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

Structure of *Escherichia coli* RutC, a member of the YjgF family and putative aminoacrylate peracid reductase of the *rut* operon

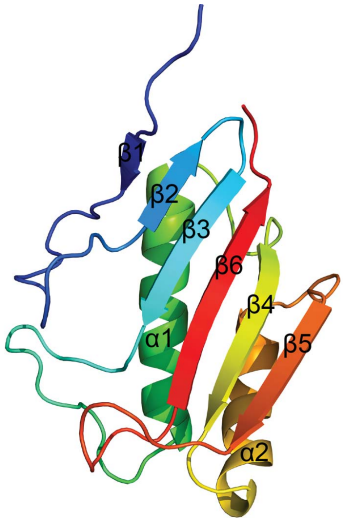
Aleksandra Alicja Knapik,[a,b,c] Janusz Jurand Petkowski,[a] Zbyszek Otwinowski,[d] Marcin Tadeusz Cymborowski,[a,b] David Robert Cooper,[a,b] Maksymilian Chruszcz,[a,b] Wanda Małgorzata Krajewska[c] and Wladek Minor[a,b]*

[a]Department of Molecular Physiology and Biological Physics, University of Virginia, 1340 Jefferson Park Avenue, Jordan Hall, Charlottesville, VA 22908, USA, [b]New York Structural Genomics Research Consortium, USA, [c]Department of Cytobiochemistry, University of Lodz, Pomorska 141/143, 90-236 Lodz, Poland, and [d]Department of Biochemistry, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA

Correspondence e-mail: wladek@iwonka.med.virginia.edu

Received 7 August 2012
Accepted 5 October 2012

PDB Reference: RutC, 3v4d



RutC is the third enzyme in the *Escherichia coli rut* pathway of uracil degradation. RutC belongs to the highly conserved YjgF family of proteins. The structure of the RutC protein was determined and refined to 1.95 Å resolution. The crystal belonged to space group $P2_12_12$ and contained six molecules in the asymmetric unit. The structure was solved by SAD phasing and was refined to an R_{work} of 19.3% (R_{free} = 21.7%). The final model revealed that this protein has a *Bacillus* chorismate mutase-like fold and forms a homotrimer with a hydrophobic cavity in the center of the structure and ligand-binding clefts between two subunits. A likely function for RutC is the reduction of peroxyaminoacrylate to aminoacrylate as a part of a detoxification process.

1. Introduction

The *rut* pathway, a novel pathway among α- and γ-proteobacteria, degrades pyrimidines to produce nitrogen at suboptimal temperatures (Kim *et al.*, 2010). The *rut* operon in *Escherichia coli* is composed of six enzymes that are responsible for reducing/degrading the uracil ring to 3-hydroxypropionic acid (RutA–F), the uracil transporter RutG and the TetR-family repressor RutR (Shimada *et al.*, 2007).

RutC is a member of the highly conserved YjgF/YER057c/UK114 family of proteins. This family (henceforth abbreviated in this paper as the YjgF family) is known to be widely distributed among eubacteria, archaea and eukaryotes. Although various studies have suggested that YjgF-family members may be involved in removing potentially toxic secondary products that result from normal metabolism (Volz, 1999; Parsons *et al.*, 2002), only YjgF has had its cellular role revealed. YjgF is required in *Salmonella enterica* for isoleucine biosynthesis when grown on pyruvate medium (Lambrecht *et al.*, 2010). It has recently been shown that it catalyzes the hydrolysis of enamines and imines, possibly as a means of liberating ammonia, although the extrapolation of this to RutC requires further investigation (Lambrecht *et al.*, 2012).

Several other family members have been shown to be physiologically important. TdcF from *E. coli* is involved in 2-ketobutyrate metabolism (Burman *et al.*, 2007), and YabJ is involved in the regulation of purine biosynthesis (Rappu *et al.*, 1999; Sinha *et al.*, 1999) and isoleucine biosynthesis (Kim *et al.*, 2001). YabJ has ribonuclease activity (Morishita *et al.*, 1999) or translation-inhibition activity (Oka *et al.*, 1995).

