

Isabelle Phan,^{a,b*} Sandhya Subramanian,^{a,b} Christian Olsen,^{a,b} Thomas E. Edwards,^{a,c} Wenjin Guo,^{a,b} Yang Zhang,^{a,b} Wesley C. Van Voorhis,^{a,d} Lance J. Stewart^{a,c} and Peter J. Myler^{a,b}

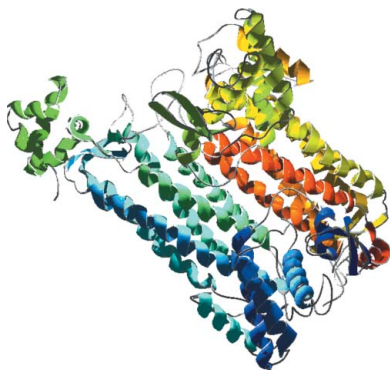
^aSeattle Structural Genomics Center for Infectious Diseases (SSGCID), USA, ^bSeattle Biomedical Research Institute, USA, ^cEmerald BioStructures, USA, and ^dDepartments of Microbiology, Genome Sciences and Immunology, University of Washington, USA

Correspondence e-mail: isabelle.phan@sbsri.org

Received 12 February 2011

Accepted 5 July 2011

PDB Reference: fumarate hydratase, 3gtd.



Structure of fumarate hydratase from *Rickettsia prowazekii*, the agent of typhus and suspected relative of the mitochondria

Rickettsiae are obligate intracellular parasites of eukaryotic cells that are the causative agents responsible for spotted fever and typhus. Their small genome (about 800 protein-coding genes) is highly conserved across species and has been postulated as the ancestor of the mitochondria. No genes that are required for glycolysis are found in the *Rickettsia prowazekii* or mitochondrial genomes, but a complete set of genes encoding components of the tricarboxylic acid cycle and the respiratory-chain complex is found in both. A 2.4 Å resolution crystal structure of *R. prowazekii* fumarate hydratase, an enzyme catalyzing the third step of the tricarboxylic acid cycle pathway that ultimately converts phosphoenolpyruvate into succinyl-CoA, has been solved. A structure alignment with human mitochondrial fumarate hydratase highlights the close similarity between *R. prowazekii* and mitochondrial enzymes.

1. Introduction

Typhus epidemics have been recurrent in human history; the pattern of infection was such that the bacterium *Rickettsia prowazekii*, the agent of typhus, could arguably determine the outcome of war, with outbreaks after World War I resulting in around three million deaths (Raoult *et al.*, 2004). Although hecatombs of this scale remain exceptional, typhus continues to ravage populations in areas of conflict, with mortality rates among infected patients of as high as 20% without antibiotics (Center for Biosecurity of UPMC; <http://upmc-biosecurity.org>). Despite its biological characteristics (environmental stability, small size, aerosol transmission, persistence in infected hosts, low infectious dose, high morbidity and substantial mortality), *R. prowazekii* may not be a primary bioweapon candidate because of its dependency on its eukaryotic host for propagation (Azad, 2007), although this view remains disputed (Walker, 2009). Nonetheless, the Center for Disease Control and Prevention (CDC) ranks *R. prowazekii* as a Category B biological agent and the Department of Health and Human Services (DHHS) classifies it as a top priority for the development of medical countermeasures, thus further encouraging efforts to understand the mechanism of action of this pathogen.

The complete genome of *R. prowazekii* contains only 834 protein-coding genes, a very small number compared with the 5000 genes found in the model bacterium *Escherichia coli*, highlighting similarities between *R. prowazekii* and mitochondrial genes as well as the absence of the genes required for anaerobic glycolysis. It has been suggested that ATP production in *Rickettsia* is the same as that in mitochondria (Andersson *et al.*, 1998). Despite the difference in size between the *Rickettsia* genome (over 1 000 000 bp) and that of human mitochondrial DNA (16 000 bp), the results of phylogenetic studies are consistent with an α -proteobacterial ancestry of the mitochondrial genome (Gray *et al.*, 2001). However, comparisons at the protein level reveal a far more complex picture, since 90% of the mitochondrial proteins are encoded in the nucleus (Gray *et al.*, 2004). One such example is fumarate hydratase, a mitochondrial enzyme from the citric acid cycle, which is encoded on nuclear chromosome 1 in humans (Craig *et al.*, 1976).

