

Published in final edited form as:

Proteins. 2010 February 15; 78(3): 773–778. doi:10.1002/prot.22623.

***Mycobacterium tuberculosis* Rv2704 is a member of the YjgF/YER057c/UK114 family**

Krishan Gopal Thakur¹, T. Praveena¹, and B. Gopal^{*}

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

Keywords

Chorismate mutase like fold; YjgF/YER057c/UK114 family; transcription; SIRAS; Sodium Iodide

Introduction

M. tuberculosis Rv2704 is a member of the highly conserved YjgF/YER057c/UK114 protein superfamily. Homologues of this protein occur in eubacteria, archaea and eukaryotes. Proteins in this family are functionally diverse and are involved in variety of enzymatic and non-enzymatic functions. The high sequence and structural similarity between members of this protein family provides an intriguing dataset to rationalize the wide variety of functional roles. This protein family is an example of minimalistic changes leading to functional diversification. This feature is best exemplified by the three close homologues of YjgF proteins in mammals (human, rat and goat) with sequence identity better than 85%. These homologues perform different functions¹⁻³. The diverse functions attributed to these proteins in this family include tumour antigen activity in the goat homologue¹, translation inhibition in human and rat homologues (hp14.5 and rp14.5)^{2,3}, endoribonuclease activity in rp14.5⁴, calpain activation in the bovine homologue⁵, molecular chaperone activity in DUK114⁶ and involvement in the regulation of purine and removal of toxic metabolites in YjgF^{7,8} and isoleucine (YjgF, YER057c, Ibm1) biosynthetic pathways⁹⁻¹¹. In addition, members of this protein family have also been shown to regulate mitochondrial maintenance (Ibm1) in yeast¹¹. Proteins from the YjgF family in plants are involved in photosynthesis and chromoplastogenesis (CHRD)¹².

The three dimensional structures of fifteen homologues of Rv2704 have been determined from bacteria, human, rat, goat and archaea^{7,13-18}. These proteins share similar homotrimeric structures with the clefts between the monomeric subunits proposed to have some functional relevance^{7,15,17} (Supplementary Table I). The best characterized protein with the YjgF fold is chorismate mutase which typifies this $\alpha+\beta$ fold and trimeric quaternary association. However, this protein shares very little sequence similarity with the other members of this family that were subsequently structurally characterized^{17,18}. Members of the YjgF family could potentially bind various metabolites. This suggestion was based on crystal structures of some homologues bound to ligands like benzoate¹⁵, acetate¹⁷, 2-ketobutyrate⁷, 1,2-ethanediol⁷, propanoate⁷ and serine⁷. In these crystal structures, the ligands bound at the cleft between the monomeric subunits of the trimer. This prompted a comparative analysis of this intersubunit cleft and an examination of role of conserved residues lining the cleft^{7,15}. No biological role could be assigned for these bound ligands

^{*}Corresponding Author Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012 Phone : (+91) 80 2293 3219 Fax: (+91) 80 23600535 bgopal@mbu.iisc.ernet.in.

¹These authors contributed equally to the work reported in this manuscript.

except perhaps for 2-ketobutyrate where it was proposed that YjgF may be involved in the removal of toxic products⁷. Even this functional implication, however, remains contestable with a recent report proposing that 2-ketobutyrate may not be the natural substrate for YjgF from *Salmonella enterica*⁹. These studies thus highlight the difficulties in structure-based functional assignment for proteins belonging to the YjgF family.

M. tuberculosis Rv2704 lies in the same operon as the principal sigma (σ) factor, σ^A (Supplementary Figure 1). σ factors are transcriptional proteins that are often regulated by their interactions with an anti- σ factor located in the same operon. Rv2704 was thus examined for its role as a putative regulator for σ^A . While we could not detect any interaction between Rv2704 and σ^A *in vitro*, it is still likely that it could regulate σ^A activity *in vivo*, perhaps involving metabolite interactions. Here we present the crystal structure of Rv2704 at 1.93Å resolution, solved using single isomorphous replacement with anomalous scattering (SIRAS). This protein is a trimer in solution with the overall structure similar to that of other YjgF homologues. The structure and biochemical features of Rv2704 could thus provide a template for more directed predictive methods for functional annotation in the YjgF family of proteins.

