*, 2006). A higher  $T_m$  corresponds to higher protein stability in a given buffer.

#### 2.4. Crystallization

Crystallization conditions were screened with a Rigaku Phoenix + Crystal Trak + Gallery 700 system (Rigaku, Japan) using the commercial crystallization screens Index and Crystal Screen (Hampton Research, USA). The crystallization conditions were optimized using the hanging-drop vapourdiffusion technique at 298 K. The drops were composed of equal volumes (1  $\mu$ l) of protein solution and reservoir solution. Small red needle-like crystals were obtained in a mixture consisting of 0.056 *M* sodium phosphate monobasic monohydrate, 1.344 *M* potassium phosphate dibasic pH 8.2. After manual optimization of the crystallization conditions, needle-like crystals ( $200 \times 30 \times 30 \mu m$ ) were obtained in 0.2 *M* NaH<sub>2</sub>PO<sub>4</sub>, 1.3 *M* K<sub>2</sub>HPO<sub>4</sub> pH 8.2.

#### 2.5. X-ray data collection and structure determination

X-ray diffraction data sets were collected at 100 K on the ID23-1 beamline at the European Synchrotron Radiation Facility, Grenoble, France using a PILATUS 6M detector (Dectris, Switzerland) at a wavelength of 0.97 Å. Prior to data collection, the crystals were soaked in reservoir solution supplemented with 20%(v/v) glycerol and then flash-cooled in liquid nitrogen. The data-collection strategy was optimized using the *BEST* program (Bourenkov & Popov, 2010) as incorporated within the *MXCuBE* interface (Gabadinho *et al.*, 2010). The X-ray data sets were processed and merged with *XDS* (Kabsch, 2010). Data-collection statistics are summarized in Table 1.





(a) SDS-PAGE (12%) of the FCC-containing fractions at different steps in purification. The lanes contain (from left to right) molecular-weight markers (labelled in kDa) and fractions after DEAE-Sepharose, Mono Q and size-exclusion chromatography. The cytochrome c subunit of TpFCCis not shown. (b) Size-exclusion chromatography of  $TpFCC-(TpCopC)_2$ on Superdex 75 10/300. The peak corresponds to a mass of 82 kDa.

The structure of TpFCC was solved by the molecularreplacement method with MOLREP (Vagin & Teplvakov, 2010) using the heterodimeric AvFCC structure (PDB entry 1fcd; Chen et al., 1994) as the starting model. Structure refinement without solvent molecules and CopC converged to an R factor of about 38%, and positive density blobs (>5 $\sigma$ ) were visible in difference electron-density maps. After several cycles of polyalanine chain tracing with Buccaneer from the CCP4 suite (Winn et al., 2011) followed by manual removal of incorrectly placed residues, six  $\beta$ -sheets and two  $\alpha$ -helices were found. Based on the  $\beta$ -sheet structure, we identified the position and orientation of TpCopC in the asymmetric unit.  $\beta$ -Sheets of TpCopC were connected based on the alignment of TpCopC with CopC from P. syringae (PDB entry 2c9p; Zhang et al., 2006). The structure was refined with REFMAC5 (Murshudov et al., 2011). Visual inspection and manual rebuilding of the model were performed with Coot (Emsley et al., 2010). Figures were prepared with PyMOL (Schrödinger). The structure was deposited in the Protein Data Bank as PDB entry 5n1t.

#### 2.6. Bioinformatics tools

The solvent-accessible surface area was calculated in PyMOL using the *get\_area* function. H atoms were added to the structure and the molecular weight was calculated using PyMOL. The theoretical molecular weight was calculated





Overall view of the  $TpFCC-(TpCopC)_2$  heterotetramer. The flavinbinding subunit is shown in grey, the cytochrome subunit in purple and the two CopC monomers in green (chain *M*) and cyan (chain *W*). The copper ion is represented by an orange sphere. The haem *c* and FAD are shown as ball-and-stick models.