The tricarboxylic acid cycle (TCA; also known as the Krebs cycle and the citric acid cycle) is a pathway that Tyler described in 1992 as

'so crucial to the metabolism of living cells that any significant defect is incompatible with life' (Tyler, 1992). The cycle is constituted by a series of biochemical reactions that lead to the progressive oxidative decarboxylation of acetyl-CoA (see Fig. 1). The step that converts fumarate to L-malate has recently been the target of studies of tumorigenesis in humans (King *et al.*, 2006) and pathogenicity in bacteria (van Ooij, 2010).

Two classes of enzymes, class I and class II fumarate hydratase (fumarase; FumC), reversibly convert fumarate to L-malate and have no detectable sequence similarity (Woods *et al.*, 1988). Class I fumarases (FumA and FumB enzymes) are homodimeric, thermolabile, iron-sulfur-containing enzymes of approximately 120 kDa. Class II fumarases (FumC enzymes) are homotetrameric, thermostable, iron-independent enzymes with a molecular mass of 200 kDa. The amino-acid sequences of mitochondrial class II FumCs are highly conserved in eukaryotes and are most closely related to the α -proteobacterial homologues (Schnarrenberger & Martin, 2002).

Defects in human FumC are the cause of fumarase deficiency, a disease characterized by progressive encephalopathy, developmental delay, hypotonia, cerebral atrophy and lactic and pyruvic acidemia (Coughlin *et al.*, 1998). Heterozygous germline mutations of FumC were found in patients with multiple cutaneous and uterine leiomyomas (MCUL). A further set of mutations is the cause of hereditary leiomyomatosis and renal cell cancer (HLRCC). Research to elucidate the mechanisms that lead to enhanced glycolysis in tumours has shown that FumC and succinate dehydrogenase (SDH)

are tumour suppressors, demonstrating for the first time how mitochondrial enzymes and their dysfunction are associated with tumorigenesis (King *et al.*, 2006). A dedicated online database of FumC gene mutations lists all reported FumC sequence variants (Bayley *et al.*, 2008).

Besides its involvement in human tumorigenesis, the TCA cycle has been targeted for its role in pathogenicity. In particular, FumC was found to be one of nine *in vivo*-induced virulence factors in *Listeria* (Wilson *et al.*, 2001) and to bind PdhS, an essential cytoplasmic histidine kinase involved in differentiation, in *Brucella* (Mignolet *et al.*, 2010). A recent paper further shows that the TCA cycle signals the switch between a pathogenic state and a mutualistic state when the *Photothabdus* bacterium changes hosts (Lango & Clarke, 2010).

To this day, the SSGCID project is the sole depositor of *Rickettsia* structures in the Protein Data Bank. Here, we present the high-resolution structure of *R. prowazekii* FumC and compare it with that of its human mitochondrial homolog.

2. Methods

2.1. Protein expression and purification

FumC from *R. prowazekii* strain Madrid E (NCBI NP_221027; *fumC* gene; EC 4.2.1.2; UniProt Q9ZCQ4) spanning the full-length protein from residues 1–461 ('ORF') was cloned into the ligation-