Interestingly, *rutC* has been identified as a gene that is induced during infection by avian pathogenic *E. coli* (APEC), which has a significant economic impact in the poultry industry throughout the world (Tuntufye *et al.*, 2012).

YjgF-family proteins are ∼14 kDa proteins that form homotrimeric assemblies. The homotrimer has a *Bacillus* chorismate mutase fold. Despite high fold similarity to other chorismate mutase fold-containing proteins, there is no obvious sequence similarity or functional connection among those proteins. Most proteins in the YjgF family have a characteristic globular shape with a large hydrophobic cavity of unknown function in the middle of the trimeric quaternary structure. Additionally, three clefts are formed at the

interfaces of adjacent monomers, which have been identified as putative ligand-binding sites (Sinha *et al.*, 1999).

Of 36 known YjgF structures, only one has a ligand bound which could be of biological significance. This is TdcF, an *E. coli* protein (Burman *et al.*, 2007; PDB entry 2uyn), in complex with 2-ketobutyrate. TdcF belongs to the *tdc* operon responsible for anaerobic degradation of L-threonine and L-serine, in which L-threonine is deaminated to 2-ketobutyrate (Burman *et al.*, 2007). The ketobutyrate molecule is bound in the putative active site formed by two adjacent monomers.

RutC is part of the uracil-degrading *rut* operon. *In vitro*, only three enzymes (RutA, RutB and RutF) of the *rut* pathway are needed to release both N atoms from the uracil ring in the form of ammonium ions (Mukherjee *et al.*, 2010). However, the *in vivo* production of toxic intermediates requires a detoxification mechanism in the living cell. This function has been proposed to be provided by both the RutC and RutD proteins by Kim *et al.* (2010). RutC could reduce peroxy-aminoacrylate to aminoacrylate, inhibiting the spontaneous hydrolysis of the amino group, so that in the next step RutD can increase the rate of hydrolysis of aminoacrylate (Lambrecht *et al.*, 2010; Knapik *et al.*, 2012). A schematic of the reaction catalyzed by RutC is shown in Fig. 1(*c*).

2. Materials and methods

2.1. Cloning, expression and purification

The *rutC* gene was extracted and amplified by PCR on *E. coli* genomic DNA using the following primers: 5′-GAATTCCATATG-CCAAAATCCGTAATTATTCCC-3′ (forward) and 5′-ATCTCGA-GTTACTTGGCGATATGCGCATTTG-3′ (reverse). Each primer confers a restriction-enzyme site at either end of the gene; these sites are *Nde*I and *Xho*I, respectively. The amplified gene was cloned into a modified pET15b vector where a linker was cloned between the *Nde*I and *Nco*I restriction sites and replaced the thrombin-cleavage site with a tobacco etch virus (TEV) protease site. This linker was made with the following primers: 5′-CATGGGCTCTTCTCATCATCATC-ATCATCATTCTTCTGGTCGTGAGAACTTGTACTTTCAAGG-CCA-3′ (forward) and 5′-TATGGCCTTGAAAGTACAAGTTCT-CACGACCAGAAGAATGATGATGATGATGATGAGAAGAG-CC-3′ (reverse). The clones were verified by sequencing and transformed into *E. coli* strain BL21 (DE3) CodonPlus RIL (Stratagene). The cells were grown in M9 minimal medium with selenomethionine (SeMet) at 310 K until the OD_{600} reached 1.2, followed by the induction of overexpression of recombinant RutC with 1 m*M* IPTG (isopropyl *β*-D-1-thiogalactopyranoside) in the presence of a cocktail of inhibitory amino acids and SeMet according to the manufacturer's specifications (Shanghai Medicilon Inc.). Induction was performed at 289 K overnight. Cell pellets were resuspended in lysis buffer (50 m*M* HEPES pH 7.5, 500 m*M* NaCl, 5% glycerol) with protease-inhibitor cocktail tablets (Roche), lysed by sonication and spun down in an ultracentrifuge. The supernatant was loaded onto an Ni–NTA affinity column (Qiagen). The column was washed with wash buffer (50 m*M* HEPES pH 7.5, 500 m*M* NaCl, 5% glycerol, 30 m*M* imidazole). RutC protein was eluted with elution buffer (50 m*M* HEPES pH 7.5, 500 m*M* NaCl, 5% glycerol, 250 m*M* imidazole). The eluted protein was incubated with polyhistidine-tagged recombinant TEV protease to remove the N-terminal affinity tag while dialyzing against 50 m*M* HEPES pH 7.5, 500 m*M* NaCl, 5% glycerol buffer at 277 K overnight. To remove cleaved polyhistidine tag and TEV protease, the sample was passed through a second Ni–NTA column. The collected flowthrough was concentrated and further purified using a size-exclusion column in 10 m*M* HEPES pH 7.5, 500 m*M* NaCl. The protein was concentrated to a final concentration of 10 mg ml^{-1}.