from the amino-acid sequence with *ProtParam*. The presence of a signal peptide in the amino-acid sequences was analyzed using the *SignalP* 4.1 server (Nielsen, 2017). Analysis of intermolecular contacts was performed with *PISA* (Krissinel & Henrick, 2007) and *STRING* (http://string-db.org). The sequence alignment and analysis were performed with *BLAST*. The sequence alignment was visualized with *TEXshade/LaTeX*. The surface charge was calculated using the *PyMOL APBS* plugin (Baker *et al.*, 2001). The OMIT map for the CopC dimer was generated in *phenix.polder* (Liebschner *et al.*, 2017)

#### 3. Results and discussion

## 3.1. Isolation and crystallization of the $TpFCC-(TpCopC)_2$ complex

The complex of TpFCC with the copper-binding protein TpCopC was isolated from a periplasmic fraction of T. paradoxus ARh 1 cells grown with thiocyanate as the electron donor and nitrogen source. The localization of the components of the complex in the periplasm is consistent with the data from a bioinformatics analysis, according to which all three subunits of the complex contain a twin-arginine trans-



#### Figure 3

(a) Fragments of the multiple sequence alignment of the flavin subunits of FCC from *T. paradoxus* (*Tp*FCC), *A. vinosum* (*Av*FCC) and *T. tepidum* (*Tt*FCC) and the translated sequences of four other FCC genes from the *T. paradoxus* genome (FCC1–FCC4). Redoxactive cysteines are highlighted in red, the residues involved in hydrogen bonding to FAD in green, the residues involved in the electron-transport chain from FAD to the haem in purple, the cysteine forming a covalent bond with FAD in blue and Trp108 in orange. Gaps are marked by three dots. (b) The loop at the *Tp*FCC–*Tp*CopC interface that adopts different conformations in *Av*FCC, *Tt*FCC and *Tp*FCC.

location motif (Palmer & Berks, 2012) enabling transport of these subunits across the cytoplasmic membrane to the periplasm. Further purification was accomplished using two anionexchange and size-exclusion chromatography steps. According to size-exclusion chromatography, the complex has a molecular weight of 82 kDa. SDS-PAGE analysis (Fig. 1a) and MALDI-TOF mass spectrometry showed that this complex is composed of the two subunits of FCC, flavin-binding (WP 006748978, molecular weight 42 kDa) and cytochrome (WP 006748977, molecular weight 7.8 kDa), and a protein of the CopC family (WP\_006748989, molecular weight 14 kDa; in all cases the calculated molecular weights of mature proteins without a signal peptide are given). Taking into account the molecular weights of the components involved in the complex, the following stoichiometry can be proposed for the complex: 1 TpFCC (heterodimer):2 TpCopC.

A Thermofluor assay was used to optimize the protein buffer prior to crystallization. The Thermofluor assay showed that TpFCC is most stable in 50 mM Tris buffer pH 7.5–8.0. Hence, the final size-exclusion chromatography step was performed in 50 mM Tris buffer pH 8.0 supplemented with 200 mM NaCl.

After size-exclusion chromatography, the sample composed of TpFCC and TpCopC was crystallized. An X-ray diffraction

data set was collected to 2.6 Å resolution (Table 1).

3.2. Overall structure of the *Tp*FCC– (*Tp*CopC)<sub>2</sub> complex

The  $TpFCC-(TpCopC)_2$  structure was solved and refined to 2.6 Å resolution. The crystal contains one heterotetramer per asymmetric unit (Fig. 2). There are two Ramachandran plot outliers: Glu67 and Phe75. Both of these are not true outliers and are attributed to the fact that the refinement was performed without Ramachandran plot restraints, as well as the relatively low resolution. The heterotetramer consists of the flavin-binding (chain A) and cytochrome (chain B) subunits of TpFCC and a homodimer of the TpCopC protein (chains M and W). The linear dimensions of the protein molecule are  $80 \times 100 \times 50$  Å. The flavin-binding subunit contains covalently bound FAD. The single-domain cytochrome subunit contains one covalently bound haem c. One of the TpCopC monomers contains a coordinated copper ion.

#### 3.3. Structure of the flavin subunit

The TpFCC subunit shares about 50% primary-structure identity with the flavinbinding subunits of AvFCC and TtFCC (Fig. 3) and contains 429 residues, 392 of which were refined. The first 37 residues are not visible in the electron-density maps. The TpFCC structure superimposes on those of AvFCC and TtFCC with  $C^{\alpha}$  r.m.s.d.s of 0.8 Å (326  $C^{\alpha}$  atoms) and 0.7 Å (319  $C^{\alpha}$  atoms), respectively. The largest differences between TpFCC and two other FCC structures are observed in the interface region between the flavin-binding subunit and TpCopC in TpFCC (Fig. 3b).