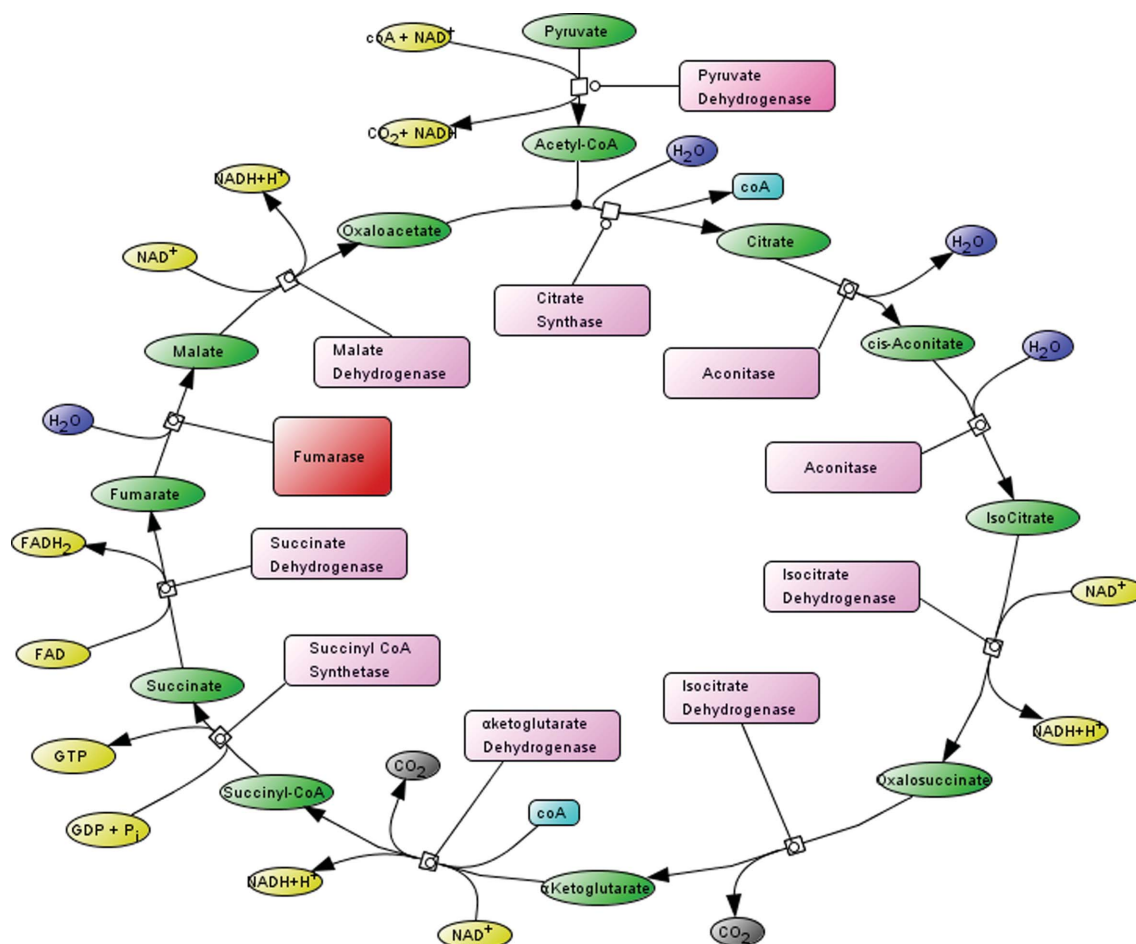


Figure 1 Chemical reaction pathway of the tricarboxylic acid cycle (TCA; also known as the Krebs cycle and the citric acid cycle); catalytic enzymes are indicated in pink boxes, with fumarase, the subject of this study, highlighted in red. This figure was prepared with *CellDesigner* (Funahashi *et al.*, 2003).

independent cloning (LIC; Aslanidis & de Jong, 1990) expression vector pAVA0421 encoding an N-terminal hexahistidine affinity tag followed by the human rhinovirus 3C protease cleavage sequence (MAHHHHHMGTLAQTQPGS-ORF).

The construct encoding the gene for FumC was transformed into chemically competent *E. coli* BL21 (DE3) Rosetta cells. An overnight culture was grown in LB broth at 310 K and was used to inoculate 2 l ZYP-5052 auto-induction medium, which was prepared as described by Studier (2005). FumC was expressed in a LEX bioreactor in the

presence of antibiotics. After 24 h at 298 K, the temperature was reduced to 288 K for a further 60 h. The sample was centrifuged at 4000g for 20 min at 277 K and the cell paste was flash-frozen in liquid nitrogen and stored at 193 K.

During the purification process, the frozen cell pellet was thawed and completely resuspended in lysis buffer (20 mM HEPES pH 7.4, 300 mM NaCl, 5% glycerol, 30 mM imidazole, 0.5% CHAPS, 10 mM MgCl₂, 3 mM β-mercaptoethanol, 1.3 mg ml⁻¹ protease-inhibitor cocktail and 0.05 mg ml⁻¹ lysozyme). The resuspended cell pellet was then disrupted on ice for 15 min with a Branson Digital 450D Sonifier (70% amplitude, with alternating cycles of 5 s pulse-on and 10 s pulse-off). The cell debris was incubated with 20 μl Benzonase nuclease at room temperature for 40 min. The lysate was clarified by centrifugation with a Sorvall RC5 at 10 000 rev min⁻¹ for 60 min at 277 K in a F14S Rotor (Thermo Fisher). The clarified solution was syringe-filtered through a 0.45 μm cellulose acetate filter (Corning Life Sciences, Lowell, Massachusetts, USA). The lysate was purified by IMAC using a HisTrap FF 5 ml column (GE Biosciences, Piscataway, New Jersey, USA) equilibrated with binding buffer (25 mM HEPES pH 7.0, 300 mM NaCl, 5% glycerol, 30 mM imidazole, 1 mM TCEP) and eluted with 500 mM imidazole in the same buffer. The eluted FumC was concentrated and further resolved by size-exclusion chromatography (SEC) using a Superdex 75 26/60 column (GE Biosciences) equilibrated in SEC buffer (20 mM HEPES pH 7.0, 300 mM NaCl, 5% glycerol and 1 mM TCEP) attached to an ÄKTA FPLC system (GE Biosciences). Peak fractions were collected and pooled based on purity-profile assessment by SDS-PAGE. Concentrated pure protein was flash-frozen in liquid nitrogen and stored at 193 K. The final concentration (39.5 mg ml⁻¹) was determined by UV spectrophotometry at 280 nm using a molar extinction coefficient of 33 015 M⁻¹ cm⁻¹ and the final purity (>97%) was assayed by SDS-PAGE.