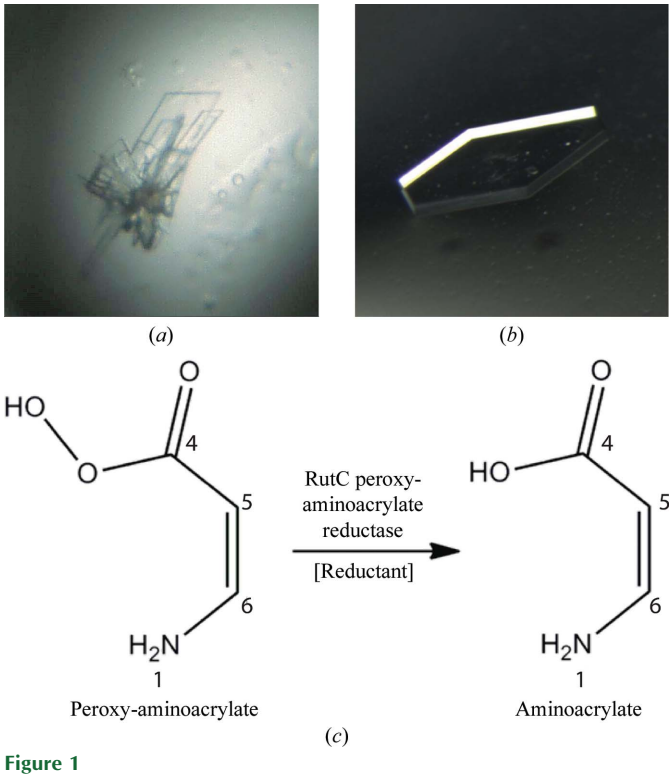


Figure 1
(*a*) Initial and (*b*) optimized crystals of RutC protein. (*c*) Schematic of the RutC-catalyzed reaction. The reductant that participates in the conversion of peroxyaminoacrylate to aminoacrylate is unknown.

2.2. Crystallization

RutC protein was crystallized by vapor diffusion in hanging drops at 293 K. Initial crystals were obtained using The PACT Suite (Qiagen) in two conditions: (i) 0.1 *M* MMT pH 4.0, 30% PEG 1500 and (ii) 0.1 *M* SPG pH 5.0, 25% PEG 1500. Both conditions gave rise to thin clustered plates that diffracted to 2.8 Å resolution (Fig. 1*a*). Crystals were optimized by the addition of a cocktail of xylitol, potassium/sodium tartrate, 2,6-pyridinecarboxylic acid, 2,4-pyridinecarboxylic acid and 3-sulfobenzoic acid. This resulted in single 'coffin-shaped' crystals (elongated hexagonal prisms) that diffracted to 1.95 Å resolution (Fig. 1*b*) and led to the final structure of the RutC protein. For data collection, crystals were flash-cooled in liquid nitrogen using a 1:1 mixture of mother liquor and 50% glycerol for cryoprotection.