The flavin-binding subunit of TpFCC contains six cysteine residues, three of which play a functional role. Thus, Cys73 forms a covalent bond with FAD, while Cys193 and Cys364 form a redox-active pair in the active site (Fig. 4*a*). The functions of the other Cys residues (Cys48, Cys301 and Cys342) are unclear. Cys301 is conserved in all structurally characterized FCCs and is located at a distance of 19 Å from the active site on the surface of the protein molecule, with its side chain buried inside the protein molecule. Cys48 and Cys342 are nonconserved and are only present in the TpFCCstructure. These residues are spatially close to each other and





*(b)* 

Figure 4

are located at the contact interface between two  $\alpha$ -helices, but a disulfide bond is not formed between these residues (the Cys48 S<sup> $\delta$ </sup>-Cys342 S<sup> $\delta$ </sup> distance is 4.2 Å).

The environment of the putative sulfide-binding site on the re side of FAD is similar in all FCC structures. It can be stated that the binding and oxidation of the sulfide ion in all FCCs proceeds by the same mechanism. Calculations with the APBS plugin show that the surface of the flavin subunit in the sulfidebinding region has a positive charge (Fig. 5). The conserved catalytic redox-active residues Cys193 and Cys364 are located directly above the plane of the isoalloxazine ring of FAD (Figs. 4a and 4c). The shortest distance from the plane of the isoalloxazine ring to the nearest  $S^{\gamma}$  atom of Cys193 is 3.5 Å. The distance between the  $S^{\gamma}$  atoms of Cys193 and Cys364 is 5.4 Å, which is substantially longer than the length of a disulfide bond (2.2 Å). The  $F_0 - F_c$  electron-density difference map also showed a large positive peak between these atoms. Therefore, Cys193 was replaced with S-mercaptocysteine (Css; Fig. 4*c*). After refinement, the distance from the  $S^{\delta}$  atom of Css193 (orange sphere in Fig. 4*a*) to Cys364 S<sup> $\gamma$ </sup> was 2.7 Å. Bound sulfur was found in the active sites of the sulfide dehydrogenases TtFCC (S2; Hirano et al., 2012) and SQR from A. aeolicus (S<sub>8</sub>; Marcia et al., 2009). Cys193 was supposed to act as a nucleophilic agent and to be involved in electron transfer to bound FAD (Cherney et al., 2010, 2012). Electron transfer to FAD is accompanied by the oxidation of Cys193 S<sup> $\gamma$ </sup>, resulting in the formation of a disulfide bond with the substrate.

The fold of the part of the flavin-binding subunit that is responsible for the binding of FAD is conserved among all FCC structures. It should be noted that FAD adopts a



(a) Functionally significant cysteine residues: Cys73 covalently bound to FAD and the redox-active pair Cys193–Cys364. Cys193 is linked to the S atom of the substrate (orange sphere). The distances between the atoms are shown by dashed lines labelled in Å. (b) The  $\pi$ -conjugated electron-transport system from FAD to the haem (HEC). The residues involved in the  $\pi$ -conjugated system are shown as ball-and-stick models. (c)  $5\sigma F_o - F_c$  OMIT map (green mesh) for S atoms (yellow spheres) in the active site of TpFCC.





(a) Electrostatic surface potential of the flavin-binding subunit of TpFCC calculated using the  $PyMOL \ APBS$  plugin. The large blue surface patch corresponds to the entrance to the sulfidebinding site. (b) Electrostatic surface potential of the cytochrome subunit of TpFCC in the region where the haem is exposed to the solvent. The haem is shown as a stick model.