2.2. Crystallization

Crystallization trials were set up according to a rational crystallization approach (Newman *et al.*, 2005) using the JCSG+ and PACT sparse-matrix screens from Emerald BioSystems and Molecular Dimensions. Protein (39.5 mg ml⁻¹, 0.4 μl) in SEC buffer (20 mM HEPES pH 7.0, 300 mM NaCl, 5% glycerol and 1 mM TCEP) was mixed with an equal volume of precipitant and equilibrated against an 80 μl reservoir in sitting-drop vapor-diffusion format in 96-well Compact Jr plates from Emerald BioSystems at 289 K. Within six weeks, crystals grew in the presence of 2.4 M sodium malonate (JCSG+ condition F9). A gradient optimization screen was designed based on this condition and crystals grew from this screen after about six weeks in 1.4 M sodium malonate pH 6.0.

2.3. Data collection and structure determination

A crystal was harvested, cryoprotected with a solution consisting of the precipitant supplemented with 20% glycerol and vitrified in liquid nitrogen. A 2.4 Å resolution data set was collected at the Advanced Light Source (Andersson *et al.*, 1998) on beamline 5.0.2 (Table 1). The data were reduced with *XDS/XSCALE* (Kabsch, 2010). The structure was determined by molecular replacement using human FumC (PDB entry 3e04; Structural Genomics Consortium, unpublished work) as a search model in *Phaser* (McCoy *et al.*, 2007) from the *CCP4* suite (Winn *et al.*, 2011). The refinement statistics are shown in Table 2. The asymmetric unit contained two protomers of the biologically relevant tetramer, with the other two protomers being generated by crystallographic symmetry. The final model was obtained after numerous iterative rounds of refinement in *REFMAC* (Murshudov *et al.*, 2011)