Materials and methods

Cloning, expression and purification of Rv2704

The gene encoding Rv2704 was PCR amplified from the genomic DNA of *M. tuberculosis* H37Rv and cloned between NheI and XhoI sites of the *E. coli* expression vector pET22b. After transforming the plasmid into Rosetta BL21(DE3) cells (Novagen, Inc.), the cells were grown in Luria broth in the presence of ampicillin (100 µg/ml) to an OD_{600nm} of 0.5–0.6. The cells were then induced with 0.3 mM IPTG (final concentration). Subsequently, the growth temperature was lowered to 290 K and cells were grown for further 12–18 h before they were spun down and stored at 193 K. Cells were lysed by sonication in a buffer containing 20 mM Tris pH 7.5 and 250 mM NaCl. The cell-free lysate was incubated with Co-NTA beads and washed with 10 bed volumes of lysis buffer in the column. Recombinant Rv2704 was eluted by a gradient of elution buffer (20 mM Tris pH 7.5, 250 mM NaCl and 200 mM Imidazole). The purity of the sample was analyzed by SDS-PAGE. The molecular weight of the recombinant protein was also verified by mass spectrometry on both MALDI-TOF as well as LC-ESI mass spectrometers (Bruker Daltonics, Inc.).

Crystallization

The initial crystallization conditions were obtained using commercial screens from Hampton Research by the hanging drop method. Recombinant Rv2704 could be concentrated to ca 80 mg/ml. Initial crystallization trials with 15 mg/ml protein concentration gave precipitates in more than 80 % of conditions. Subsequently reducing the protein concentration to 3 mg/ml provided initial crystals within 3-4 days of setting up the crystallization trays. Diffraction quality crystals were obtained in a condition containing 15% PEG 8000, 0.1 M Bis-Tris Propane (pH 6.5) and any one of several additives – N-docecyl-N,N-dimethylamine-N-oxide (0.5 % w/v), 1,6 hexane diol (3% v/v), glucose monohydrate (3% w/v) or sodium fluoride (50 mM) using the microbatch method. Rv2704 crystals were cryo-protected using 15% (v/v) glycerol in mother liquor prior to flash freezing at 100 K. NaI and CsCl derivatives were made adopting the quick soak methods described earlier^{19,20}. Transferring Rv2704 crystals from the crystallization droplet under oil to a solution on a cover slip containing cryo-protectant and derivative salts led to a substantial drop in diffraction quality. To circumvent this problem, the crystallization solution containing two-fold excess of the derivative salt was transferred to the well containing crystals (layered below mineral oil) in 1:1 ratio and gently mixed using a loop. These derivative crystals were soaked for various time points and

were tested for diffraction quality and isomorphism. The optimal conditions that gave isomorphous crystals with a reasonable anomalous signal involved soaking Rv2704 crystals in 500 mM NaI for ca 3 min. No anomalous signal was observed for the CsCl soaked crystals.

Structure determination and refinement

Molecular replacement trials using the structures of different YjgF homologues were performed using PHASER²¹ and MOLREP²². Although, MR solutions could be obtained with different models, the R_{free} did not drop below ca 35% even after repeated cycles of model building and refinement using COOT²³ and REFMAC²⁴. The electron density for the first 20 residues was poor and thus this segment could not be modeled. Phase information for the Rv2704 structure was thus obtained by Single Isomorphous Replacement with Anomalous Scattering (SIRAS) using the halide/alkali quick soak methods^{19,20,25,27}. Although both CsCl and NaI were examined, we were successful in obtaining anomalous signal with the NaI derivative alone. Diffraction data were collected at the home source (1.5418Å) on a Microstar UltraII X-Ray generator (Bruker AXS) and a MAR345 detector (Mar Research, Inc). Data were indexed and integrated using iMOSFLM²⁸ and were scaled using SCALA²⁹. The online Auto-Rickshaw server^{21,30-34} at EMBL was used to locate the positions of the Iodine atoms. Six heavy atom sites were located with occupancies ranging from 0.4 to 1. Several cycles of model building were performed using COOT²³ and the structure was refined using REFMAC²⁴. The locations of bound water molecules were initially identified by Arp/wArp³¹ and were subsequently added using COOT²³. The data collection, phasing and refinement statistics are summarized in Table I. The coordinates and structure factors have been deposited in the Protein Data Bank with the accession code 3I7T.