Figure 6

Cartoon representation of the cytochrome subunit of TpFCC (rainbowcoloured) aligned with the first domain of the dihaem subunit of TtFCC(grey). The haem groups are shown as ball-and-stick models. Haem iron is shown as an orange sphere.

somewhat different conformation in *Tp*FCC and *Av*FCC. In the *Tp*FCC and *Tt*FCC structures the pentoxy moiety of the FAD molecule adopts the proper configuration (2*R*,3*S*,3*S*), while in the AvFCC structure this moiety has the inverse configuration (2*S*,3*R*,3*R*). The FAD molecule is covalently bound to the protein by the FAD C8a–Cys73 S<sup> $\gamma$ </sup> thioether bond (1.7 Å). Cys73 is conserved in all FCC structures and is replaced by proline in SQR structures, in which the FAD molecule is bound noncovalently. The conformation of FAD is stabilized by eight hydrogen bonds. Most of the residues that are involved in these hydrogen bonds are conserved in all known FCC structures: Gly44, Val109, Pro190, Arg192, Asp321, Gly332 and Phe333 (the numbering corresponds to the *Tp*FCC structure).

Despite the conservation of the polypeptide chain fold around FAD, there are substitutions of some functional resi-

dues in the TpFCC structure. Glu36, which forms a hydrogen bond to the ribose O3 atom of FAD in AvFCC and TtFCC, is replaced by Asn65 in TpFCC. Of the five copies of the gene for FCC in the T. paradoxus genome, only TpFCC as characterized in this study contains asparagine at this position; the other copies of the gene for FCC in the T. paradoxus genome contain a glutamic acid residue (FCC1-FCC4 in Fig. 3). The TpFCC structure contains Pro74 on the si side of the isoalloxazine ring of FAD, which corresponds to Tyr43 in TtFCC and AvFCC. The other copies of the FCC gene in the T. paradoxus genome contain tyrosine at this position. Therefore, the environment of the FAD molecule in the TpFCC structure differs from that in the other structurally characterized FCC, as

well as from the other four copies of FCC from the *T. paradoxus* genome.

In *Tp*FCC FAD is less solvent-accessible compared with AvFCC and *Tt*FCC owing to shielding of the adenine group of FAD by the side group of Trp108. The shortest distance from the indole ring of Trp108 (the C<sup>c3</sup> atom) to the plane of the purine ring of FAD is 3.3 Å, thus preventing a water molecule from approaching the purine ring of FAD. This shielding results in a decrease in the solvent-accessible surface area of FAD by 200 Å<sup>2</sup>.

The aromatic groups of FAD, Phe333, Tyr365 and Trp415 of the flavin-binding subunit and the haem of the cytochrome subunit are coplanar and form an extended  $\pi$ -interaction system (Fig. 4b). This coplanar  $\pi$ - $\pi$  system is conserved and forms a potential electron-transport chain from FAD to the haem of the cytochrome subunit in all FCC structures. In the other copies of the FCC gene Phe333 is replaced by tyrosine or histidine, *i.e.* by a residue containing an aromatic side chain that is capable of being involved in  $\pi$ - $\pi$  interactions.

#### 3.4. Cytochrome c subunit of TpFCC

The cytochrome subunit of FCC (chain *B*) is composed of 75 residues arranged as four  $\alpha$ -helices (Fig. 6). The polypeptide chain fold is similar to those in the N-terminal domains of the two-domain cytochrome subunits of  $A\nu$ FCC and TtFCC. The sequence identity and r.m.s.d. between equivalent  $C^{\alpha}$  atoms are 40% and 0.82 Å (67 C<sup> $\alpha$ </sup> atoms), respectively, for TtFCC, and 40% and 0.98 Å (68 C<sup> $\alpha$ </sup> atoms), respectively, for  $A\nu$ FCC.

The Fe atom of haem *c* is hexacoordinated by His15 N<sup> $\varepsilon$ </sup> and Met53 S<sup> $\delta$ </sup> in axial positions. The position of the imidazole ring of His15 is stabilized by a His15 N<sup> $\delta$ </sup>...Pro26 O hydrogen bond. One propionate group of the haem is exposed to the solvent, while another is covered by the polypeptide chain (the Ser45–Met53 loop).

