Structure of the *Bacillus anthracis* dTDP-L-rhamnose-biosynthetic enzyme dTDP-4-dehydrorhamnose 3,5-epimerase (RfbC)

Aleksander Shornikov,^{a,†} Ha Tran,^{a,†} Jennifer Macias,^{a,†} Andrei S. Halavaty,^{b,c} George Minasov,^{b,c} Wayne F. Anderson^{b,c} and Misty L. Kuhn^{a*}

Received 8 August 2017

Accepted 31 October 2017

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

† These authors contributed equally to this project.

Keywords: RfbC; dTDP-4-dehydrorhamnose 3,5-epimerase; *Bacillus anthracis*; Anthrax.

PDB reference: RfbC, 3ryk

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

The exosporium layer of *Bacillus anthracis* spores is rich in L-rhamnose, a common bacterial cell-wall component, which often contributes to the virulence of pathogens by increasing their adherence and immune evasion. The biosynthetic pathway used to form the activated L-rhamnose donor dTDP-L-rhamnose consists of four enzymes (RfbA, RfbB, RfbC and RfbD) and is an attractive drug target because there are no homologs in mammals. It was found that co-purifying and screening RfbC (dTDP-6-deoxy-D-xylo-4-hexulose 3,5-epimerase) from *B. anthracis* in the presence of the other three *B. anthracis* enzymes of the biosynthetic pathway yielded crystals that were suitable for data collection. RfbC crystallized as a dimer and its structure was determined at 1.63 Å resolution. Two different ligands were bound in the protein structure: pyrophosphate in the active site of one monomer and dTDP in the other monomer. A structural comparison with RfbC homologs showed that the key active-site residues are conserved across kingdoms.

1. Introduction

Bacillus anthracis is a Gram-positive, endospore-forming bacterium and is the causative agent of Anthrax, a potentially lethal disease. Infection occurs when *B. anthracis* spores enter through cutaneous openings, are ingested or are inhaled (Bozue *et al.*, 2005). L-Rhamnose (6-deoxy-L-mannose) plays an important role in the cell-wall structure of many bacterial species. It has been found to contribute to the virulence of a number of species, including the Gram-negative *Salmonella enterica* and *Vibrio cholerae*, where it is present as a part of the O-antigen, and to be essential for the growth of Gram-positive bacteria such as *Streptococcus pyogenes* (Mistou *et al.*, 2016; Giraud *et al.*, 2000; Allard *et al.*, 2001). In *B. anthracis*, L-rhamnose is not found in the cell wall of the vegetative cells, but is a major component of the spore exosporium layer (Fox *et al.*, 2003). In the closely related *Bacillus cereus*, L-rhamnose is the second most abundant sugar, constituting 6.4% of the exosporium (Daubenspeck *et al.*, 2004; Stewart, 2015; Maes *et al.*, 2016; Matz *et al.*, 1970). L-Rhamnose-deficient strains of *B. anthracis* exhibit reduced host adherence, but this deficiency does not translate into reduced virulence (Chitlaru *et al.*, 2011; Bozue *et al.*, 2005).

L-Rhamnose is biosynthesized in many bacterial and archaeal species as the activated precursor dTDP-L-rhamnose from α -D-glucose-1-phosphate and dTTP in a four-step pathway. dTDP-6-deoxy-D-xylo-4-hexulose 3,5-epimerase (RfbC or RmlC) catalyzes the third reaction: a double epimerization producing dTDP-6-deoxy-L-lyxo-4-hexulose

Table 1

Macromolecule-production information.

RfbA, RfbB, RfbC and RfbD are members of the dTDP-L-rhamnose-biosynthetic pathway and encode a glucose-1-phosphate thymidyltransferase, a dTDP-glucose 4,6-dehydratase, a dTDP-4-dehydrorhamnose 3,5-epimerase and a dTDP-4-dehydro- β -L-rhamnose reductase, respectively. All four proteins were purified together and this combination of proteins was then used for crystallization of RfbC.