2.3. Data collection and processing, structure solution and refinement

X-ray diffraction data were collected from RutC protein crystals cooled to 100 K on the 21ID-G beamline of the Advanced Photon Source (APS) at Argonne National Laboratory. The data were collected at the selenium absorbance peak for phasing. Data collection and reduction were performed with the *HKL*-2000 package (Otwinowski & Minor, 1997). Solution of the structure by single-wavelength anomalous diffraction (SAD) phasing and building of the initial models were performed using *HKL*-3000 (Minor *et al.*, 2006), which is integrated with *SHELXC*/*D*/*E* (Sheldrick, 2008), *MLPHARE* (Otwinowski, 1991), *DM* (Cowtan & Zhang, 1999),

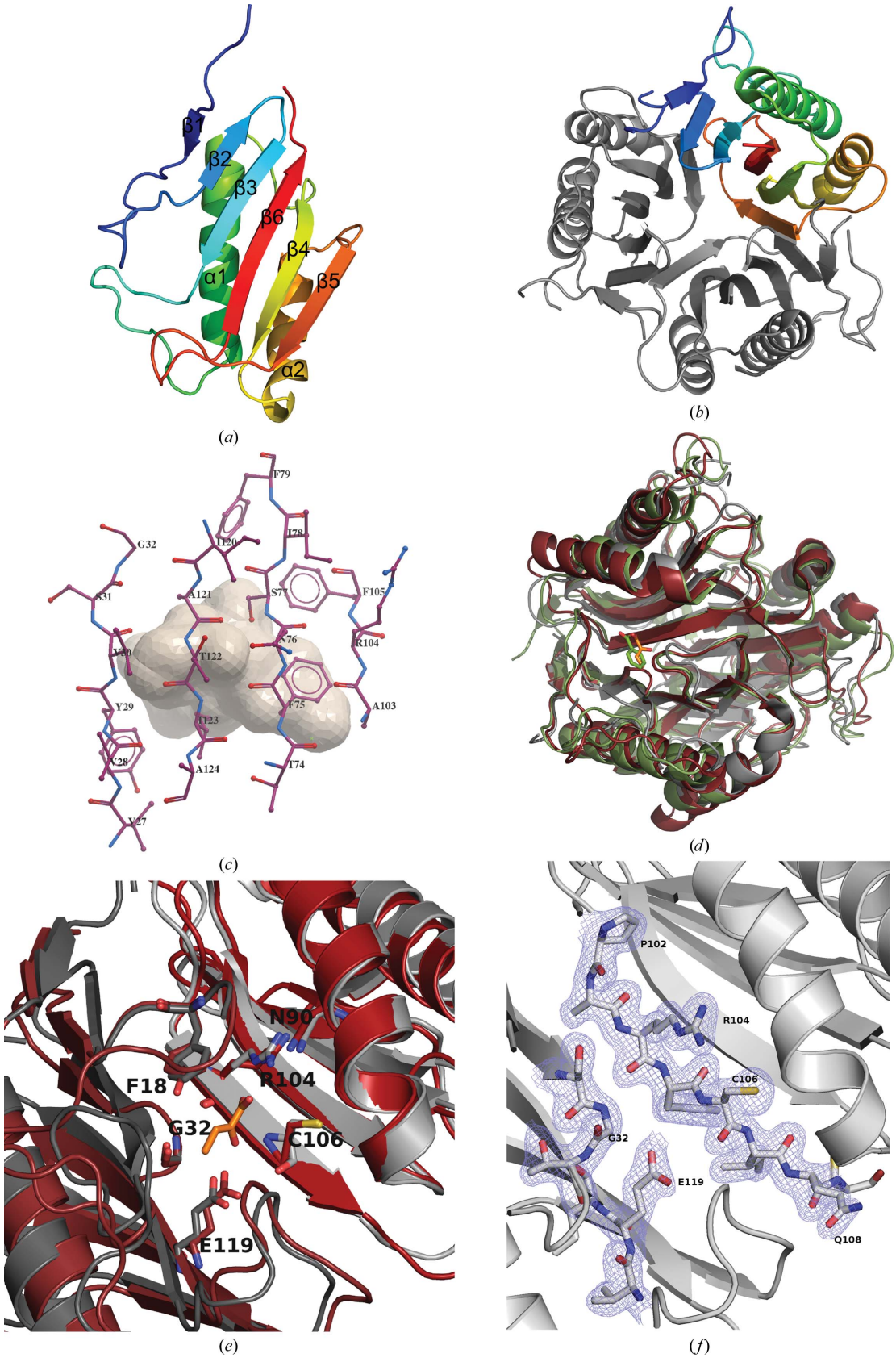


Figure 2
Cartoon representation of (*a*) a monomer and (*b*) the trimeric form of the RutC protein. (*c*) A model of the tubular cavity and the residues composing its surface (represented as sticks), as calculated by *ICM-PRO*. (*d*) Superposition of RutC (gray) and its two structural homologs TdcF (burgundy) and hp14.5 (green). Two ligands are visible in the putative active-site cavity of TdcF (2-ketobutyrate; orange) and hp14.5 (benzoate; lime green). (*e*) Close-up of the active sites of RutC and TdcF. Only RutC residues are labelled. (*f*) A $2F_o - F_c$ electron-density map of the RutC active site. Some residues that would obscure the view have been omitted.

Table 1
Data-collection and refinement statistics for RutC from *E. coli*.

Values in parentheses are for the highest resolution shell.

Data collection	
Wavelength (Å)	0.9787
Unit-cell parameters (Å, °)	a = 92.3, b = 179.2, c = 45.2, $\alpha = \beta = \gamma$ = 90.00
Space group	$P2_12_12$
Resolution range (Å)	50–1.95 (1.98–1.95)
No. of observations	534145
No. of unique reflections	55640
Data completeness (%)	99.9 (100)
R_{merge}	0.094 (0.633)
$\langle I/\sigma(I)\rangle$	34.3 (3.0)
Refinement	
No. of protein atoms	5666
Mean B value for protein atoms (Å^2)	43.3