Results and discussion

Rv2704 has a chorismate mutase-like fold

There is a monomer of Rv2704 in the asymmetric unit of the crystal. The three fold axis of the Rv 2704 trimer is perfectly aligned with the crystallographic three fold axis (Figure 1A). This observation is consistent with the data from size exclusion chromatography experiments carried out using a Superdex G75 column (GE Healthcare) which suggests that Rv2704 is a trimer in solution (Figure 1B). This 142aa long polypeptide folds into a single compact domain as is the case with chorismate mutase, the canonical member of the fold. An alignment of YjgF-like protein sequences for which structural data is available suggests poor sequence conservation in the N-terminal segment (~20aa) of Rv2704 and an eleven residue extension in the C-terminal domain. Not surprisingly, these regions show structural differences when compared to the other YjgF proteins. The N-terminal region of Rv2704 adopts a α helical conformation spanning Ser12 to Gly19. This α helix is not seen in any of the other crystal structures of YjgF proteins determined thus far (Figure 1A). An interesting observation is that this helix (α 1) of Rv2704 is replaced by a loop with a high B-factor in other structures and has been proposed to play role in facilitating ligand binding¹⁵. 21 residues at the C-terminus of Rv2704 could not be modeled due to poor electron density. The introduction of an α helix thus differentiates the structure of Rv2704 from other YjgF proteins. The arrangement of the secondary structures in Rv2704 is β 1 α 1 β 2 β 3 α 2 β 4 α 3 β 5 β 6. This arrangement is different from that of other YjgF proteins which have the arrangement β 1 β 2 β 3 α 1 β 4 α 2 β 5 β 6^{7,15,17} (Supplementary Figure 2B). The other difference between the structure of Rv2704 and YjgF proteins is in a loop that connects α 2 and β 3. Multiple sequence alignment and structural superimposition suggests that this connecting loop is the shortest in Rv2704 (Figures 1A and 1C).

Structural features of Rv2704

The comparatively large (ca 15) dataset of YjgF protein structures allows for a comparative analysis of surface features, cavities and clefts in proteins with this fold. There are two prominent clefts in Rv2704 that are also conserved in a few YjgF homologues. One cleft is located at the trimeric interface (Type-I) while the other lies at the interface of two monomers i.e. AB, BC and CA (Type-II). Type-II clefts are often characterized by the presence of bound ligands. In a comparative analysis of 226 ligand binding clefts, the first and second largest pockets (clefts) included more than 80% of the binding sites³⁵. Indeed, it has been variously suggested that these clefts could serve as potential active sites for a catalytic role. However, no functional role for the cleft between the trimeric interface (type I) has been assigned in any of the YjgF proteins characterized till date. The type II clefts (between adjacent monomers) on the other hand, have been shown to bind various ligands—benzoate (in human hp14.5)¹⁵ or 2-ketobutyrate, acetate, ethylene glycol or serine (in *E.coli* TdcF)⁷. It is worth noting that while the type I cleft is less conserved (altogether absent in some cases), the type II cleft is maintained in all known YjgF homologues. Structural and sequence analyses of the residues lining the type-II cleft shows substantial variations in architecture, volume, charge distribution and sequence composition. Thus ligand specificity may not be hardwired for this binding site. For example, side-chain specific interactions are mediated by Arg105 and Glu114 in *E.coli* TdcF. The NH₂ and NE of the conserved Arg105 forms a bi-dentate salt-bridge with the carboxylate of 2-ketobutyrate. An Arg is present at a structurally equivalent position in several YjgF proteins. In Rv2704, Arg105 is replaced by Thr98 (Supplementary Figure 2A). However, Glu114 of Rv2704 is structurally equivalent to Glu120 of *E. coli* TdcF. In the crystal structure of TdcF solved in complex with several ligands, a majority of the conformational changes were restricted to the loop connecting β 1 and β 2. This loop becomes more ordered in the ligand bound form⁷. This loop lining the putative active-site cleft was also noted to have a high temperature factor in the human homologue¹⁵, a finding that is consistent across all crystal structures that were examined.