	RfbA	RfbB	RfbC	RfbD
Source organism	<i>B. anthracis</i> strain Ames	<i>B. anthracis</i> strain Ames	<i>B. anthracis</i> strain Ames	<i>B. anthracis</i> strain Ames
Cloning/expression vector	pMCSG7	pMCSG19c	pMCSG19c	pMCSG7
Expression host	<i>E. coli</i> BL21(DE3) Magic cells	<i>E. coli</i> BL21(DE3) Magic cells	<i>E. coli</i> BL21(DE3) Magic cells	<i>E. coli</i> BL21(DE3) Magic cells
Complete amino-acid sequence of the construct produced	MHHHHHSSGVDLGTENLYFQS NAMKGIILAGGTGSRLYPITK VTNKHLLPVGRYPMIYHAVYK LKQCDITDIMIITGKEHMGDV VSFLGSGQEFVGSFTYRVQDK AGGIAQALGLCEDFVGNDRMV VILGDNIFSDDIRPYVEEFTN QKEGAKVLLQSVDDPERFGVA NIQNRKIIEIEEKPKPKSSY AVTGIYLYDSKVFSYIKELKP SARGELEITDINNWLKRGVL TYNEMSGWWTDAAGTHVSLQRA NALARDINFGKQFNGE	MHHHHHSSGVDLGTENLYFQS NAMNIVLTGGAGFIGSNFVHY MLQSYETYKIINFDAITYSGN LNNVKSIOQHPNYYFVKGEIQ NGELLEHVIKERDVQIVNFA AESHVDRSIEENPIPFYDTNVI GTVTLLELVKKYPHKLKVQS TDEVYGSGLGKTGRFTEETPLA PNSPYSSSKASADMIALAYLA TYQLPVIIVTRCSNNYGPYQYP EKLIPLMVTNALEGKKLPLYG DGLNVRDHLVHTDHCSAIDVV LHKGRVGEVYNIGGNEKNTV EVVEQIITLLGKTKKIDIEYVT DRLGHDRRYAINAEKMKNEFD WEPKYTFEQGLQETVQWYEKN EEWWKPLKK	MHHHHHSSGVDLGTENLYFQS NAMKVIETNFTDAKLLEPRLF GDDRGFFTESYNKKVLETGLV THSFVQDNVSYSAAEAGTIRGL HFQKNPKAQTKLIQVMQGAII DVIVDLRKDSPTFKQWRGYIL SADNHRQLLVKGFAGHGFCTL VPHTIVMYKVDEYYSADHDSG VLWNDKELAIWPVTSPILSD KDRILPLLQECEDSF	MHHHHHSSGVDLGTENLYFQS NAMKERVITGANGQLGKQLQ EELNPEEYDIYPFDKLLDIT NISQVQQVQVEIRPHIIHCA AYTKVDQAEKERDLAYVINAI GARNVAVASQLVGAALVYIST DYVFQGDREPEGYDEFHNPAPI NIYGASKYAGEQFVKELHNKY FIVRTSWLYGKYGNFVKMTI RLGKEREIEISVVDQIGSPTY VADLNVMINKLIHTSLYGYTH VSNITGSCSWFEFAKKIFSYAN MKVNVLPVSTEEFGAAARPK YSIFQHNMLRLNGFLQMPSE EGLERFFIETKSH

(dTDP-4-dehydro-L-rhamnose). RmlC-like epimerases belong to the diverse cupin superfamily, which includes phosphomannose isomerases, transcription factors and storage proteins (Dunwell *et al.*, 2001). Along with the related 3,4-ketoisomerases, they are responsible for much of the diversity of the final products in various biosynthetic pathways that share the first two steps of the dTDP-L-rhamnose-biosynthetic pathway (Giraud *et al.*, 2000; Thoden & Holden, 2014; Salinger *et al.*, 2015; Dong *et al.*, 2003; Kubiak *et al.*, 2012). Here, we report the crystal structure of RfbC from *B. anthracis* determined at 1.63 Å resolution in complex with dTDP and pyrophosphate. We compared it with homologous structures of other 3,5-epimerases as well as with 3- and 5-mono-epimerases.

