Structure of the *Bacillus anthracis* dTDP-L-rhamnose-biosynthetic enzyme glucose-1-phosphate thymidyltransferase (RfbA)

Jackson Baumgartner,^{a,†} Jesi Lee,^{a,†} Andrei S. Halavaty,^{b,c} George Minasov,^{b,c} Wayne F. Anderson^{b,c} and Misty L. Kuhn^{a,*}

Received 7 August 2017

Accepted 22 October 2017

Edited by P. Dunten, Stanford Synchrotron Radiation Lightsource, USA

† These authors contributed equally to this project.

Keywords: RfbA; glucose-1-phosphate thymidyltransferase; *Bacillus anthracis*; Anthrax.

PDB reference: RfbA, 4ecm

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

^aDepartment of Chemistry and Biochemistry, San Francisco State University, USA, ^bDepartment of Biochemistry and Molecular Genetics, Northwestern University Feinberg School of Medicine, USA, and ^cCenter for Structural Genomics of Infectious Diseases (CSGID), USA. *Correspondence e-mail: mkuhn@sfsu.edu

L-Rhamnose is a ubiquitous bacterial cell-wall component. The biosynthetic pathway for its precursor dTDP-L-rhamnose is not present in humans, which makes the enzymes of the pathway potential drug targets. In this study, the three-dimensional structure of the first protein of this pathway, glucose-1-phosphate thymidyltransferase (RfbA), from *Bacillus anthracis* was determined. In other organisms this enzyme is referred to as RmlA. RfbA was co-crystallized with the products of the enzymatic reaction, dTDP- α -D-glucose and pyrophosphate, and its structure was determined at 2.3 Å resolution. This is the first reported thymidyltransferase structure from a Gram-positive bacterium. RfbA shares overall structural characteristics with known RmlA homologs. However, RfbA exhibits a shorter sequence at its C-terminus, which results in the absence of three α -helices involved in allosteric site formation. Consequently, RfbA was observed to exhibit a quaternary structure that is unique among currently reported glucose-1-phosphate thymidyltransferase bacterial homologs. These structural analyses suggest that RfbA may not be allosterically regulated in some organisms and is structurally distinct from other RmlA homologs.

1. Introduction

Anthrax is an infectious disease that has historically plagued workers in a variety of industries involving animals or animal byproducts (Sternbach, 2003). The most deadly form is inhalation Anthrax, which ultimately results in pleural effusion, with a 45% mortality rate even when treated (Hendricks *et al.*, 2014). During the 20th century, the number of fatalities decreased significantly owing to improved hygiene practices; however, Anthrax still poses a serious public health threat as a potential bioterror agent or a re-emerging infectious disease. The causative agent of Anthrax is the endospore-forming Gram-positive bacterium *Bacillus anthracis*, which is present in soil worldwide (Sternbach, 2003).

Bacterial cell walls consist of multiple layers of peptidoglycan and polysaccharides, which provide permeability for small molecules and a protective structure. L-Rhamnose is incorporated into the bacterial cell wall as rhamnose-containing cell-wall polysaccharides (Mistou *et al.*, 2016; Doran & Mattingly, 1982; Wagner *et al.*, 1978) using the activated sugar dTDP-L-rhamnose. Several studies have shown that the loss of rhamnose leads to impeded cell growth and cell-wall division abnormalities (van der Beek *et al.*, 2015; van Sorge *et al.*, 2014); therefore, the enzymes involved in its production may be targets for therapeutics to combat the disease.

Table 1

Macromolecule-production information.

The *rfbA* gene encodes the glucose-1-phosphate thymidyltransferase/dTDP- α -D-glucose pyrophosphorylase.

Source organism	<i>B. anthracis</i> strain Ames
Cloning and expression vector	pMCSG7 (ampicillin-resistant)
Expression host	<i>E. coli</i> BL21(DE3) Magic cells
Complete amino-acid sequence of the construct produced	MHHHHHSSGVDLGTENLYFQSNAMKGII AGGTGSRLYPITKVTNKHLLPVGRYPMI YHAVYKLLKQCDITDIMITGKEHMGDVV SFLGSGQEFVGSFTYRVQDKAGGIAQAL GLCEDFVGNDRMVILGDNIFSDDIRPY VEEFTNQKEGAKVLLQSVDDPERFGVAN IQNRKIIIEIEEKPKPKSSYAVTGIYLY DSKVFSYIKELKPSARGELEITDINNWI LKRGLVTYNEMSGWWTAGTHVSLQRAN ALARDINFGKQFNGE

Table 2

Crystallization.