Sequence-structure correlation between members of the YjgF family

The multiple sequence alignment presented in Figure 1C shows that the closest protein from the YjgF family shares 37% sequence identity over 109 aligned residues with an overall sequence identity of 28%. Structural superimposition of the Rv2704 monomer with these YjgF proteins suggests that the fold is fairly well conserved. The root mean squared deviations (rmsd) over 114 Ca atoms between these structures vary from 1.4-1.8 Å with the sequence identity ranging from 21-35% (Supplementary Table I). The seven invariant residues in this alignment differ from the conservation pattern reported earlier³⁶. Thus the conserved Arg105 which is believed to play important role in the activity of YjgF proteins is replaced by Thr98 while the catalytic cysteine residue in *E. coli* YjgF¹⁸ is replaced by methionine in Rv2704 (Supplementary Fig. 2A). Another aspect, perhaps of interest from a bioinformatics perspective, is the difference in the physico-chemical properties among mammalian, archaeal and bacterial YjgF proteins. High acid stability is a characteristic feature of mammalian and archaeal YjgF proteins. Unlike homologues from human (hp14.5)¹⁵, rat (rp14.5)³ and archaea,¹⁶ Rv2704 is not stable in acidic conditions as it immediately precipitates in 5% perchloric acid solution.

The probable functions of *E. coli* TdcF and YabJ could be inferred from the role of the other gene products in the same operon. *E. coli* TdcF is a part of tdc operon which is involved in degradation of L-threonine and L-serine⁷, and YabJ forms part of *pur* operon³⁷. Unfortunately the functional annotation of majority of the YjgF proteins is difficult as their genes lie in operons with other hypothetical genes. As Rv2704 lies in the same operon as the principal σ factor σ^A , the probable role of Rv2704 as an anti- σ factor was examined. σ^A -Rv2704 interactions were examined using a co-expression and co-purification strategy. Both

genes (σ^A also referred to as Rv2703 and Rv2704) were cloned into the pET-Duet expression vector (Novagen, Inc.). σ^A was cloned in the Multiple Cloning Site 1 (MCS1) that would incorporate a poly-histidine tag in the recombinant protein and Rv2704 was cloned in MCS2. Both proteins were seen to express and a purification strategy using a Co-NTA affinity chromatography was adopted to obtain this complex. While this strategy was successful in several cases (data not shown), σ^A was not seen to interact with Rv2704 in this experiment. Rv2704 did not coelute with σ^A . In this context, it is important to note that while σ^A is an essential gene involved in the transcription of all house-keeping genes in *M. tuberculosis*, Rv2704 is not. Thus while direct interaction between Rv2704 and σ^A does not appear likely, the structure of Rv2704 suggests a different possibility. The finding that the members of the YjgF family bind different metabolites suggests that Rv2704 could modulate its function in the presence of different metabolites. The structural and biochemical features of Rv2704 could thus serve as a template to further examine the role of this protein in *M. tuberculosis*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by the Wellcome Trust, United Kingdom. The X-ray facility is supported by Grants-in-Aid from the Department of Science and Technology and the Department of Biotechnology, Government of India. BG is an International Senior Research Fellow of the Wellcome Trust, U.K.