2. Materials and methods

2.1. Protein expression and purification

All four proteins of the dTDP-L-rhamnose-biosynthetic pathway (RfbA, RfbB, RfbC and RfbD) from *B. anthracis* strain Ames were cloned into the ampicillin-resistant vectors outlined in Table 1 using previously described procedures (Kwon & Peterson, 2014). All constructs contained a polyhistidine tag followed by a TEV protease cleavage site. Each clone was transformed independently into *Escherichia coli* BL21(DE3) Magic cells (kanamycin-resistant; Dieckman *et al.*, 2002). Starter cultures of each transformant (four in total) were grown in LB overnight at 310 K and were used the next day to inoculate 0.25 l cultures of Terrific Broth (Product No. BP2468500, Fisher BioReagents). Each of the four individual cultures was grown until the OD_{600 nm} reached 0.6, they were incubated on ice, and 0.5 mM IPTG was then added to induce protein expression with shaking overnight at 298 K. The following day, 250 ml of each of the individual cultures containing the RfbA, RfbB, RfbC or RfbD protein were

combined (all four together), centrifuged and the pellet was resuspended in 60 ml lysis buffer (10 mM Tris HCl pH 8.3, 500 mM NaCl, 10% glycerol, 0.01% IGEPAL CA630, 5 mM β -mercaptoethanol). The sample was sonicated and centrifuged as described previously (Kuhn *et al.*, 2013), and the crude extract (supernatant) containing all four proteins was loaded onto a 5 ml Ni-NTA affinity column. The mixture of proteins was purified and concentrated as described previously (Kuhn *et al.*, 2013). The polyhistidine tags were not removed from the proteins. We attempted to crystallize the RfbC protein after purifying it in the absence of the other proteins of the pathway, but did not obtain crystals. The yield of each protein (RfbA, RfbB, RfbC and RfbD) when they were purified separately was 22, 45, 25 and 43 mg l⁻¹, respectively. Thus, the relative proportions of each protein in the mixture used for crystallization in this study were approximately 5.5 mg RfbA:11.25 mg RfbB:6.25 mg RfbC:10.75 mg RfbD.

2.2. Crystallization

We thoroughly screened the RfbC protein alone for crystallization but did not obtain any crystals. Since RfbC is a member of the dTDP-L-rhamnose-biosynthetic pathway and we had clones for the other members of the pathway, we co-crystallized RfbC in the presence of the other three proteins. The mixture containing all four proteins (RfbA, RfbB, RfbC and RfbD) that were purified together (total concentration of 7.5 mg ml⁻¹) in 10 mM Tris-HCl pH 8.3, 500 mM NaCl, 5 mM β -mercaptoethanol (BME), 1 mM dTTP, 1 mM Glc1P, 1 mM NAD⁺, 1 mM NADP⁺, 1 mM MgCl₂ was set up for crystallization in 96-well microplates using the sitting-drop vapor-diffusion method with a ratio of 1 μ l protein solution to 1 μ l reservoir solution. The RfbC protein crystallized at room temperature in 0.2 M diammonium citrate pH 5.0, 20% (w/v)

Table 2
Crystallization.

Method	Vapor diffusion, sitting drop
Plate type	96-well microplate
Temperature (K)	295
Protein concentration (mg ml ⁻¹)	7.5 (all four proteins together)
Buffer composition of protein solution	10 mM Tris-HCl pH 8.3, 500 mM NaCl, 5 mM BME, 1 mM dTTP, 1 mM Glc1P, 1 mM NAD ⁺ , 1 mM NADP ⁺ , 1 mM MgCl ₂
Composition of reservoir solution	The JCSG+ Suite condition A3: 0.2 M diammonium citrate pH 5.0, 20%(w/v) PEG 3350
Volume and ratio of drop	1 µl, 1:1(v:v) ratio
Volume of reservoir (µl)	100

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	APS beamline 21-ID-D
Wavelength (Å)	0.97919
Temperature (K)	100
Detector	MAR Mosaic 300 mm CCD
Space group	<i>P</i> 3
<i>a</i> , <i>b</i> , <i>c</i> (Å)	86.65, 86.65, 45.10
α , β , γ (°)	90.00, 90.00, 120.00
Resolution range (Å)	30.00–1.63 (1.66–1.63)
No. of unique reflections	47039 (2273)
Completeness (%)	99.9 (98.5)
Multiplicity	4.8 (4.4)
$\langle I/\sigma(I) \rangle$	26.5 (2.9)
$R_{\text{r.i.m.}}$ [†]	0.033 (0.315)
Overall <i>B</i> factor from Wilson plot (Å ²)	24.6

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

PEG 3350 (condition A3 of The JCSG+ Suite from Qiagen; Table 2).