Method	Vapor diffusion, sitting drop
Plate type	96-well microplate
Temperature (K)	295
Protein concentration (mg ml ⁻¹)	7.5
Buffer composition of protein solution	10 mM Tris-HCl pH 8.3, 500 mM NaCl, 5 mM BME, 2 mM MgCl ₂ , 2 mM dTDP- α -D-glucose
Composition of reservoir solution	0.2 M sodium sulfate, 0.1 M bis-tris propane pH 8.5, 20% (w/v) PEG 3350
Volume and ratio of drop	1 μ l:1 μ l
Volume of reservoir (μ l)	100

In bacteria, dTDP-L-rhamnose is produced from α -D-glucose-1-phosphate (Glc1P) using a four-enzyme biosynthetic pathway. The first enzyme of this pathway, Glc1P-thymidyltransferase (RfbA; Rfb homologs are frequently termed Rml in other organisms), transfers thymidyl monophosphate from dTTP to Glc1P, which forms dTDP- α -D-glucose and pyrophosphate (PP_i). Three additional enzymes in the pathway, RfbB, RfbC and RfbD, carry out further reactions to yield the final product dTDP-L-rhamnose, which is thought to allosterically regulate RfbA (Melo & Glaser, 1965; Blankenfeldt *et al.*, 2000). RfbA from *B. anthracis* is only 245 residues in length, while its RmlA homologs are 295 residues in length. RmlA homologs from Gram-negative bacteria have previously been reported (Blankenfeldt *et al.*, 2000; Sivaraman *et al.*, 2002; Barton *et al.*, 2001), but information regarding Gram-positive RfbA or RmlA homologs is limited. Here, we report the first three-dimensional crystal structure of an RfbA protein from *B. anthracis* in complex with dTDP- α -D-glucose and PP_i, which provides new insight into the differences between RfbA and RmlA proteins.

2. Materials and methods

2.1. Protein expression and purification

The wild-type *rfbA* gene (GenBank ID AAP25186) from *B. anthracis* strain Ames was cloned into the ampicillin-resistant pMCSG7 vector, which contains an N-terminal polyhistidine tag followed by a TEV cleavage site and the start codon of the *rfbA* gene. It was transformed into kanamycin-

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	Beamline 21-ID-G, APS
Wavelength (Å)	0.97856
Temperature (K)	100
Detector	MAR Mosaic 300 mm CCD
Space group	<i>P</i> 6 ₂₂
<i>a</i> , <i>b</i> , <i>c</i> (Å)	135.07, 135.07, 85.14
α , β , γ (°)	90.00, 90.00, 120.00
Resolution range (Å)	30.00–2.30 (2.34–2.30)
No. of unique reflections	20938 (1022)
Completeness (%)	99.9 (100.0)
Multiplicity	10.7 (10.9)
$\langle I/\sigma(I) \rangle$	22.9 (5.1)
<i>R</i> _{int}	0.059 (0.293)
Overall <i>B</i> factor from Wilson plot (Å ²)	32.4

† Estimated $R_{int} = R_{merge}[N/(N-1)]^{1/2}$, where *N* is the data multiplicity.

resistant *Escherichia coli* BL21(DE3) Magic cells as described previously (Kwon & Peterson, 2014; Table 1). The cells were grown at 310 K in 1.5 l Terrific Broth until the OD_{600 nm} reached 0.6–0.8. The cells were then placed on ice for 20 min, after which protein expression was induced with 0.6 mM IPTG. The proteins were expressed overnight with shaking at 298 K and harvested as described previously (Kuhn *et al.*, 2013). The cell pellets were resuspended in 100 ml lysis buffer [10 mM Tris-HCl pH 8.3, 500 mM NaCl, 5 mM β -mercaptoethanol (BME), 10% (v/v) glycerol, 0.01% (v/v) IGEPAL CA630], sonicated and centrifuged. The soluble fraction was applied onto a 5 ml Ni-NTA affinity column and purified and concentrated as described previously (Kuhn *et al.*, 2013). The yield of pure protein was 22 mg l⁻¹. The N-terminal polyhistidine tag was not removed prior to crystallization and the protein was immediately set up for crystallization.