The most substantial difference between the TpFCC and the TtFCC and AvFCC structures is that TpFCC contains a haem c-binding domain. In the dihaem structures of TtFCCand AvFCC the propionate groups of the two haems are buried at the interdomain interface of the two-domain cytochrome c and are linked to each other by hydrogen bonds. In contrast, one propionate group of haem c in TpFCC is exposed to the solvent. In TpFCC the haem solvent accessibility is ~300 Å<sup>2</sup> higher than those of both haems c in other FCC structures (the total surface area of the haem is 1434 Å<sup>2</sup>).

The surface of the cytochrome subunit is predominantly negatively charged and bears a positive charge in the region where the haem is exposed to the solvent (Fig. 5b), which is the most probable binding site for the electron acceptor.

#### 3.5. Structure of the copper-binding protein TpCopC

Structure refinement of the TpFCC heterodimer converged to R factors of ~38%. The difference electron-density maps showed large peaks (>5 $\sigma$ ) at distances of greater than 10 Å from the nearest polypeptide chain. After fitting of the TpCopC dimer to this density and subsequent refinement, the R and  $R_{\text{free}}$  factors decreased to 18 and 25%, respectively.

The crystal structure contains one TpCopC homodimer per asymmetric unit (Fig. 7*a*). Residues His1–Thr127 (chain *M*) and His1–Gly125 (chain *W*) were located in electron-density maps. The C-termini of these chains are disordered; six (chain *W*) and four (chain *M*) residues were not observed in the electron-density maps. Residues 81–83 in chain *M* and Ser81 in chain *W* were also not visible. These disordered residues are located in the Asp79–Gln83 loop at a distance of 15 Å from the copper-binding site. We attributed the absence of electron density for residues 81–83 in chain *M* to conformational rearrangements upon dissociation of the copper(II) ion from the copper-binding site (see below).

The TpCopC monomer consists of two four-stranded  $\beta$ -sheets and one  $\alpha$ -helix. The  $\beta$ -sheets form a sandwich structure. Leu52–Glu61 form an  $\alpha$ -helix which lies on the external surface of one of the  $\beta$ -sheets and is stabilized by three intramolecular hydrogen bonds. The  $\alpha$ -helix Leu52–Glu61 is a distinguishing feature of TpCopC and is not present in the other structurally characterized CopC (Fig. 7c). The  $\alpha$ -helix of chain W is located at the intermolecular interface between TpCopC and the flavin-binding subunit of TpFCC (chain A). Therefore, the  $\alpha$ -helix apparently plays a key role in formation of the TpFCC–(TpCopC)<sub>2</sub> complex.

*Tp*CopC has a His1-Xxx-His3 motif at the N-terminus and belongs to the  $C_{0-2}$  CopC subfamily (Lawton *et al.*, 2016). This subfamily is characterized by the presence of one high-affinity copper(II)-binding site per subunit. It has been proposed that it is His3 that provides the high binding affinity of  $C_{0-2}$  CopC for copper(II) ions owing to the replacement of asparagine in the coordination sphere of the copper ion (Lawton *et al.*, 2016; Wijekoon *et al.*, 2015).

In TpCopC, chain W binds one copper(II) ion in the active site at the N-terminus (Fig. 7b). This copper ion has a square-planar geometry and is coordinated by the N-terminal amino

group, His1 N<sup> $\delta$ </sup>, Asp110 O<sup> $\delta$ </sup> and His112 N<sup> $\delta$ </sup>. A similar geometry of the coordination environment of the copper(II) ion has been observed in other known CopC structures. His3 is not involved in copper binding and is located at a distance of 6– 7 Å from the copper(II)-binding site. Therefore, the presence of the His1-Xxx-His3 motif is not responsible for the high affinity of C<sub>0-2</sub> CopC for copper(II) ions. The main difference in the copper-binding site in *Tp*CopC is that the amino group of His1 forms an additional His1 N····Glu28 O<sup> $\varepsilon$ </sup> hydrogen bond. In the structure of CopC from *P. syringae* (Zhang *et al.*, 2006) the His1 N–Glu27 O<sup> $\varepsilon$ </sup> distance is 4.2 Å, *i.e.* this hydrogen bond is absent. Apparently, the hydrogen bond between His1 N and Glu28 O<sup> $\varepsilon$ </sup> is responsible for the higher affinity of C<sub>0-2</sub> CopC for copper(II) ions owing to an increase in the nucleophilicity of the amino group of His1.