2.3. Data collection and processing

Crystals were transferred to reservoir solution and considered to be cryoready prior to flash-cooling in liquid nitrogen. X-ray data were collected on the LS-CAT 21-ID-D beamline at the Advanced Photon Source (APS) at Argonne National Laboratory. The data were indexed, scaled and integrated with *HKL-2000* (Otwinowski & Minor, 1997). Data-collection and processing statistics are summarized in Table 3.

2.4. Structure solution and refinement

The structure of RfbC was determined in complex with dTDP and pyrophosphate (PDB entry 3ryk) by molecular replacement using one monomer from PDB entry 1dzt (48% sequence identity; Giraud *et al.*, 2000) as the search model in *Phaser* (McCoy *et al.*, 2007) from the *CCP4* package (Winn *et al.*, 2011). Since we used a mixture of proteins during crystallization, the protein in the crystal was identified prior to structure solution by first calculating Matthews coefficients for each protein and then determining the crystal parameters and resolution. The molecular-replacement solutions were tested based on the Matthews coefficients for each protein, which revealed that the protein that crystallized was RfbC. No density for the other proteins was observed. The initial

Table 4
Structure refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	28.85–1.63 (1.67–1.63)
Completeness (%)	99.8 (96.7)
No. of reflections, working set	44611 (3149)
No. of reflections, test set	2425 (201)
Final R_{cryst}	0.123 (0.153)
Final R_{free}	0.160 (0.253)
No. of non-H atoms	
Protein	2814
Ligand	34
Solvent	266
Total	3114
R.m.s. deviations	
Bonds (Å)	0.010
Angles (°)	1.852
Average <i>B</i> factors (Å ²)	
Protein	26.25
Ligand	
TYD	29.80
POP	36.40
Water	43.27
Ramachandran plot [†]	
Favored regions (%)	90.3
Additionally allowed (%)	8.4
Outliers (%)	0.6

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

structure was rebuilt with *ARP/wARP* (Morris *et al.*, 2003), modified manually using *Coot* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010) and refined using *REFMAC* v.5.5 (Murshudov *et al.*, 2011). In the final stages of refinement, translation–libration–screw (TLS) groups were obtained from the *TLS Motion Determination (TLSMD)* server (Painter & Merritt, 2006; <http://skuld.bmsc.washington.edu/~tlsmd/>). The RfbC structure was validated using the PDB validation server (<http://deposit.pdb.org/validate/>) and *MolProbity* (Chen *et al.*, 2010; Davis *et al.*, 2007; <http://molprobity.biochem.duke.edu/>). The statistics for structure refinement are presented in Table 4. The sequence identities between RfbC and RfbA, RfbB and RfbD are 8.7, 9.2 and 11.8%, respectively.

3. Results and discussion

Since the three-dimensional structures of all four dTDP-L-rhamnose-biosynthetic pathway enzymes from a single Gram-positive bacterium have not previously been determined, we sought to crystallize them from *B. anthracis*. Despite screening each protein separately for crystals, we found that it was necessary to co-purify and screen all four proteins together to obtain crystals of the RfbC protein alone.