References

1. Ceciliani F, Faotto L, Negri A, Colombo I, Berra B, Bartorelli A, Ronchi S. The primary structure of UK114 tumor antigen. *FEBS Lett.* 1996; 393(2-3):147–150. [PubMed: 8814279]
2. Oka T, Tsuji H, Noda C, Sakai K, Hong YM, Suzuki I, Munoz S, Natori Y. Isolation and characterization of a novel perchloric acid-soluble protein inhibiting cell-free protein synthesis. *J Biol Chem.* 1995; 270(50):30060–30067. [PubMed: 8530410]
3. Schmiedeknecht G, Kerkhoff C, Orso E, Stohr J, Aslanidis C, Nagy GM, Knuechel R, Schmitz G. Isolation and characterization of a 14.5-kDa trichloroacetic-acid-soluble translational inhibitor protein from human monocytes that is upregulated upon cellular differentiation. *Eur J Biochem.* 1996; 242(2):339–351. [PubMed: 8973653]
4. Morishita R, Kawagoshi A, Sawasaki T, Madin K, Ogasawara T, Oka T, Endo Y. Ribonuclease activity of rat liver perchloric acid-soluble protein, a potent inhibitor of protein synthesis. *J Biol Chem.* 1999; 274(29):20688–20692. [PubMed: 10400702]
5. Melloni E, Michetti M, Salamino F, Pontremoli S. Molecular and functional properties of a calpain activator protein specific for mu-isoforms. *J Biol Chem.* 1998; 273(21):12827–12831. [PubMed: 9582310]
6. Farkas A, Nardai G, Csermely P, Tompa P, Friedrich P. DUK114, the *Drosophila* orthologue of bovine brain calpain activator protein, is a molecular chaperone. *Biochem J.* 2004; 383(Pt 1):165–170. [PubMed: 15250825]
7. Burman JD, Stevenson CE, Sawers RG, Lawson DM. The crystal structure of *Escherichia coli* TdcF, a member of the highly conserved YjgF/YER057c/UK114 family. *BMC Struct Biol.* 2007; 7:30. [PubMed: 17506874]
8. Schmitz G, Downs DM. Reduced transaminase B (IlvE) activity caused by the lack of yjgF is dependent on the status of threonine deaminase (IlvA) in *Salmonella enterica* serovar Typhimurium. *J Bacteriol.* 2004; 186(3):803–810. [PubMed: 14729707]
9. Christopherson MR, Schmitz GE, Downs DM. YjgF is required for isoleucine biosynthesis when *Salmonella enterica* is grown on pyruvate medium. *J Bacteriol.* 2008; 190(8):3057–3062. [PubMed: 18296521]

10. Enos-Berlage JL, Langendorf MJ, Downs DM. Complex metabolic phenotypes caused by a mutation in *yjgF*, encoding a member of the highly conserved YER057c/YjgF family of proteins. *J Bacteriol.* 1998; 180(24):6519–6528. [PubMed: 9851994]
11. Kim JM, Yoshikawa H, Shirahige K. A member of the YER057c/yjgF/Uk114 family links isoleucine biosynthesis and intact mitochondria maintenance in *Saccharomyces cerevisiae*. *Genes Cells.* 2001; 6(6):507–517. [PubMed: 11442631]
12. Leitner-Dagan Y, Ovadis M, Zuker A, Shklarman E, Ohad I, Tzfira T, Vainstein A. CHR1, a plant member of the evolutionarily conserved YjgF family, influences photosynthesis and chromoplastogenesis. *Planta.* 2006; 225(1):89–102. [PubMed: 16845531]
13. Deaconescu AM, Roll-Mecak A, Bonanno JB, Gerchman SE, Kycia H, Studier FW, Burley SK. X-ray structure of *Saccharomyces cerevisiae* homologous mitochondrial matrix factor 1 (Hmf1). *Proteins.* 2002; 48(2):431–436. [PubMed: 12112709]
14. Deriu D, Briand C, Mistiniene E, Naktinis V, Grutter MG. Structure and oligomeric state of the mammalian tumour-associated antigen UK114. *Acta Crystallogr D Biol Crystallogr.* 2003; 59(Pt 9):1676–1678. [PubMed: 12925811]
15. Manjasetty BA, Delbruck H, Pham DT, Mueller U, Fieber-Erdmann M, Scheich C, Sievert V, Bussow K, Niesen FH, Weihofen W, Loll B, Saenger W, Heinemann U. Crystal structure of *Homo sapiens* protein hp14.5. *Proteins.* 2004; 54(4):797–800. [PubMed: 14997576]
16. Miyakawa T, Lee WC, Hatano K, Kato Y, Sawano Y, Miyazono K, Nagata K, Tanokura M. Crystal structure of the YjgF/YER057c/UK114 family protein from the hyperthermophilic archaeon *Sulfolobus tokodaii* strain 7. *Proteins.* 2006; 62(2):557–561. [PubMed: 16323205]
17. Sinha S, Rappu P, Lange SC, Mantsala P, Zalkin H, Smith JL. Crystal structure of *Bacillus subtilis* YabJ, a purine regulatory protein and member of the highly conserved YjgF family. *Proc Natl Acad Sci U S A.* 1999; 96(23):13074–13079. [PubMed: 10557275]
18. Volz K. A test case for structure-based functional assignment: the 1.2 Å crystal structure of the *yjgF* gene product from *Escherichia coli*. *Protein Sci.* 1999; 8(11):2428–2437. [PubMed: 10595546]
19. Dauter Z, Dauter M. Entering a new phase: using solvent halide ions in protein structure determination. *Structure.* 2001; 9(2):R21–26. [PubMed: 11250204]
20. Dauter Z, Dauter M, Rajashankar KR. Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. *Acta Crystallogr D Biol Crystallogr.* 2000; 56(Pt 2):232–237. [PubMed: 10666615]
21. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr.* 2007; 40(Pt 4):658–674. [PubMed: 19461840]
22. Vagin A, Teplyakov A. *MOLREP*: an Automated Program for Molecular Replacement. *J Appl Cryst.* 1997; 30:1022–1025.
23. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004; 60(Pt 12 Pt 1):2126–2132. [PubMed: 15572765]
24. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr.* 1997; 53(Pt 3):240–255. [PubMed: 15299926]
25. Dauter M, Dauter Z. Phase determination using halide ions. *Methods Mol Biol.* 2007; 364:149–158. [PubMed: 17172764]
26. Korolev S, Dementieva I, Sanishvili R, Minor W, Otwinowski Z, Joachimiak A. Using surface-bound rubidium ions for protein phasing. *Acta Crystallogr D Biol Crystallogr.* 2001; 57(Pt 7):1008–1012. [PubMed: 11418770]
27. Nagem RA, Dauter Z, Polikarpov I. Protein crystal structure solution by fast incorporation of negatively and positively charged anomalous scatterers. *Acta Crystallogr D Biol Crystallogr.* 2001; 57(Pt 7):996–1002. [PubMed: 11418768]
28. Leslie, A. Joint CCP4+ESF-EAMCB Newsletter on Protein Crystallography. 1992. No6
29. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr.* 1994; 50(Pt 5):760–763. [PubMed: 15299374]
30. Cowtan KD, Zhang KY. Density modification for macromolecular phase improvement. *Prog Biophys Mol Biol.* 1999; 72(3):245–270. [PubMed: 10581970]