2.2. Crystallization

The RfbA protein was screened for crystallization at a concentration of 7.5 mg ml⁻¹ (0.27 mM) in 10 mM Tris-HCl pH 8.3, 500 mM NaCl, 5 mM BME, 2 mM MgCl₂, 2 mM dTDP- α -D-glucose (Sigma-Aldrich) against The Classics, PACT, JCSG+ and AmSO₄ Suites (Qiagen). 1 μ l protein solution was added to 1 μ l reservoir solution in a sitting-drop vapor-diffusion 96-well microplate at 298 K. Crystals of RfbA grew in 0.2 M sodium sulfate, 0.1 M bis-tris propane pH 8.5, 20% (w/v) PEG 3350 (The PACT Suite condition H8; Table 2).

2.3. Data collection and processing

Crystals of RfbA were transferred from the drop and soaked in equal proportions of crystallization precipitant and a solution consisting of 3.6 M ammonium sulfate and 50% sucrose prior to flash-cooling in liquid nitrogen. Data were collected on the LS-CAT 21-ID-G beamline at the Advanced Photon Source (APS) at Argonne National Laboratory. HKL-3000 (Minor *et al.*, 2006) was used to index, scale and integrate the data. Statistics for data collection and data processing are described in Table 3.

Table 4
Structure refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	29.24–2.30 (2.36–2.30)
Completeness (%)	99.9 (100.0)
No. of reflections, working set	19775 (1416)
No. of reflections, test set	1073 (81)
Final R_{cryst}	0.153 (0.154)
Final R_{free}	0.186 (0.206)
No. of non-H atoms	
Protein	1940
Ligand	45
Solvent	226
Total	2211
R.m.s. deviations	
Bonds (Å)	0.011
Angles (°)	1.526
Average B factors (Å ²)	
Protein	30.9
Ligands	
DAU	33.9
POP	55.7
Water	36.7
Ramachandran plot†	
Favored regions (%)	90.5
Additionally allowed (%)	9.0
Outliers (%)	0.5

† Ramachandran plot statistics are based on *PROCHECK* (Laskowski *et al.*, 1993).

2.4. Structure solution and refinement

The structure of RfbA in complex with dTDP- α -D-glucose and PP_i (PDB entry 4ecm) was determined using molecular replacement with *Phaser* (McCoy *et al.*, 2007) from the *CCP4* suite (Winn *et al.*, 2011). All residues of one monomer of PDB entry 3hl3 (100% identity; Center for Structural Genomics of Infectious Diseases, unpublished work) were used as a model for molecular replacement, and the initial structure was rebuilt with *ARP/wARP* (Morris *et al.*, 2003) and manually modified using *Coot* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010). *REFMAC* v.5.5 was used to refine the structure, and translation–libration–screw (TLS) groups were used from the *TLSMD* server (Painter & Merritt, 2006; <http://skuldbmsc.washington.edu/~tlsmd/>) during the latter stages of refinement. The quality of the structure was determined using *MolProbity* (Chen *et al.*, 2010; <http://molprobity.biochem.duke.edu/>) and the PDB validation server (<http://deposit.pdb.org/validate/>). Structure-refinement statistics are presented in Table 4. The structure contained a Phe residue (−4) and an Ala residue (0) from the protease-cleavage site of the construct. The remaining residues of the affinity tag and the last two residues of the C-terminus (Gly244 and Glu245) were disordered.

3. Results and discussion

3.1. Overall three-dimensional structure of RfbA

The crystal structure of RfbA in complex with the reaction products dTDP- α -D-glucose and PP_i (PDB entry 4ecm) was determined at 2.3 Å resolution (Fig. 1*a* and 1*b*). RfbA crystallized in space group *P*6₂2 with *D*₂ symmetry and one monomer in the asymmetric unit. PP_i was not added for co-crystallization and is likely to be a contaminant from the