3.1. Crystal structure of RfbC

The RfbC crystals diffracted to 1.63 Å resolution and the structure was determined in space group *P*3 (PDB entry 3ryk). It is a member of the diverse cupin superfamily that is characterized by a β -sandwich or ‘jelly-roll’-like fold. The asymmetric unit holds a single dimer and each monomer contains the first 174 of 181 residues; the remaining residues are disordered. An alanine residue from the affinity tag is

observed just prior to the N-terminal methionine residue. β -Strands are the prevalent secondary-structural element, with each monomer containing 13 β -strands and just two α -helices (Figs. 1*a* and 1*b*). The β -strands are arranged in an antiparallel manner forming two sheets of five and eight β -strands, which together form a flattened, sandwich-like β -barrel (Fig. 1*c*). The five-stranded β -sheet is formed by β 9, β 8, β 11, β 6 and β 13, and the eight-stranded β -sheet is comprised of β 1, β 2, β 10, β 7, β 12 and β 5 of one monomer and β 4 and β 3 of a second monomer (Fig. 1). This domain swapping between monomers further stabilizes the dimer. The active sites are located within cavities formed by the β 5– β 6 end of the barrel structure of each monomer (Fig. 1*c*).

3.2. Homologs of RfbC

DALI was used to identify structurally similar proteins in the Protein Data Bank (PDB) with chain A of PDB entry 3ryk

as the query (Holm & Rosenström, 2010). The results showed 27 structures with Z-scores above 20 that represented 11 bacterial and archaeal species and are clustered by similarity in an unrooted tree (Fig. 2*a*). RmlC di-epimerases from *Pseudomonas aeruginosa*, *Salmonella typhimurium* and archaeal enzymes form one branch. Mono- and di-epimerases from Actinobacteria (*Mycobacterium tuberculosis*, *Streptomyces spheroides*, *Streptomyces bikiniensis* and *Amylocatopsis orientalis*) form a second clade. Finally, *Streptococcus suis* RmlC di-epimerase structures form a distinct third branch. We selected five structures situated in different clusters to compare with our RfbC structure: three 3,5-epimerases [PDB entries 1dzt (Giraud *et al.*, 2000), 2ixk and 2ixl (Dong *et al.*, 2007)], one 3-epimerase (PDB entry 4hmz; Kubiak *et al.*, 2012) and one 5-epimerase (PDB entry 1oi6; Merkel *et al.*, 2004) (Fig. 2). The sequence identities of these five homologs compared with RfbC range from 50% for the *P. aeruginosa*

Figure 1

Three-dimensional crystal structure of *B. anthracis* RfbC. (a) Cartoon representation of a monomer of RfbC colored from blue at the N-terminus to red at the C-terminus. (b) Topology diagram of the RfbC monomer. Helices are shown as cylinders and β -strands are shown as arrows. Secondary structure is colored in the same manner as in (a). (c) RfbC dimer colored in blue and orange for each monomer. dTDP is shown as sticks in the active site of the blue monomer and PP_i is shown as sticks in the active site of the orange monomer. (a) and (c) were created using PyMOL (Schrödinger) and (b) was created using Pro-origami (Stivala *et al.*, 2011) and Inkscape (<https://inkscape.org/en/>).

and *S. typhimurium* RmlC di-epimerases (PDB entries 2ixk and 1dzt) to 22% for the *A. orientalis* mono-epimerase (PDB entry 1oi6). Overall, 21 residues (11.6%) were conserved across the selected homologs (Fig. 2*b*), and highly conserved structural regions included the secondary structures around the active site (the $\beta 6$ strand, $\beta 3$ – $\beta 4$ turn and $\alpha 2$ helix). Other secondary-structural elements located away from the active site were also conserved, notably $\beta 8$ and the ten-residue loop connecting it to strand $\beta 9$. The $\beta 8$ – $\beta 9$ strands form the center of the Greek-key motif of the ‘jelly-roll’ fold.

3.3. Active site

Three strictly conserved catalytic residues are essential for the function of RmlC: His62, Lys72 and Tyr132 (Dong *et al.*, 2003, 2007; Figs. 2*b* and 3*a*). The reaction involves two sequential proton-abstraction/reprotonation steps: first at C5 and then at C3. Asp168 activates His62 in a manner similar to the catalytic mechanism of a serine protease, and the pK_a of the α -carbon is lowered by the stabilization effect of Lys72 on the resulting enolate anion. Tyr132 acts as an acid in the reaction by reprotonating the enolate from the opposite side of the ring (Dong *et al.*, 2007). All three of these residues reside within one monomer of the dimer, but residues from both monomers from the dimer are required to bind the substrate and product (Fig. 3). Thus, RfbC is functional as a dimer.