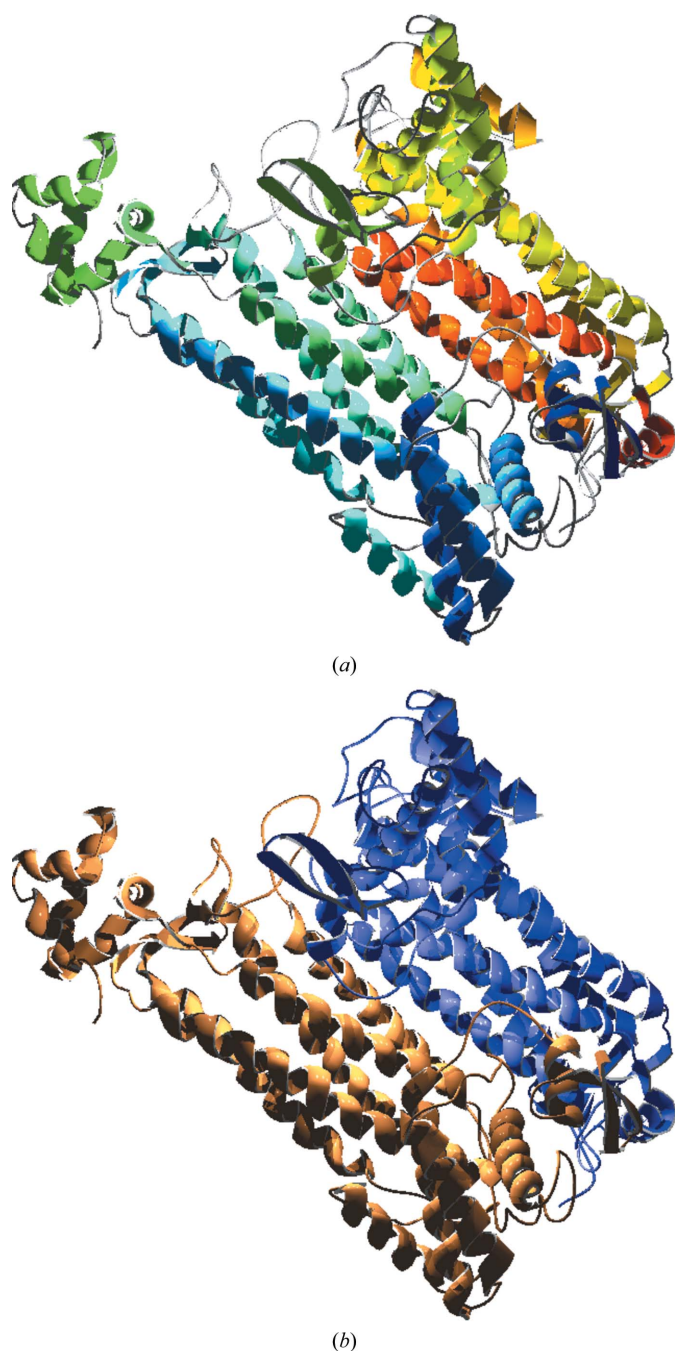


Figure 2
Ribbon diagram of the homodimeric unit structure of *R. prowazekii* FumC showing (a) the overall fold gradient-coloured from red (N-terminus) to blue (C-terminus) and (b) the backbone trace of chain A (blue) and chain B (brown) of the dimeric unit. This figure and all other structure figures in this paper (except for Fig. 5) were prepared using the *POV-Ray* renderer (<http://povray.org>) and *DeepView* (Guex & Peitsch, 1997).

Table 1

Data-collection statistics.

Values in parentheses are for the highest of 20 resolution shells.

Space group	<i>P</i> 3 ₁ 21
Unit-cell parameters (Å)	<i>a</i> = <i>b</i> = 144.9, <i>c</i> = 106.21
Wavelength (Å)	1.0
Resolution range (Å)	19.74–2.4 (2.46–2.40)
No. of unique reflections	46831
Completeness (%)	99.7 (97.5)
<i>R</i> _{merge} †	0.12 (0.73)
Mean <i>I</i> / σ (<i>I</i>)	9.5 (2.7)

$$\dagger \sum_h \sum_l |I_l(h) - \langle I(h) \rangle| / \sum_h \sum_l I_l(h).$$

Table 2

Refinement and model statistics.

Values in parentheses are for the highest of 20 resolution shells.

Resolution range (Å)	19.74–2.4 (2.46–2.40)
<i>R</i> _{cryst} †	0.194
<i>R</i> _{free} ‡	0.226
R.m.s.d. bonds (Å)	0.0080
R.m.s.d. angles (°)	1.043
Protein atoms	6571
Nonprotein atoms	213
Mean <i>B</i> factor (Å ²)	20.275
Residues in favored region (%)	98
Residues in allowed region (%)	100
<i>MolProbity</i> § score [percentile]	1.4 [99th]

† $R_{cryst} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$. The free *R* factor was calculated using 5% of the reflections omitted from the refinement (Winn *et al.*, 2011). ‡ Chen *et al.* (2010).

Table 3

Best pairwise backbone *C*^α r.m.s.d. between the FumC structure from *Rickettsia* and those from human, *E. coli* and *S. cerevisiae* (yeast) calculated using *MultiProt* (Shatsky *et al.*, 2004).

Structure pairs	R.m.s.d. (Å)	Alignment size (residues)
3gtd (<i>Rickettsia</i>)/1kq7 (<i>E. coli</i>)	0.98	416
3gtd (<i>Rickettsia</i>)/1yfm (yeast)	0.98	418
3gtd (<i>Rickettsia</i>)/3e04 (human)	0.99	423

and manual rebuilding in *Coot* (Emsley & Cowtan, 2004). The final model consisted of residues Asn3–Glu457 with no internal gaps for protomer *A*, residues Asn3–Pro316 and Met321–Leu406 for protomer *B*, 198 water molecules, two malonate molecules (one bound to each protomer) and a sodium ion assigned based on the crystallization conditions (sodium malonate), *B* factors and coordination distances of ~2.5 Å (Zheng *et al.*, 2008). The structure was assessed and corrected for geometry and fitness using *MolProbity* (Chen *et al.*, 2010).

3. Discussion

The *R. prowazekii* FumC structure was determined in complex with the product analog malonate. Like the first reported FumC structure from *E. coli* (Weaver *et al.*, 1995), *R. prowazekii* FumC crystallized as a homodimer containing two subunits of the normally tetrameric enzyme (see Fig. 2), in which each chain forms an elongated central four-helix bundle capped by two compact domains at the N- and C-termini. Fig. 3 shows the tetrameric assembly predicted by the *PISA* quaternary-structure tool (Krissinel & Henrick, 2007), including the malonate ligand in the active site.