In chain M the copper-binding site is unoccupied. Metalcontaining proteins are known to undergo photoreduction under X-ray irradiation (De la Mora *et al.*, 2012; Antonyuk & Hough, 2011; Bowman *et al.*, 2016). Therefore, we attributed the absence of copper in chain M to the photoreduction of copper(II) to copper(I) followed by dissociation of copper(I) from the copper-binding site. The absence of a copper ion causes the difference in the structures of chains M and W. The TpCopC monomers superimpose with an r.m.s.d. of 0.4 Å over 334 equivalent C<sup> $\alpha$ </sup> atoms. The most substantial difference in the spatial arrangement of the monomers is observed in the Val109–His112 loop that is involved in the coordination of the copper ion.

The fact that the copper(II) ion is predominantly reduced in chain M of TpCopC may be associated with the crystal packing of TpFCC–(TpCopC)<sub>2</sub>. Chain M of TpCopC forms an intermolecular contact with the cytochrome subunit of the symmetry-related TpFCC–(TpCopC)<sub>2</sub> molecule. His1 of chain M is at a distance of ~14 Å from the haem of this cytochrome. The short distance ensures electron transfer from the closely located haem as a source of electrons to the copper-binding site of chain M as an electron acceptor.

## 3.6. Intersubunit contacts in the *Tp*FCC–(*Tp*CopC)<sub>2</sub> heterotetramer

There are strong contacts between the flavin and cytochrome subunits of the functional TpFCC heterodimer and between the TpCopC monomers. All of the FCCs described in the literature exist as stable heterodimers and all CopCs form stable homodimers.

Complexes of FCC with CopC have not been described previously. The interface between TpFCC and TpCopC is located far (>20 Å) from the residues that are essential for sulfide oxidation. In TpCopC the interface with TpFCC is also far from the copper-binding sites. Therefore, TpCopC is not a component of the sulfide-oxidizing machinery and TpFCC is not directly involved in copper binding.

Contacts between the flavin-binding subunit and individual TpCopC monomers are weak. The interface between the CopC dimer and the flavin subunit is stabilized by seven hydrogen bonds and four salt bridges (Fig. 8). The following

four residues are involved in salt bridges: Arg214 and Arg271 of the flavin-binding subunit of TpFCC and Glu11 of chain M and Glu61 of chain W of the TpCopC dimer. Glu61 belongs to the  $\alpha$ -helix of TpCopC.

In TpCopC, the residues forming salt bridges are located inside the  $\alpha$ -helix. This  $\alpha$ -helix was not observed in the CopC structures described previously and is not directly associated with the copper-binding function. Apparently, this  $\alpha$ -helix is required for interactions with other proteins and can be considered to be a hallmark of the CopC subfamily.

The free energy of dissociation of TpFCC and TpCopC is close to zero, although *PISA* scores this complex as nonstable. This means that the complex is rather unstable and dissociates under standard conditions. The crystallization conditions (1.3 *M* phosphate pH 8.2), which slightly resemble the typical environment in the periplasm of haloalkaliphilic bacteria (2 *M* ionic strength, pH 10), and the high protein concentration used for crystallization (40 mg ml<sup>-1</sup>) probably shift the equilibrium towards formation of the  $TpFCC-(TpCopC)_2$  complex.

3.7. Putative biological function of the *Tp*FCC–(*Tp*CopC)<sub>2</sub> complex

The biological function of the  $T_pFCC-(T_pCopC)_2$  complex is unclear. In the genome, the gene cluster encoding  $T_pFCC$  is present at a distance of ten genes or 14 000 bp from that encoding  $T_pCopC$ . A search for biological partners of  $T_pFCC$ and  $T_pCopC$  using *STRING* did not reveal interactions between them.