added dTDP- α -D-glucose. Therefore, it was modeled with an occupancy of 0.6 and still had atomic *B* values that were higher than the *B* values from the Wilson plot (the electron density surrounding the ligands is shown in Supplementary Fig. S1). Several disordered residues, including the polyhistidine tag and the last two residues of the C-terminus, were not present in the structure. The presence of the Ramachandran plot outlier (Table 4) can be explained by the fact that an alternative conformation was built for the main chain between Thr21 and Asn22 because a single conformation gave negative density for the carbonyl O atom of Thr21. The outlier was modeled with an occupancy of 0.5. The mixed β -sheet of the RfbA monomer forms the core of the protein and is sandwiched between two clusters of α -helices, which results in the three-layered $\alpha/\beta/\alpha$ fold common to nucleotide-diphosphosugar transferases (Fig. 1*b*). The biological assembly of RfbA is a homotetramer in which the active sites of each monomer face in opposite directions (Fig. 1*a*).

3.2. Active site of RfbA

The active site of RfbA is located at one edge of the core β -sheet and is surrounded by α -helices where both dTTP and Glc1P bind. Two small antiparallel β -sheets formed by β 2/ β 3 and β 7/ β 13 have previously been shown to contain residues that recognize the sugar-phosphate moiety of Glc1P in the homologous bacterial RmlAs (Giraud & Naismith, 2000). The larger nucleotide-binding region of the protein includes the N-terminal β -strands 1–7 and α -helices 1–4 and 8, while the smaller sugar-binding region is comprised of residues 116–221 (Sivaraman *et al.*, 2002). dTDP- α -D-glucose and PP_i were bound in the RfbA structure and form several hydrogen bonds to specific active-site residues (Fig. 1*f*). For instance, dTDP- α -D-glucose forms hydrogen bonds to Gly8, Arg13, His24, Gln80, Asn109, Gly143, Glu157 and Lys158, and PP_i forms hydrogen bonds to Gly11, Ser12, Arg13 and the catalytically critical Lys23 residue (Blankenfeldt *et al.*, 2000; Figs. 1*c* and 1*d*).

Although 2 mM MgCl₂ was added during co-crystallization, no magnesium ions were observed in the crystal structure. Magnesium ions are known to be responsible for stabilization of the phosphate groups of the substrates and products of the reaction, and are absolutely required for catalysis (Glaser & Kornfeld, 1961). The binding site for magnesium has been determined in the structure of the RfbA homolog RffH from *E. coli* (PDB entry 1mc3; Sivaraman *et al.*, 2002). When this structure is overlaid with that of RfbA, magnesium is found to be coordinated to the Asp108 and Asp219 residues (Fig. 1*e*). Asp219 in the RfbA structure was observed in two conformations: one with the side chain pointing towards the dTDP- α -D-glucose and the other with the side chain pointing away from it.

The N^ε atom of His24 that makes a hydrogen bond to the 3'-OH of the deoxyribose of dTDP- α -D-glucose (Figs. 1*c* and 2*a*) is not conserved compared with RmlA, which instead has a strictly conserved glutamine that serves the same function. Previous mutagenesis studies have suggested that mutating

Figure 1

Three-dimensional crystal structure of RfbA. (a) Ribbon representation of the RfbA tetramer. Each monomer is represented in a different color. (b) RfbA monomer with secondary structure labeled from the N-terminus (blue) to the C-terminus (red). (c) RfbA active site with dTDP- α -D-glucose highlighted as thick green sticks. PP_i is shown as balls and thin sticks. (d) RfbA active site with PP_i highlighted as thick sticks. dTDP- α -D-glucose is shown as balls and thin sticks. (e) Superposition of the active sites of RfbA and RffH (PDB entry 1mc3). The Mg²⁺ ion is modeled from the RffH structure. (f) Surface representation of hydrophobic residues in the active site of RfbA that contribute to the stabilization of the thymidine and glucose moieties of dTDP- α -D-glucose. The residues involved in these patches are mostly conserved. The hydrophobic bed at the top-right of the figure interacts with the thymidine moiety; the representative residues are Leu6, Gly9, Ala83, Ile86 and Leu106. The lower-left bed forms hydrophobic interactions with the glucose moiety, which are mainly formed by Phe142, Val168 and Trp217.

this residue alters the activity of the enzyme towards other purine and pyrimidine nucleotides (Jakeman *et al.*, 2008;

Moretti *et al.*, 2011). The specificity pocket for the thymidine base in RfbA is formed by Gly8, Gly9, Gln80, Ala83, Gly85

(a)

(b)

(c)