The RfbC protein structure showed dTDP bound in the active site of one monomer and pyrophosphate (PP_i) in the active site of a second monomer of the dimer; however, the aforementioned catalytic residues do not interact directly with these ligands and the orientations of the phosphates on dTDP are rotated away from the active site (Fig. 3). The nucleobase of dTDP bound in one monomer is stacked between the aromatic rings of Tyr138 and Phe26* (where the asterisk refers to a residue of the swapped $\beta 3$ – $\beta 4$ hairpin monomer), and is stabilized by a network of hydrogen bonds through water molecules and the Glu28*, Gln46 and Asn48 residues. The Ser165, Asp168 and Lys167 residues form hydrogen bonds directly to the phosphates of dTDP, and Arg59 and Asp21* form hydrogen bonds to the phosphates through water molecules. Since the phosphates of dTDP face outwards, the residues located on the $\beta 3$ – $\beta 4$ loop assume different conformations from the structures of homologs with a full substrate analog bound, and the loop is shifted by as much as 2.4 Å. Arg59 and Arg23* are strictly conserved in RmlC-like epimerases and typically form salt bridges with the phosphates of the substrate. Owing to the improper conformation of the dTDP compared with the native substrate, Arg23* does not make contact with the phosphate in the RfbC structure (Fig. 3*c*). Similar interactions are seen in the monomer containing only PP_i .

3.4. Alternative conformations of key residues

The RfbC protein (PDB entry 3ryk) was co-crystallized with ligands that lack the pyranosyl moiety of the substrate. As a result, some residues in the active site assumed alternative

conformations (Figs. 3*a* and 3*c*). For example, His62 in the RfbC structure is observed in two alternative conformations: one matches the conformation observed in structures in complex with sugars and the other is rotated almost 120° around the C^α – C^β bond. Additionally, Lys167 of RfbC appears in two alternative conformations: one within a hydrogen-bond distance of the α -phosphate O atom and the other rotated about 90° around the C^α – C^β bond towards the bulk solvent (Fig. 3). It is not apparent from the multiple sequence alignment that this lysine residue is conserved across di- and mono-epimerases, as it is replaced by arginine in the sequences of PDB entries 1oi6 and 4hmz and by alanine in that of PDB entry 2ixl in the alignment (Fig. 2*b*). Upon further examination of the three-dimensional structures of these proteins, however, Lys167 of RfbC and Lys151 of PDB entry 2ixl are positioned similarly and form a hydrogen bond to the phosphate group of the substrate. A key difference between mono- and di-epimerases thus is that mono-epimerases have an Arg–Glu acid–base pair and di-epimerases have a Lys–Glu/Asp acid–base pair. One exception to this is the RmlC di-epimerase from *M. tuberculosis*, which has an Arg–Glu acid–base pair.

3.5. Source of ligands during co-crystallization

Since dTDP and PP_i were not added during co-crystallization we are not certain of their source, but we suspect two likely scenarios. (i) PP_i could have been produced from the RfbA enzymatic reaction since dTTP, glucose 1-phosphate and RfbA were added during crystallization. (ii) Alternatively, dTTP could have been hydrolyzed to produce dTDP, and PP_i is a known contaminant of nucleotides. Two groups have previously reported unexpected dTDP electron density in the active sites of RfbC homologs. For instance, the *P. aeruginosa* RmlC protein (PDB entry 2ixl) was crystallized with dTDP-xylose; however, only electron density corresponding to dTDP was observed. Evidence suggested that the sugar moiety was located outside the binding site and was therefore disordered (Dong *et al.*, 2007). In another study, dTDP was observed at the active site of the 3,4-ketoisomerase QdtA from *Thermoanaerobacterium thermosaccharolyticum* (PDB entry 4zu7; Thoden *et al.*, 2015), which has the same fold and active-site location as RmlC-like epimerases. Its presence was attributed to binding during protein expression and purification (Thoden *et al.*, 2015). Examination of the electron-density map of the RfbC structure does not show additional density beyond the β -phosphate group of dTDP bound in one monomer, and no additional density is seen surrounding PP_i in the other monomer.