Structure alignment of the *R. prowazekii* FumC monomer with the human enzyme using *MultiProt* (Shatsky *et al.*, 2004) showed an r.m.s.d. of 0.99 Å over 92% of the sequence. The average *C*^α r.m.s.d. of a global alignment of FumC structures from *Rickettsia* (bound to

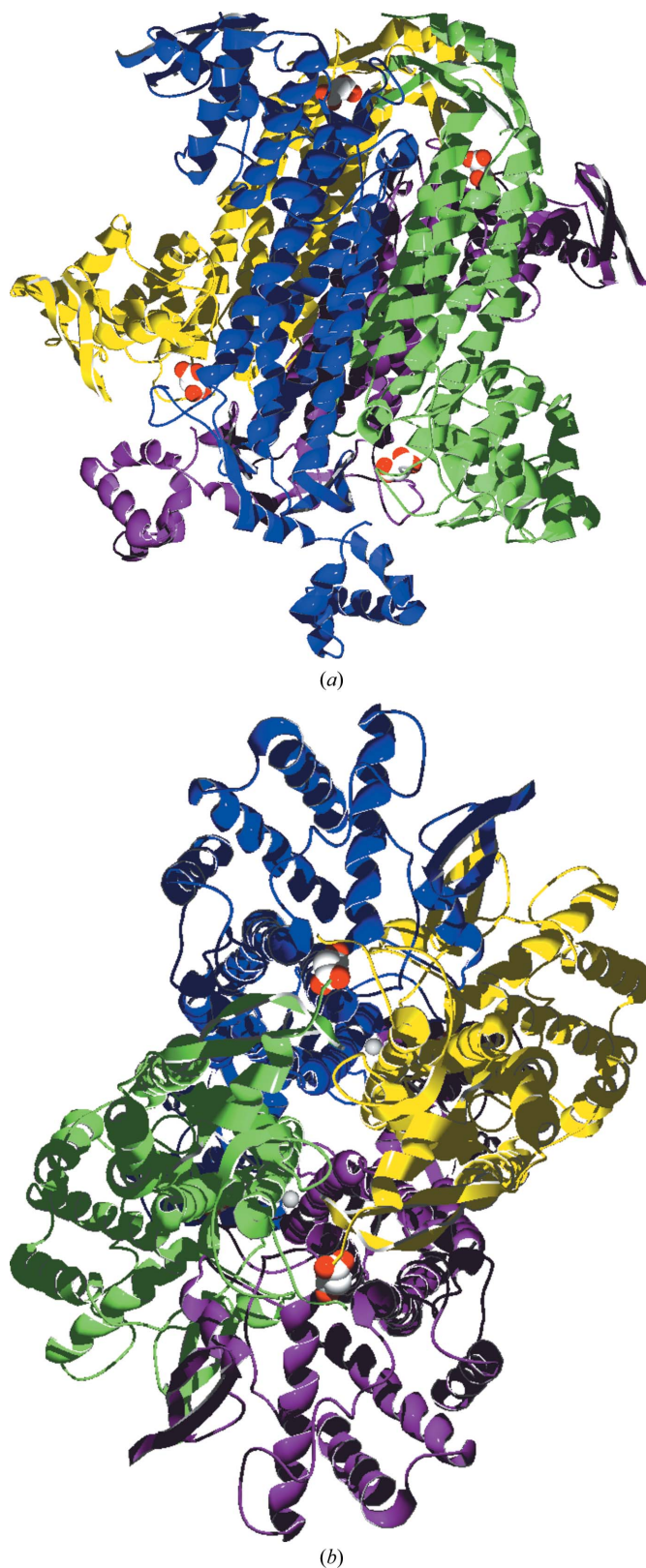


Figure 3

Two views of the tetrameric assembly predicted by *PISA* from the 3gtd coordinates (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html) showing the schematic backbone trace of the four subunits modelled for dimer 1 chain *A* (blue) and *B* (yellow) and dimer 2 chain *A* (magenta) and *B* (green). (a) The side view of each chain bound to the ligand malonate shown in CPK. (b) The two sodium ions at the interface of each dimer can be seen near the central axis of symmetry.