No. of water atoms	342
Mean B value for water atoms (Å^2)	45.0
R_{work} (%)	19.5
R_{free} (%)	21.7
R.m.s.d. bond lengths (Å)	0.018
R.m.s.d. bond angles (°)	1.7
Geometry	
Ramachandran outliers (%)	0
Ramachandran favored (%)	98.8
Residues with bad bonds (%)	0
Residues with bad angles (%)	0.13
Clashscore	5.65
Clashscore percentile	96

ARP/*wARP* (Perrakis *et al.*, 1999), *CCP*4 (Winn *et al.*, 2011), *SOLVE* (Terwilliger & Berendzen, 1999) and *RESOLVE* (Terwilliger, 2002). Model refinement was performed with *REFMAC*5 (Murshudov *et al.*, 1997, 2011; Winn *et al.*, 2011) and *Coot* (Emsley & Cowtan, 2004). *MolProbity* (Chen *et al.*, 2010) and *ADIT* (Yang *et al.*, 2004) were used to validate the structure. Data-collection and refinement statistics are presented in Table 1. The model was deposited in the PDB with accession code 3v4d.

More detailed structure analyses (calculation of the cavity volumes and superposition) were performed using the *Internal Coordinate Mechanics Professional* (*ICM-PRO*) software from Molsoft LLC (Abagyan & Totrov, 1997). The figures were created using *PyMOL* (The PyMOL Molecular Graphics System v.1.2r3pre, Schrödinger LLC) and *ICM-PRO*. R.m.s.d.s and percentage identities were calculated using the *DALI* sever (Holm & Rosenström, 2010). The sequence alignment was rendered using *PySHADE* (Porebski *et al.*, manuscript in preparation).

3. Results and discussion

3.1. Overall structure and multimeric assembly of the RutC protein

The structure of *E. coli* RutC was determined by the SAD technique and refined to 1.95 Å resolution. A RutC monomer is a single-domain, 13.7 kDa, 128-residue-long polypeptide chain with a *Bacillus* chorismate-mutase-like fold. The monomer consists of two α-helices and six β-strands. The β-strands form a flat sheet, with strands β1 and β5 twisted. The helices are packed tightly adjacent to the sheet, parallel to the plane of the sheet and to each other (Fig. 2*a*).