Figure 2

Comparison of RfbA with its homologs. (a) Multiple sequence alignment of RfbA and its homologs. Those from *P. aeruginosa* (PDB entry 1g3l), *A. thermoaerophilus* (PDB entry 4ho4) and *Salmonella enterica* (PDB entry 1iin, K24Q mutant; Barton *et al.*, 2001) are RmlAs. The *E. coli* (PDB entry 1mc3) homolog is annotated as RfIH. The GalU homologs are from *Burkholderia vietnamiensis* (PDB entry 5iif; Seattle Structural Genomics Center for Infectious Disease, unpublished work) and *Sphingomonas elodea* (PDB entry 2ux8), respectively. bcRfbA (IG_03809) and bcRmlA (IG_05050) are both from *Bacillus cereus* HuA2-9. Residues highlighted in blue are strictly conserved, whereas those marked in yellow represent functional conservation. In PDB entry 1g3l the residues in the allosteric pocket are highlighted in red (contributed from the same chain) and brown (contributed by the adjacent monomer). The DALI server (Holm & Rosenström, 2010), Clustal Omega (Goujon *et al.*, 2010) and ESPript (Xavier & Gouet, 2014) were used to generate the alignment. (b) Superposition of RfbA from *B. anthracis* (yellow) and RmlA from *P. aeruginosa* (gray) with dTDP-α-D-glucose in cyan ball-and-stick representation in the active site. The three C-terminal helices that are only present in RmlA are highlighted in red. (c) Superposition of RfbA from *B. anthracis* (yellow) and GalU from *S. elodea* (gray) with dTDP-α-D-glucose in cyan ball-and-stick representation in the active site. The additional C-terminal helix (GalUα10) and the additional region between α2 and β5 of GalU (GalUα3) are highlighted in purple.

Figure 3

Tetrameric organization of RfbA and homologs. (a) Tetrameric assembly of RfbA. A circle diagram representing the orientation of each monomer is colored as in Fig. 1(a). The diagonal lines represent the orientation of the core β -sheet. (b) Tetrameric assembly of RmlA from *P. aeruginosa*. The red arrows indicate the locations of the allosteric sites of RmlA. The orientation of monomers in the RmlA tetramer is different from that in RfbA and GalU. (c) Tetrameric assembly of GalU from *S. elodea*. The colors of the circle diagrams are equivalent to those for RfbA. (d) Superposition of RfbA and RmlA tetramers (coloring as described in Fig. 2). The dTDP- α -D-glucose of RfbA is depicted in cyan ball-and-stick representation, while the dTDP-L-rhamnose bound in the allosteric site of RmlA is shown as green spheres. Top: ribbon representation of the RfbA and RmlA tetramers superimposed. Bottom: bird's-eye view of the region where the allosteric site forms in RmlA. One monomer of RmlA is shown in a ribbon representation (and overlaid with one monomer of RfbA in yellow), while the other is shown as a surface representation. The surface representation clearly shows that the three C-terminal helices of RmlA heavily contribute not only to the dimer interface, but also to formation of the allosteric binding pocket.

and Ile86, and is too small for purines to enter; the hydrogen bond from Gly8 to the C2 carbonyl of thymidine is likely to aid in pyrimidine specificity. Additionally, Gln80 and the amide backbone of Gly85 make hydrogen bonds to the C4 carbonyl of thymidine, which may prevent cytosine from binding in the pocket and explains the ability of some RmlA homologs to use UTP in addition to dTTP. We found that RfbA was capable of using both dTTP and UTP, but not ATP, GTP or CTP (see Supplementary Table S1 and accompanying text for further details).