3.6. Interactions between Rfb proteins

Sequential enzymes of some metabolic pathways have been shown to form protein complexes (metabolons), or exist as bifunctional or multifunctional fused enzymes in different organisms (Srere, 1987; Krüger *et al.*, 2012). As a result, we proceeded to screen various combinations of proteins of the L-rhamnose-biosynthetic pathway from *B. anthracis* for

crystallization when our efforts to crystallize the single *B. anthracis* RfbC protein alone failed. To our knowledge there is currently no evidence for multi-protein complex formation for members of this pathway; however, the fact that we obtained crystals of the RfbC protein only when all four

proteins were added together for crystallization suggests the possibility that the other proteins stabilized RfbC for crystallization and a multi-protein complex may exist. We used a similar method to obtain crystals of the *B. anthracis* protein RfbB (PDB entry 6bi4). Examples of bifunctional enzymes

Figure 3

RfbC active-site and dimer formation. (a) Active site of RfbC (PDB entry 3ryk) showing three strictly conserved catalytic residues and alternative conformations of His62 and Lys72. One monomer of the dimer is shown in orange and the other is shown in blue. dTDP is shown as white sticks, while dTDP-L-rhamnose is shown as purple sticks and is modeled from the structure with PDB code 2ixl. (b) C5' epimerization reaction adapted from Dong *et al.* (2007). Residues that participate in the reaction are numbered according to the *B. anthracis* RfbC protein sequence. (c) Comparison of the structures with PDB codes 3ryk and 2ixl. The 3ryk dimer is shown as blue and orange ribbons and the 2ixl dimer is shown as purple and yellow ribbons. 3ryk contains dTDP and PP_i (shown as sticks) and 2ixl shows dTDP-L-rhamnose as sticks. Each structure shows an enlarged view of the key conserved residues in homologous RmlC enzymes. Alternative conformations are seen for His62 and Lys167. The protein structures in (a) and (c) were created using PyMOL, and ChemDraw was used to create the scheme shown in (b).

from other nucleotide-activated biosynthetic pathways that produce different sugars to dTDP-L-rhamnose include the RHM2/MUM4 protein from *Arabidopsis thaliana*, which produces UDP-L-rhamnose (Watt *et al.*, 2004), and the CapF protein from *Staphylococcus aureus*, which produces UDP-N-acetyl-L-fucosamine (Miyafusa *et al.*, 2012). Therefore, the strategy of combining enzymes that are members of the same metabolic pathway together for crystallization when they are unable to be crystallized alone may extend the crystallographer's toolbox and provide new insight into potential protein complexes.

4. Conclusions

Overall, the RfbC structure from *B. anthracis* is quite similar to those of other homologs. The approach of co-crystallizing members of a metabolic pathway together may be a good salvage strategy to obtain crystals of proteins that fail common crystallization trials.

Funding information

This project was funded in part with Startup Funds from San Francisco State University (to MLK). Additional funding for this project included Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH), US Department of Health and Human Services under Contract Nos. HHSN272200700058C and HHSN272201200026C to WFA.