31. Morris RJ, Zwart PH, Cohen S, Fernandez FJ, Kakaris M, Kirillova O, Vonnrhein C, Perrakis A, Lamzin VS. Breaking good resolutions with ARP/wARP. *J Synchrotron Radiat.* 2004; 11(Pt 1): 56–59. [PubMed: 14646134]
32. Panjikar S, Parthasarathy V, Lamzin VS, Weiss MS, Tucker PA. Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr D Biol Crystallogr.* 2005; 61(Pt 4):449–457. [PubMed: 15805600]
33. Sheldrick GM, Schneider TR. SHELXL: high-resolution refinement. *Methods Enzymol.* 1997; 277:319–343. [PubMed: 18488315]
34. Terwilliger TC. Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr.* 2000; 56(Pt 8):965–972. [PubMed: 10944333]
35. Glaser F, Morris RJ, Najmanovich RJ, Laskowski RA, Thornton JM. A method for localizing ligand binding pockets in protein structures. *Proteins.* 2006; 62(2):479–488. [PubMed: 16304646]
36. Parsons L, Bonander N, Eisenstein E, Gilson M, Kairys V, Orban J. Solution structure and functional ligand screening of HI0719, a highly conserved protein from bacteria to humans in the YjgF/YER057c/UK114 family. *Biochemistry.* 2003; 42(1):80–89. [PubMed: 12515541]
37. Weng M, Nagy PL, Zalkin H. Identification of the *Bacillus subtilis* pur operon repressor. *Proc Natl Acad Sci U S A.* 1995; 92(16):7455–7459. [PubMed: 7638212]
38. Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, Pupko T, Ben-Tal N. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.* 2005; 33(Web Server issue):W299–302. [PubMed: 15980475]
39. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994; 22(22):4673–4680. [PubMed: 7984417]
40. Gouet P, Courcelle E, Stuart DI, Metz F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics.* 1999; 15(4):305–308. [PubMed: 10320398]