Gel-filtration chromatography shows that the RutC protein forms a trimeric assembly in solution. In the crystal lattice, there are two trimers in the asymmetric unit. The RutC homotrimer forms a globular structure with a tubular inner cavity (Fig. 2*b*). Six helices (two from each monomer) form the outer rim of the trimer. The barrel-like core of the structure is formed by 12 β-strands (four from each monomer). One β-strand from each monomer (β5) is twisted with respect to the others and this twist causes these strands from each monomer to close off one end of the large (400 Å^3 in volume) tubular cavity inside the trimer structure. The surface of the cavity (Fig. 2*c*) consists mostly of hydrophobic residues, some of which are well conserved among the YjgF family of proteins. The nonconserved amino acids involved in the formation of the cavity are similar in hydrophobicity to the consensus residues present in other RutC orthologs (Burman *et al.*, 2007; Volz, 1999). Some of the residues involved in cavity formation are also responsible for formation of the quaternary structure. They are well conserved and form hydrophobic (Leu15, Pro17, Tyr86, Tyr93, Phe97 and Pro102) and polar patches organized in two rings. One is formed by nine phenylalanines, three from each chain (Phe75, Phe79 and Phe105); the second, which is more polar, is made up of Tyr29, His125 and Val27 from all chains.

Table 2
Residue compositions of the putative active sites of RutC and three analyzed homologs.

Residues which are 100% conserved are shown in bold.

RutC	TdcF	hp14.5	YabJ
Phe18	Tyr17	Tyr21	Tyr17
Gly32	**Gly31**	**Gly35**	**Gly31**
Thr58	Asn56	Asn62	Asn56
Asn90	**Asn88**	**Asn93**	**Asn88**
Arg104	**Arg105**	**Arg107**	**Arg102**
Cys106	Cys107	Ala109	Cys104
Glu119	**Glu120**	**Glu122**	**Glu117**

3.2. Homologs of RutC and the putative active site

The cleft located between two monomers has been shown to bind various ligands: 2-ketobutyrate, ethylene glycol or serine in *E. coli* TdcF (Burman *et al.*, 2007), acetate in YabJ (Sinha *et al.*, 1999) and benzoate in the human protein hp14.5 (Manjasetty *et al.*, 2004). Previous analysis (Thakur *et al.*, 2010) has shown that those clefts vary in volume, charge distribution and amino-acid composition among YjgF-family representatives, and it was proposed that the cavity on the border of two monomers forms a ligand-binding site (Fig. 2*d*). The putative active site in the YjgF family is characterized by 7–9 residues which are highly conserved among YjgF proteins and are all located within the cleft (Volz, 1999). Residues from two monomers participate in the construction of one active site. These residues are also conserved in RutC and are listed in Table 2.

To map a putative ligand-binding site in the RutC protein, we superimposed its structure on those of other YjgF structures with bound ligands. The structure of the TdcF protein from *E. coli* bound with its physiological ligand 2-ketobutyrate and the structure of hp14.5 (*Homo sapiens*) with its ligand benzoate were both used for superposition. The putative binding site in RutC superposes with the TdcF 2-ketobutyrate binding site, which suggests that this cleft could be a physiological binding site (Figs. 2*e* and 2*f*). The β1–β2 loop forms a lid for this binding site in two of the six copies in the asymmetric unit, but is disordered in the other four chains.

RutC and TdcF share 32% sequence identity and have an r.m.s.d. of 1.6 Å for 128 aligned C$^\alpha$ atoms, while RutC and hp14.5 share 24% sequence identity and have an r.m.s.d. of 1.6 Å for 112 aligned C$^\alpha$ atoms. A sequence alignment of the RutC, TdcF, hp14.5 and YabJ proteins is presented in Fig. 3. The residues that are directly involved in binding ligands are identical in all three of the aligned proteins (Arg104 and Glu119 in RutC). The invariant arginine creates hydrogen bonds to the ligands of TdcF and hp14.5. In the case of 2-ketobutyrate, its carboxyl group creates a bidentate salt bridge with the Arg105 side chain.