3.3. The *rfbA* and *rmlA* genes encode different proteins in some organisms

RfbA is a 245-amino-acid protein, which is approximately 50 amino acids shorter than all known RmlA homologs, including that from the Gram-positive bacterium *Aneurini-bacillus thermoaerophilus* (PDB entry 4ho4; T. J. Chen, W. T. Chien, C. C. Lin & W. C. Wang, unpublished work; Fig. 2a). With the exception of these 50 residues, the majority of the RfbA sequence, including the active-site residues, is especially well conserved amongst RmlA homologs (Fig. 2a). Bacterial homologs were chosen for the multiple sequence alignment based on the similarity in sequence and tertiary structure of the monomers. RfbA is present primarily in *Bacillus* and *Paenibacillus*; however, several other organisms, such as *Nitrospira* and *Methanosarcina*, also contain proteins similar to RfbA rather than RmlA, *i.e.* lacking the ~50 C-terminal amino acids. Several strains of the closely related bacteria *B. cereus* and *B. thuringiensis* have both RfbA and RmlA paralogs in their genomes. For instance, in the *B. cereus* HuA2-9 genome IG_03809 encodes RfbA and is present in the same operon as the remaining genes for the dTDP-L-rhamnose-pathway enzymes (*rfbB*, *rfbC* and *rfbD*), while IG_05050 encodes RmlA and is only accompanied by the *rfbB* gene in its operon.

3.4. RfbA lacks the allosteric site present in RmlA

The absence of these 50 residues in RfbA leads to a lack of three helices at its C-terminus compared with RmlA (PDB entry 1g3l; Blankenfeldt *et al.*, 2000; Fig. 2b), which are known to be part of the dimerization domain and significantly contribute to the formation of an allosteric site located between monomers (Fig. 3d). The C-terminus of RfbA most closely resembles its homolog UDP-glucose pyrophosphorylase (GalU) from *Helicobacter pylori* (PDB entry 3juk; Kim *et al.*, 2010), which catalyzes a similar reaction to RmlA except that it uses UTP to form UDP-Glc rather than dTTP to form dTDP- α -D-glucose. The structure of GalU (PDB entry 2ux8; Aragão *et al.*, 2007) is quite similar to that of RfbA, with the exception of the presence of three additional α -helices in GalU: an α 3 helix between the α 2 helix and the β 5 strand, and an α 10 helix at the C-terminus (Fig. 2c).

RmlA homologs contain conserved allosteric site residues Leu45, Tyr114–Glu120, Gly218–Gly220, Ala251, Glu255, Ile256 and Arg259; however, RfbA shares only one of these conserved residues (Asp114; Fig. 2a) and lacks any significant

conservation of the residues that make contact with the allosteric effector dTDP-L-rhamnose in the RmlA structure from *Pseudomonas aeruginosa* (PDB entry 1g3l). Two major loops of RmlA form part of the allosteric site: Tyr114–Glu120 and Gly218–Gly220. These loops are longer in RmlA than in RfbA. The latter of these two loops in RmlA (Gly218–Gly220) is involved in forming a portion of the allosteric site of the adjacent monomer of the dimer. The peptide backbone of this loop is more extended and the side chain of Arg219 embeds itself deep into the opposite monomer. In contrast, the comparable Ser214–Trp216 loop of RfbA is more compact, with side chains oriented towards the interior of the monomer. The near-complete lack of conserved allosteric site residues combined with the lack of key secondary structure involved in allosteric site formation suggests that the RfbA protein is not allosterically regulated like RmlA.

3.5. The tetrameric assembly of RfbA differs from that of RmlA

The biological assembly of all known RmlA crystal structures is homotetrameric, with a diamond-shaped hole in the center that is created by the arrangement of its monomers (Fig. 3b). The arrangement of RfbA monomers is different from that in RmlA because the specific RfbA monomer interactions cause the α -helices that would otherwise create the diamond shape seen in RmlA to instead be on the outer edge of the tetramer (Fig. 3a). The additional C-terminal helices of RmlA lie on top of the core β -sheet of each monomer, and the structure of the full RmlA tetramer does not align well with the tetramer of RfbA (Fig. 3d). Instead, the RfbA tetrameric assembly most closely resembles that of GalU, where the additional helices of GalU lie within the center and at the edges of the tetramer (Fig. 3c).

4. Conclusions

Since the tetrameric arrangement of RfbA is different from that of RmlA, the interactions needed to form the allosteric site between monomers are not present and may therefore indicate that RfbA is not allosterically regulated by dTDP-L-rhamnose. This conclusion is further supported by the fact that RfbA lacks a significant portion of the residues involved in forming the allosteric site and that the structure of RfbA more closely resembles that of GalU, which lacks allosteric regulation. However, further studies are needed to address this question and to determine the metabolic rationale for, or the evolutionary regulatory implications of, the observation that some organisms contain genes encoding either RfbA, RmlA or both in their genomes.

5. Related literature

The following references are cited in the Supporting Information for this article: Fusari *et al.* (2006).