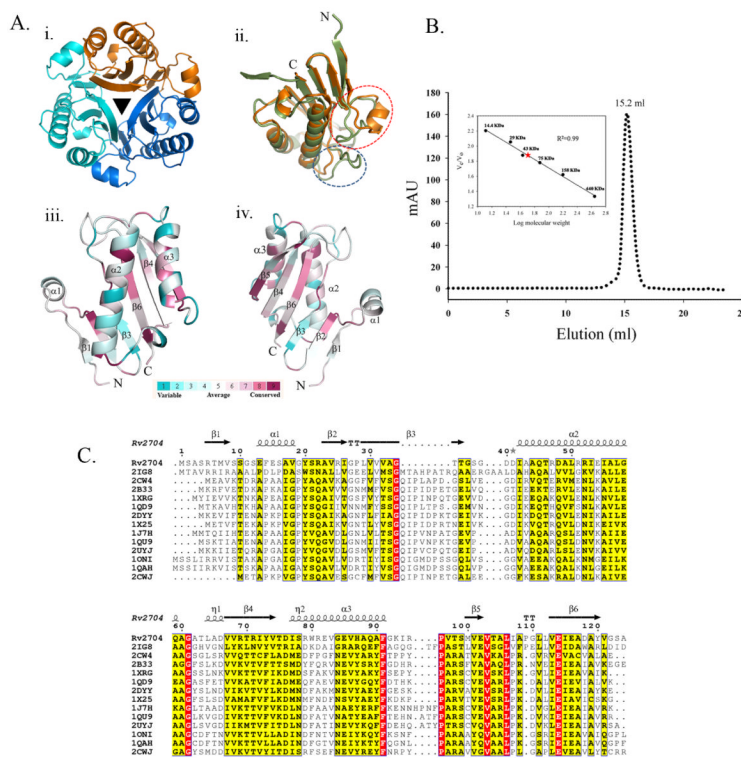


Fig. 1.
(A) Structural features of Rv2704. **i.** The trimeric arrangement of Rv2704 monomers. **ii.** The crystal structure of Rv2704 differs from other YjgF homologs in a loop segment that adopts an α helical structure ($\alpha 1$) (red circle) and in a loop that connects $\alpha 2$ and $\beta 3$ which is shorter than those in other YjgF proteins (blue circle). **iii.** Sequence conservation mapped on to the crystal structure of Rv2704 reveals highly conserved regions in YjgF proteins. This figure was drawn using ConSurf³⁸. **(B) Size exclusion profile of Rv2704.** Rv2704 is a trimer in solution. **(C) Multiple sequence alignment** of YjgF proteins reveals seven invariant residues (highlighted in red) and a high overall sequence similarity (shown in yellow). The C-terminal extension of Rv2704 (Val125-Gly142) is not shown in this sequence alignment. This alignment and presentation were made using CLUSTALW³⁹ and ESPript⁴⁰.

TABLE I

Crystallographic Data and Refinement Statistics

I. Data collection		
	Native	NaI derivative
Wavelength	1.5418 Å	1.5418 Å
Resolution (Å)	33.96 – 1.94 (2.05-1.94) ^a	33.19 – 1.93 (2.03-1.93)
Unit cell parameters	a = b = c = 67.92 Å	a = b = c = 67.88 Å
Space group	P2 ₁ 3	P2 ₁ 3
Total number of reflections	44811 (6489)	144413 (20323)
Number of unique reflections	7959 (1141)	8092 (1149)
Completeness (%)	99.7 (100)	100 (100)
Multiplicity	5.6 (5.3)	17.8 (17.7)
R _{merge} ^b (%)	12.1 (50.3)	10 (42.4)
<I>/σ(I)	8.7 (3.1)	21.5 (6.6)
II. Refinement statistics		
Resolution range		33.19 – 1.93
rmsd _{bonds} (Å)		0.009
rmsd _{angles} (°)		1.241
R _{cryst} ^c / R _{free} ^d (%)		20.5/24.7
No. of protein atoms		877
No. of solvent atoms		61
III. Ramachandran plot statistics		
Residues in most favoured regions (%)		94.1
Residues in allowed regions (%)		5.9
Residues in generously allowed regions (%)		0
Residues disallowed regions (%)		0

^aValues for outer shells are given in parentheses.

^bR_{merge} = $\sum_j \sum_i |I_j - \langle I \rangle| / \sum_j I_j$ where I_j is the intensity of the j^{th} reflection and $\langle I \rangle$ is the average intensity.

^cR_{cryst} = $\sum_{hkl} |F_o - F_c| / \sum_{hkl} |F_o|$

^dR_{free} was calculated as for R_{cryst} but on 5 % of the data excluded from the refinement calculation.