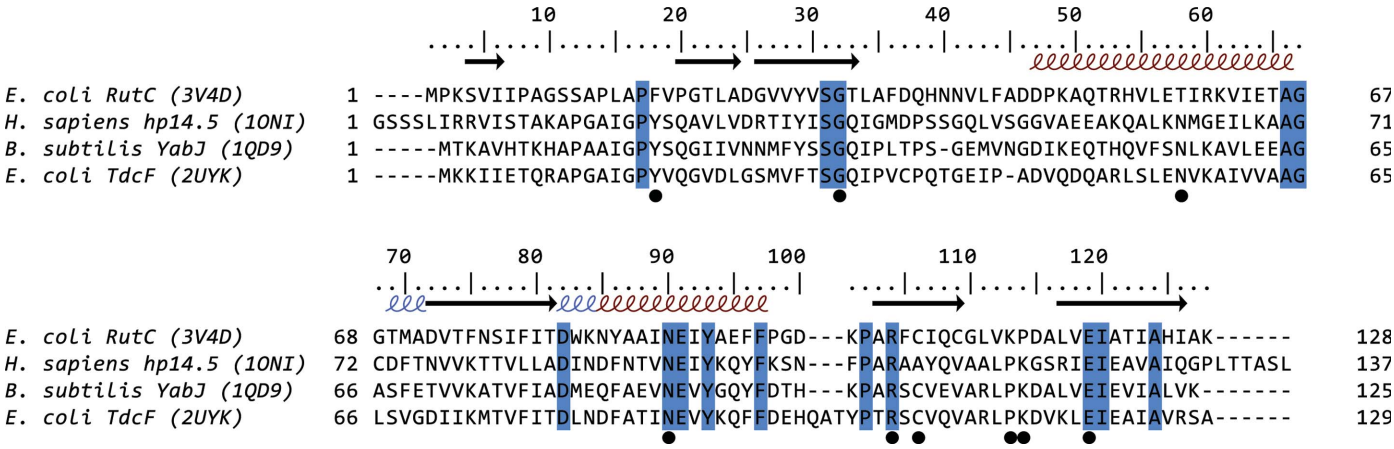


Figure 3
Alignment of YjgF-family members with RutC. The amino-acid sequences of the YjgF-family members assayed in this study were aligned using *ClustalW* (Thompson *et al.*, 1994) and manually edited using *BioEdit* (Hall, 1999). The numbering of the residues is based on the *E. coli* RutC sequence. Residues that are 100% conserved are highlighted; residues that are involved in active-site formation are marked with a dot. The secondary structure was calculated using *PROMOTIF* (Hutchinson & Thornton, 1996) on the basis of the RutC structure. Red and blue helices represent α-helices and 3_{10}-helices, respectively.

The second important residue is glutamic acid (Glu119 in RutC), which is present in all three aligned proteins and in TdcF creates hydrogen bonds to 2-ketobutyrate and other ligands. It has been suggested by Kim *et al.* (2010) that the $XC^{106}XXC^{109}$ motif could participate in the reduction of the peracid form of aminoacrylate. The conserved residue Cys107 in TdcF also creates hydrogen bonds to 2-ketobutyrate. If RutC catalyzes this proposed reaction, Cys106 in RutC could be in an analogous position. Another cysteine residue in the vicinity of the putative active site in RutC is Cys109. It is located 9 Å away from Cys106 and is not conserved among members of the YjgF family. In the homologous protein YjgF from *E. coli* (Volz, 1999), it was reported that a conserved cysteine residue (Cys107) underwent an unusual post-translational modification, which was identified as a thiosulfate or thiophosphate group with an unknown function (Feng *et al.*, 2008). The corresponding cysteine residue in RutC does not show this modification. The importance of this possible modification is still unknown, but it suggests that this residue may be of importance to the function and regulation of this protein.

4. Conclusions

The described RutC structure is the third solved structure of a protein from the *rut* operon. It has a characteristic trimeric *Bacillus* chorismate mutase fold. The *rut* operon is a recently discovered novel biochemical pathway and its enzymatic components might be direct targets for further drug development.

The authors would like thank Dr Matthew D. Zimmerman for valuable discussions and Rachel Vigour for proofreading the manuscript. The work described in this paper was supported by NIH grant GM094662 to the New York Structural Genomics Consortium, and GM53163. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor for the support of this research program (Grant 085P1000817).