We can speculate as to how TpFCC and TpCopC form a complex. It was hypothesized that together the CopC and CopD proteins are responsible for the incorporation of copper into fully folded copper-containing proteins in the cytoplasm



#### Figure 7

(a) Overview of the dimeric TpCopC consisting of chains M (green) and W (cyan). (b) The copper(II)-binding site in chain W. The copper ion (orange sphere) has a square-planar coordination formed by His1 N, His1 N<sup>§</sup>, Asp110 O<sup>§</sup> and His112 N<sup>§</sup> (black dashed lines). The His1 N atom forms a hydrogen bond to Glu28 O<sup>§</sup> (yellow dashed line). The  $2F_o - F_c$  electron-density map is shown as a mesh at  $1\sigma$  (grey) and  $5\sigma$  (purple). (c) Alignment of TpCopC (Tpar) and the CopCs from *E. coli* (PDB entry 1lyq; Ecol; Wernimont *et al.*, 2003), *P. syringae* (PDB entry 2c9q; Psyr; Zhang *et al.*, 2006) and *M. trichosporium* (PDB entry 5icu; Mtri; Lawton *et al.*, 2016). The copper(I)-binding site is shown in orange and the copper(II)-binding site is in blue. Glu28 is coloured red. The residues of TpCopC that interact with TpFCC are shown in dark grey (chain M) and grey (chain W). Secondary-structure elements of TpCopC are depicted above the sequence.



Figure 8

Interface between TpFCC and the TpCopC dimer. Residues involved in the interface formation are coloured cyan (CopC, chain W), green (CopC, chain M) and purple (TpFCC). Residues involved in salt-bridge formation are shown as ball-and-stick models. Salt bridges are depicted as orange dashed lines.

(Lawton *et al.*, 2016). The gene for periplasmic copperdependent thiocyanate dehydrogenase (TcDH; PDB entry 5f75, unpublished work) is located directly at the 5' end of the CopCD operon. Therefore, in the periplasmic space TpCopC may be associated with the thiocyanate-oxidation system as a protein involved in the incorporation of copper ions into TcDH. TpFCC, in turn, may be associated with the metabolism of sulfur produced by the TcDH-catalyzed oxidation of thiocyanate.

Therefore, it can be hypothesized that the  $TpFCC-(TpCopC)_2$  heterotetramer, the structure of which is determined in this study, is part of a large periplasmic copperdependent complex. This complex contains TpFCC, TpCopC and TcDH. The role of the complex is to oxidize thiocyanate and utilize the resulting sulfur.

#### Acknowledgements

MALDI-TOF MS analysis was carried out with the equipment of the Shared-Access Equipment Centre 'Industrial Biotechnology' of the Federal Research Center 'Fundamentals of Biotechnology' at the Russian Academy of Sciences.

#### **Funding information**

This work was supported by the Russian Science Foundation (Project No. 14-24-00172, structure determination and analysis) and the Russian Foundation for Basic Research (Project No. 16-04-01818, protein isolation and crystal-lization). DYS was supported by RFBR grant 16-04-00035.