Acknowledgements

This project was funded in part with Startup Funds from San Francisco State University (to MLK).

Funding information

Funding for this research was provided by: National Institute of Allergy and Infectious Diseases (contract Nos. HHSN272200700058C and HHSN272201200026C to Wayne F. Anderson).

References

- Aragão, D., Fialho, A. M., Margues, A. R., Mitchell, E. P., Sá-Correia, I. & Frazão, C. (2007). *J. Bacteriol.* **189**, 4520–4528.
- Barton, W. A., Lesniak, J., Biggins, J. B., Jeffrey, P. D., Jiang, J., Rajashankar, K. R., Thorson, J. S. & Nikolov, D. B. (2001). *Nature Struct. Biol.* **8**, 545–551.
- Beek, S. L. van der, Le Breton, Y., Ferenbach, A. T., Chapman, R. N., van Aalten, D. M., Navratilova, I., Boons, G. J., McIver, K. S., van Sorge, N. M. & Dorfmüller, H. C. (2015). *Mol. Microbiol.* **98**, 946–962.
- Blankenfeldt, W., Asuncion, M., Lam, J. & Naismith, J. H. (2000). *EMBO J.* **19**, 6652–6663.
- 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.* **D66**, 12–21.
- Doran, T. I. & Mattingly, S. J. (1982). *Infect. Immun.* **36**, 1115–1122.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Fusari, C., Demonte, A. M., Figueroa, C. M., Aleanzi, M. & Iglesias, A. A. (2006). *Anal. Biochem.* **352**, 145–147.
- Giraud, M.-F. & Naismith, J. H. (2000). *Curr. Opin. Struct. Biol.* **10**, 687–696.
- Glaser, L. & Kornfeld, S. (1961). *J. Biol. Chem.* **236**, 1795–1799.
- Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J. & Lopez, R. (2010). *Nucleic Acids Res.* **38**, W695–W699.
- Hendricks, K. A., Wright, M. E., Shadomy, S. V., Bradley, J. S., Morrow, M. G., Pavia, A. T., Rubinstein, E., Holty, J. C., Messonnier, N. E., Smith, T. L., Pesik, N., Treadwell, T. A. & Bower, W. A. (2014). *Emerg. Infect. Dis.* **20**, e130687.
- Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
- Jakeman, D., Young, J. L., Huestis, M. P., Peltier, P., Daniellou, R., Nugier-Chauvin, C. & Ferrières, V. (2008). *Biochemistry*, **47**, 8719–8725.
- Kim, H., Choi, J., Kim, T., Lokanath, N. K., Ha, S. C., Suh, S. W., Hwang, H.-Y. & Kim, K. K. (2010). *Mol. Cells*, **29**, 397–405.
- Kuhn, M. L., Majorek, K. A., Minor, W. & Anderson, W. F. (2013). *Protein Sci.* **22**, 222–230.
- Kwon, K. & Peterson, S. N. (2014). *Methods Mol. Biol.* **1140**, 61–74.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
- 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.
- Melo, A. & Glaser, L. (1965). *J. Biol. Chem.* **240**, 398–405.
- Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta Cryst.* **D62**, 859–866.
- Mistou, M. Y., Sutcliffe, I. C. & van Sorge, N. M. (2016). *FEMS Microbiol. Rev.* **40**, 464–479.
- Moretti, R., Chang, A., Peltier-Pain, P., Bingman, C. A., Phillips, G. N. & Thorson, J. S. (2011). *J. Biol. Chem.* **286**, 13235–13243.
- Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). *Methods Enzymol.* **374**, 229–244.
- Painter, J. & Merritt, E. A. (2006). *Acta Cryst.* **D62**, 439–450.
- Sivaraman, J., Sauvé, V., Matte, A. & Cygler, M. (2002). *J. Biol. Chem.* **277**, 44214–44219.
- Sorge, N. M. van et al. (2014). *Cell Host Microbe*, **15**, 729–740.
- Sternbach, G. (2003). *J. Emerg. Med.* **24**, 463–467.
- Wagner, M., Wagner, B. & Rýc, M. (1978). *J. Gen. Microbiol.* **108**, 283–294.
- Winn, M. D. et al. (2011). *Acta Cryst.* **D67**, 235–242.
- Xavier, R. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.