References

- Allard, S. T. M., Giraud, M.-F. & Naismith, J. H. (2001). *Cell. Mol. Life Sci.* **58**, 1650–1665.
- Bozue, J. A., Parthasarathy, N., Phillips, L. R., Cote, C. K., Fellows, P. F. F., Mendelson, I., Shafferman, A. & Friedlander, A. M. (2005). *Microb. Pathog.* **38**, 1–12.
- 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.
- Chitlaru, T., Altboum, Z., Reuveny, S. & Shafferman, A. (2011). *Immunol. Rev.* **239**, 221–236.
- Daubenspeck, J. M., Zeng, H., Chen, P., Dong, S., Steichen, C. T., Krishna, N. R., Pritchard, D. G. & Turnbough, C. L. Jr (2004). *J. Biol. Chem.* **279**, 30945–30953.
- Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., Snoeyink, J., Richardson, J. S. & Richardson, J. S. (2007). *Nucleic Acids Res.* **35**, W375–W383.
- Dieckman, L., Gu, M., Stols, L., Donnelly, M. I. & Collart, F. R. (2002). *Protein Expr. Purif.* **25**, 1–7.
- Dong, C., Major, L. L., Allen, A., Blankenfeldt, W., Maskell, D. & Naismith, J. H. (2003). *Structure*, **11**, 715–723.
- Dong, C., Major, L. L., Srikanthasani, V., Errey, J. C., Giraud, M.-F., Lam, J. S., Graninger, M., Messner, P., McNeil, M. R., Field, R. A., Whitfield, C. & Naismith, J. H. (2007). *J. Mol. Biol.* **365**, 146–159.
- Dunwell, J. M., Culham, A., Carter, C. E., Sosa-Aguirre, C. R. & Goodenough, P. W. (2001). *Trends Biochem. Sci.* **26**, 740–746.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst. D* **60**, 2126–2132.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst. D* **66**, 486–501.
- Fox, A., Stewart, G. C., Waller, L. N., Fox, K. F., Harley, W. M. & Price, R. L. (2003). *J. Microbiol. Methods*, **54**, 143–152.
- Giraud, M.-F., Leonard, G. A., Field, R. A., Berlind, C. & Naismith, J. H. (2000). *Nature Struct. Mol. Biol.* **7**, 398–402.
- Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J. & Lopez, R. (2010). *Nucleic Acids Res.* **38**, W695–W699.
- He, Z., Zhang, H., Gao, S., Lercher, M. J., Chen, W.-H. & Hu, S. (2016). *Nucleic Acids Res.* **44**, W236–W241.
- Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
- Krüger, B., Liang, C., Prell, F., Fieselmann, A., Moya, A., Schuster, S., Völker, U. & Dandekar, T. (2012). *Metabolites*, **2**, 940–958.
- Kubiak, R. L., Phillips, R. K., Zmudka, M. W., Ahn, M. R., Maka, E. M., Pyeatt, G. L., Roggensack, S. J. & Holden, H. M. (2012). *Biochemistry*, **51**, 9375–9383.
- 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.
- Maes, E., Krzewinski, F., Garenaux, E., Lequette, Y., Coddeville, B., Trivelli, X., Ronse, A., Faille, C. & Guerardel, Y. (2016). *J. Biol. Chem.* **291**, 9666–9677.
- Matz, L. L., Beaman, T. C. & Gerhardt, P. (1970). *J. Bacteriol.* **101**, 196–201.
- 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.
- Merkel, A. B., Major, L. L., Errey, J. C., Burkart, M. D., Field, R. A., Walsh, C. T. & Naismith, J. H. (2004). *J. Biol. Chem.* **279**, 32684–32691.
- Mistou, M. Y., Sutcliffe, I. C. & van Sorge, N. M. (2016). *FEMS Microbiol. Rev.* **40**, 464–479.
- Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). *Methods Enzymol.* **374**, 229–244.
- Miyafusa, T., Caaveiro, J. M. M., Tanaka, Y. & Tsumoto, K. (2012). *Biochem. J.* **443**, 671–680.
- 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.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Painter, J. & Merritt, E. A. (2006). *Acta Cryst. D* **62**, 439–450.
- Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
- Salinger, A. J., Brown, H. A., Thoden, J. B. & Holden, H. M. (2015). *Protein Sci.* **24**, 1633–1639.
- Srere, P. A. (1987). *Annu. Rev. Biochem.* **56**, 89–124.
- Stewart, G. C. (2015). *Microbiol. Mol. Biol. Rev.* **79**, 437–457.
- Stivala, A., Wybrow, M., Wirth, A., Whisstock, J. & Stuckey, P. (2011). *Bioinformatics*, **27**, 3315–3316.
- Thoden, J. B. & Holden, H. M. (2014). *Protein Sci.* **23**, 683–692.
- Thoden, J. B., Vinogradov, E., Gilbert, M., Salinger, A. J. & Holden, H. M. (2015). *Biochemistry*, **54**, 4495–4506.
- Watt, G., Leoff, C., Harper, A. D. & Bar-Peled, B. (2004). *Plant Physiol.* **134**, 1337–1346.
- Winn, M. D. *et al.* (2011). *Acta Cryst. D* **67**, 235–242.