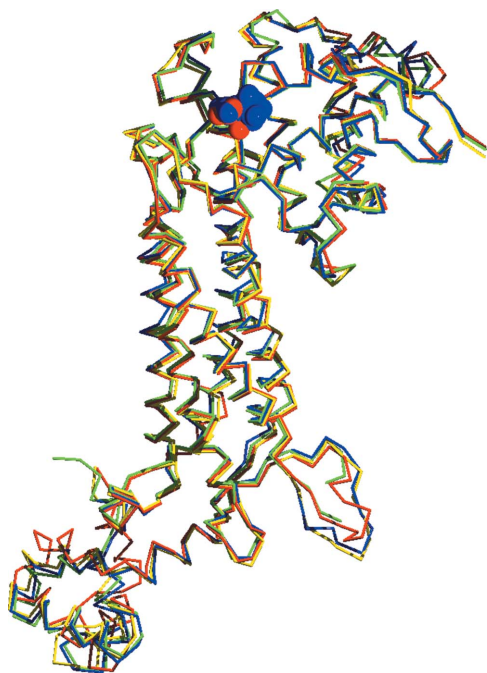


Figure 4

Superposition of the backbone C^α traces of the FumC monomers from *Rickettsia* (PDB entry 3gtd; red), human (PDB entry 3e04; green; Structural Genomics Consortium, unpublished work), *E. coli* (PDB entry 1kq7; blue; Estévez *et al.*, 2002) and *S. cerevisiae* (PDB entry 1yfm; yellow; Weaver *et al.*, 1998) showing the conserved overall fold and the deviations at the C-terminus at the bottom left region of the structure. The ligands for the *Rickettsia* and *E. coli* structures, malonate (red) and citric acid (blue), respectively, are represented in CPK colors.

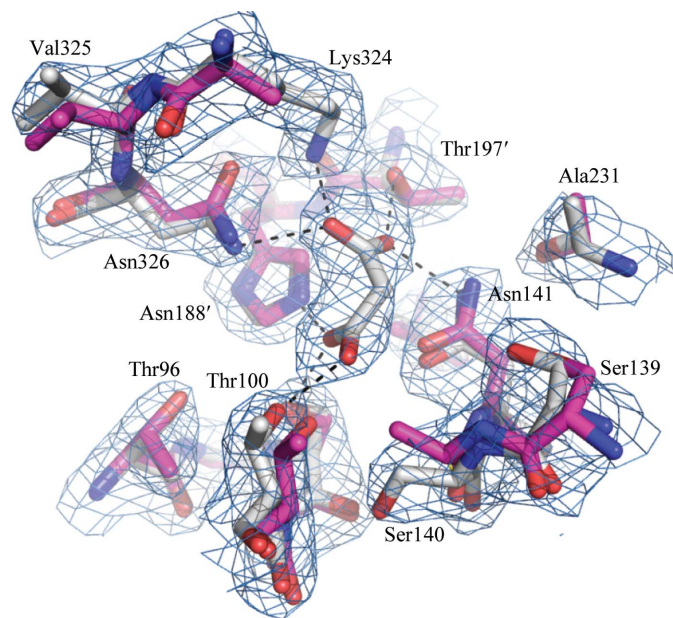


Figure 5

Schematic representation of the ligand environment in the *Rickettsia* FumC monomer complexed with malonate superimposed on the corresponding residues in the human structure (PDB entry 3e04). The backbone (gray for *Rickettsia*, magenta for human) and side chains of residues located within 6 Å of the ligand are shown. Hydrogen bonds of less than 3 Å are shown as dashed lines. Residues are numbered according to the *Rickettsia* FumC numbering system and residues from the second protomer that comprise the active site are identified with the ' notation (Thr187' and His188'). The $2|F_o| - |F_c|$ electron-density map is shown in light blue mesh contoured at 1.0σ . Note that the human structure is apo and thus Lys371 (equivalent to Lys324 in *Rickettsia*) appears disordered. This figure was generated using PyMOL (DeLano, 2002).

the product analog malonate in the active site), human (unbound; PDB entry 3e04; Structural Genomics Consortium, unpublished work), *E. coli* (bound to the competitive inhibitor citrate in the active site and to *S*-malate in the B site; Estévez *et al.*, 2002) and *Saccharomyces cerevisiae* (unbound; Weaver *et al.*, 1998) is 0.90 Å over 410 residues (see Table 3 for pairwise r.m.s.d.s). The largest deviation is found in the C-terminal region; otherwise the backbone structure is remarkably conserved, including the active site (see Fig. 4). The residues located within 6 Å of the ligand in the *Rickettsia* structure, Thr96, Ser98, Ser139, Ser140, Asn141, Ala231 and Leu358, are 100% conserved in the three other species and adopt almost identical conformations, even in the unbound structures: the r.m.s.d. for all atoms over those eight residues is 0.83 Å from the human structure, 1.09 Å from that from *E. coli* and 1.15 Å from that from *S. cerevisiae*. The only visible difference between the human and *Rickettsia* pockets is the tilting of the Ser140 hydroxyl group away from the active site in the human structure (see Fig. 5).

FumC displays some essential features of a good drug target: it is clearly involved in a crucial biological pathway, is functionally well characterized and possesses a druggable binding site. However, the structural evidence obtained in the present study strongly indicates that this enzyme is an unsuitable target for therapeutic intervention against *Rickettsia* owing to the very high degree of conservation between the human and *R. prowazekii* structures in terms of both the global fold and the binding site.