References

- Antonyuk, S. V. & Hough, M. A. (2011). Biochim. Biophys. Acta, 1814, 778–784.
- Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). Proc. Natl Acad. Sci. USA, 98, 10037–10041.
- Berben, T., Sorokin, D. Y., Ivanova, N., Pati, A., Kyrpides, N., Goodwin, L. A., Woyke, T. & Muyzer, G. (2015). *Stand. Genomic Sci.* 10, 105.
- Bourenkov, G. P. & Popov, A. N. (2010). Acta Cryst. D66, 409-419.
- Bowman, S. E. J., Bridwell Rabb, J. & Drennan, C. L. (2016). Acc. Chem. Res. 49, 695–702.
- Bradford, M. M. (1976). Anal. Biochem. 72, 248-254.
- Brito, J. A., Sousa, F. L., Stelter, M., Bandeiras, T. M., Vonrhein, C., Teixeira, M., Pereira, M. M. & Archer, M. (2009). *Biochemistry*, **48**, 5613–5622.
- Chen, Z., Koh, M., Van Driessche, G., Van Beeumen, J., Bartsch, R., Meyer, T., Cusanovich, M. & Mathews, F. (1994). *Science*, **266**, 430– 432.
- Cherney, M. M., Zhang, Y., James, M. N. G. & Weiner, J. H. (2012). J. Struct. Biol. 178, 319–328.
- Cherney, M. M., Zhang, Y., Solomonson, M., Weiner, J. H. & James, M. N. G. (2010). J. Mol. Biol. 398, 292–305.
- De la Mora, E., Lovett, J. E., Blanford, C. F., Garman, E. F., Valderrama, B. & Rudino-Pinera, E. (2012). *Acta Cryst.* D68, 564–577.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). Acta Cryst. D66, 486–501.
- Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N. & Nordlund, P. (2006). *Anal. Biochem.* **357**, 289–298.
- Fukumori, Y. & Yamanaka, T. (1979). J. Biochem. 85, 1405-1414.
- Gabadinho, J. et al. (2010). J. Synchrotron Rad. 17, 700-707.
- Grant, W. D. & Sorokin, D. Y. (2011). *Extremophiles Handbook*, edited by K. Horikoshi, pp. 27–54. Tokyo: Springer.
- Hirano, Y., Kimura, Y., Suzuki, H., Miki, K. & Wang, Z.-Y. (2012). *Biochemistry*, **51**, 6556–6567.
- Kabsch, W. (2010). Acta Cryst. D66, 125-132.
- Kostanjevecki, V., Brigé, A., Meyer, T. E., Cusanovich, M. A., Guisez, Y. & van Beeumen, J. (2000). *J. Bacteriol.* **182**, 3097–3103.
- Krissinel, E. & Henrick, K. (2007). J. Mol. Biol. 372, 774-797.
- Laemmli, U. K. (1970). Nature (London), 227, 680-685.
- Lawton, T. J., Kenney, G. E., Hurley, J. D. & Rosenzweig, A. C. (2016). Biochemistry, 55, 2278–2290.
- Liebschner, D., Afonine, P. V., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C. & Adams, P. D. (2017). Acta Cryst. D73, 148–157.
- Marcia, M., Ermler, U., Peng, G. & Michel, H. (2009). Proc. Natl Acad. Sci. USA, 106, 9625–9630.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D67, 355–367.
- Nielsen, H. (2017). Methods Mol. Biol. 1611, 59-73.
- Palmer, T. & Berks, B. (2012). Nature Rev. Microbiol. 10, 483-496.
- Shahak, Y. & Hauska, G. (2008). Sulfur Metabolism in Phototrophic Organisms, edited by R. Hell, C. Dahl, D. Knaff & T. Leustek, pp. 319–335. Dordrecht: Springer.
- Sorokin, D. Y., Banciu, H., Robertson, L. A., Kuenen, J. G., Muntyan, M. S. & Muyzer, G. (2013). *The Prokaryotes: Prokaryotic Physiology and Biochemistry*, edited by E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt & F. Thompson, pp. 529–554. Berlin: Springer.
- Sorokin, D. Y., Berben, T., Melton, E. D., Overmars, L., Vavourakis, C. D. & Muyzer, G. (2014). *Extremophiles*, 18, 791–809.
- Sorokin, D. Y., Gorlenko, V. M., Tourova, T. P., Tsapin, A. I., Nealson, K. H. & Kuenen, G. J. (2002). *Int. J. Syst. Evol. Microbiol.* 52, 913– 920.
- Sorokin, D. Y., Rusanov, I. I., Pimenov, N. V., Tourova, T. P., Abbas, B. & Muyzer, G. (2010). FEMS Microbiol. Ecol. **73**, 278–290.

- Sorokin, D. Y., Tourova, T. P., Lysenko, A. M. & Kuenen, J. G. (2001). *Appl. Environ. Microbiol.* 67, 528–538.
- Vagin, A. & Teplyakov, A. (2010). Acta Cryst. D66, 22-25.
- Visser, J. M., de Jong, G., Robertson, L. A. & Kuenen, J. G. (1997). Arch. Microbiol. 167, 295–301.
- Wernimont, A. K., Huffman, D. L., Finney, L. A., Demeler, B.,

O'Halloran, T. V. & Rosenzweig, A. C. (2003). J. Biol. Inorg. Chem. 8, 185–194.

- Wijekoon, C. J. K., Young, T. R., Wedd, A. G. & Xiao, Z. (2015). Inorg. Chem. 54, 2950–2959.
- Winn, M. D. et al. (2011). Acta Cryst. D67, 235-242.
- Zhang, L., Koay, M., Maher, M. J., Xiao, Z. & Wedd, A. G. (2006). J. Am. Chem. Soc. 128, 5834–5850.