References

Abagyan, R. & Totrov, M. (1997). *Proteins*, **29**, Suppl. 1, 215–220.
Burman, J. D., Stevenson, C. E., Sawers, R. G. & Lawson, D. M. (2007). *BMC Struct. Biol.* **7**, 30.
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.
Cowtan, K. D. & Zhang, K. Y. J. (1999). *Prog. Biophys. Mol. Biol.* **72**, 245–270.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Feng, J., Zhu, M., Schaub, M. C., Gehrig, P., Roschitzki, B., Lucchinetti, E. & Zaugg, M. (2008). *Cardiovasc. Res.* **80**, 20–29.
Hall, T. A. (1999). *Nucleic Acids Symp. Ser.* **41**, 95–98
Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
Hutchinson, E. G. & Thornton, J. M. (1996). *Protein Sci.* **5**, 212–220.
Kim, J.-M., Yoshikawa, H. & Shirahige, K. (2001). *Genes Cells*, **6**, 507–517.
Kim, K.-S., Pelton, J. G., Inwood, W. B., Andersen, U., Kustu, S. & Wemmer, D. E. (2010). *J. Bacteriol.* **192**, 4089–4102.
Knapik, A. A., Petkowski, J. J., Otwinowski, Z., Cymborowski, M. T., Cooper, D. R., Majorek, K. A., Chruszcz, M., Krajewska, W. M. & Minor, W. (2012). *Proteins*, **80**, 2359–2368.
Lambrecht, J. A., Browne, B. A. & Downs, D. M. (2010). *J. Biol. Chem.* **285**, 34401–34407.
Lambrecht, J. A., Flynn, J. M. & Downs, D. M. (2012). *J. Biol. Chem.* **287**, 3454–3461.
Manjasetty, B. A., Delbrück, H., Pham, D. T., Mueller, U., Fieber-Erdmann, M., Scheich, C., Sievert, V., Büssow, K., Niesen, F. H., Weihofen, W., Loll, B., Saenger, W., Heinemann, U. & Neisen, F. H. (2004). *Proteins*, **54**, 797–800.
Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta Cryst.* D**62**, 859–866.
Morishita, R., Kawagoshi, A., Sawasaki, T., Madin, K., Ogasawara, T., Oka, T. & Endo, Y. (1999). *J. Biol. Chem.* **274**, 20688–20692.
Mukherjee, T., Zhang, Y., Abdelwahed, S., Ealick, S. E. & Begley, T. P. (2010). *J. Am. Chem. Soc.* **132**, 5550–5551.
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.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Oka, T., Tsuji, H., Noda, C., Sakai, K., Hong, Y., Suzuki, I., Muñoz, S. & Natori, Y. (1995). *J. Biol. Chem.* **270**, 30060–30067.
Otwinowski, Z. (1991). *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–86. Warrington: Daresbury Laboratory.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Parsons, L., Bonander, N., Eisenstein, E., Gilson, M., Kairys, V. & Orban, J. (2002). *Biochemistry*, **42**, 80–89.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Rappu, P., Shin, B. S., Zalkin, H. & Mäntsälä, P. (1999). *J. Bacteriol.* **181**, 3810–3815.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Shimada, T., Hirao, K., Kori, A., Yamamoto, K. & Ishihama, A. (2007). *Mol. Microbiol.* **66**, 744–757.
Sinha, S., Rappu, P., Lange, S. C., Mäntsälä, P., Zalkin, H. & Smith, J. L. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 13074–13079.
Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1937–1940.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Thakur, K. G., Praveena, T. & Gopal, B. (2010). *Proteins*, **78**, 773–778.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Tuntufye, H. N., Lebeer, S., Gwakisa, P. S. & Goddeeris, B. M. (2012). *Appl. Environ. Microbiol.* **78**, 3343–3351.
Volz, K. (1999). *Protein Sci.* **8**, 2428–2437.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yang, H., Guranovic, V., Dutta, S., Feng, Z., Berman, H. M. & Westbrook, J. D. (2004). *Acta Cryst.* D**60**, 1833–1839.