The authors thank the whole SSGCID team. This research was funded under Federal Contract No. HHSN272200700057C from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services.

References

- Andersson, S. G., Zomorodipour, A., Andersson, J. O., Sicheritz-Pontén, T., Alsmark, U. C., Podowski, R. M., Näslund, A. K., Eriksson, A. S., Winkler, H. H. & Kurland, C. G. (1998). *Nature (London)*, **396**, 133–140.
- Aslanidis, C. & de Jong, P. J. (1990). *Nucleic Acids Res.* **18**, 6069–6074.
- Azad, A. F. (2007). *Clin. Infect. Dis.* **45**, S52–S55.
- Bayley, J. P., Launonen, V. & Tomlinson, I. P. (2008). *BMC Med. Genet.* **9**, 20.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst. D* **66**, 12–21.
- Coughlin, E. M., Christensen, E., Kunz, P. L., Krishnamoorthy, K. S., Walker, V., Dennis, N. R., Chalmers, R. A., Elpeleg, O. N., Whelan, D., Pollitt, R. J., Ramesh, V., Mandell, R. & Shih, V. E. (1998). *Mol. Genet. Metab.* **63**, 254–262.
- Craig, I., Tolley, E. & Bobrow, M. (1976). *Cytogenet. Cell Genet.* **16**, 118–121.
- DeLano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst. D* **60**, 2126–2132.
- Estévez, M., Skarda, J., Spencer, J., Banaszak, L. & Weaver, T. M. (2002). *Protein Sci.* **11**, 1552–1557.
- Funahashi, A., Tanimura, N., Morohashi, M. & Kitano, H. (2003). *BIOSILICO*, **1**, 159–162.
- Gray, M. W., Burger, G. & Lang, B. F. (2001). *Genome Biol.* **2**, REVIEWS1018.
- Gray, M. W., Lang, B. F. & Burger, G. (2004). *Annu. Rev. Genet.* **38**, 477–524.
- Guex, N. & Peitsch, M. C. (1997). *Electrophoresis*, **18**, 2714–2723.
- Kabsch, W. (2010). *Acta Cryst. D* **66**, 125–132.
- King, A., Selak, M. A. & Gottlieb, E. (2006). *Oncogene*, **25**, 4675–4682.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Lango, L. & Clarke, D. J. (2010). *Mol. Microbiol.* **77**, 1394–1405.
- 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.
- Mignolet, J., Van der Henst, C., Nicolas, C., Deghelt, M., Dotreppe, D., Letesson, J. J. & De Bolle, X. (2010). *J. Bacteriol.* **192**, 3235–3239.
- 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. D* **67**, 355–367.

- Newman, J., Egan, D., Walter, T. S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J. L., Stuart, D. I. & Perrakis, A. (2005). *Acta Cryst.* **D61**, 1426–1431.
- Ooij, C. van (2010). *Nature Rev. Microbiol.* **8**, 614.
- Raoult, D., Woodward, T. & Dumler, J. S. (2004). *Infect. Dis. Clin. North Am.* **18**, 127–140.
- Schnarrenberger, C. & Martin, W. (2002). *Eur. J. Biochem.* **269**, 868–883.
- Shatsky, M., Nussinov, R. & Wolfson, H. J. (2004). *Proteins*, **56**, 143–156.
- Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
- Tyler, D. (1992). *The Mitochondrion in Health and Disease*. New York: VCH.
- Walker, D. H. (2009). *Vaccine*, **27**, D52–D55.
- Weaver, T., Lees, M., Zaitsev, V., Zaitseva, I., Duke, E., Lindley, P., McSweeney, S., Svensson, A., Keruchenko, J., Keruchenko, I., Gladilin, K. & Banaszak, L. (1998). *J. Mol. Biol.* **280**, 431–442.
- Weaver, T. M., Levitt, D. G., Donnelly, M. I., Stevens, P. P. & Banaszak, L. J. (1995). *Nature Struct. Biol.* **2**, 654–662.
- Wilson, R. L., Tvinnereim, A. R., Jones, B. D. & Harty, J. T. (2001). *Infect. Immun.* **69**, 5016–5024.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Woods, S. A., Schwartzbach, S. D. & Guest, J. R. (1988). *Biochim. Biophys. Acta*, **954**, 14–26.
- Zheng, H., Chruszcz, M., Lasota, P., Lebioda, L. & Minor, W. (2008). *J. Inorg. Biochem.* **102**, 1765–1776.