Journal of Structural Biology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Technical Note

Structure of the *Bacillus anthracis* dTDP-L-rhamnose biosynthetic pathway enzyme: dTDP-α-D-glucose 4,6-dehydratase, RfbB

Trevor Gokey[a], Andrei S. Halavaty[b,c], George Minasov[b,c], Wayne F. Anderson[b,c], Misty L. Kuhn[a,*]

[a] Department of Chemistry and Biochemistry, San Francisco State University, USA
[b] Department of Biochemistry and Molecular Genetics, Northwestern University Feinberg School of Medicine, USA
[c] Center for Structural Genomics of Infectious Diseases (CSGID), USA

ARTICLE INFO

Keywords:
Cell wall polysaccharide producing enzymes
Co-crystallization of metabolic enzymes
Dehydratase
Protein crystallization
Therapeutic drug target

ABSTRACT

Many bacteria require L-rhamnose as a key cell wall component. This sugar is transferred to the cell wall using an activated donor dTDP-L-rhamnose, which is produced by the dTDP-L-rhamnose biosynthetic pathway. We determined the crystal structure of the second enzyme of this pathway dTDP-α-D-glucose 4,6-dehydratase (RfbB) from *Bacillus anthracis*. Interestingly, RfbB only crystallized in the presence of the third enzyme of the pathway RfbC; however, RfbC was not present in the crystal. Our work represents the first complete structural characterization of the four proteins of this pathway in a single Gram-positive bacterium.

1. Background on L-rhamnose and enzymes of the dTDP-L-rhamnose biosynthetic pathway

Numerous bacteria contain L-rhamnose as a structural component of their cell envelopes (Mistou et al., 2016). This sugar is critical in many bacteria because it is incorporated into cell wall polysaccharides in Gram-positive bacteria (Mistou et al., 2016) and lipopolysaccharides in many Gram-negative bacteria (Jiang et al., 1991; Macpherson et al., 1994; Marolda and Valvano, 1995; Rocchetta et al., 1999). The activated donor dTDP-L-rhamnose is used to transfer L-rhamnose to components of the cell envelope and is synthesized using a series of four enzymes that comprise the dTDP-L-rhamnose biosynthetic pathway (Giraud and Naismith, 2000). Genes that encode these enzymes are annotated inconsistently across organisms, as *rml*, *rfb*, or *rff* A, B, C, and D. The effect of preventing incorporation of L-rhamnose into bacterial cell walls varies by species, but it causes severe growth and cell division abnormalities in Gram-positive bacteria like *Streptococcus pyogenes* and *Streptococcus mutans* (Beek et al., 2015). Since this sugar is not found in eukaryotes, targeting enzymes from biosynthetic pathways that produce bacterial cell wall components is one avenue for therapeutic drug development toward bacterial pathogens (Adibekian et al., 2011; Mäki and Renkonen, 2004; Mistou et al., 2016).

B. anthracis is a Gram-positive bacterium that causes the disease Anthrax and is currently classified as a Category A Priority Pathogen by the US National Institute of Allergy and Infectious Diseases due to its ease of transmission and high mortality rate. Throughout history this disease has resulted in the death of significant numbers of humans and livestock around the world, but it is largely seen as a disease of the past for humans, except as a bioterrorist agent or potential re-emerging disease. The dTDP-L-rhamnose biosynthetic enzymes have been widely studied across a variety of bacteria, however, the full pathway of enzymes has not been structurally characterized from a single Gram-positive bacterium. Therefore, we have determined the structures of all four enzymes of the pathway from this important human pathogen (Protein Data Bank (PDB) accession codes: RfbA (4ecm; Baumgartner et al., 2017), RfbB (6bi4), RfbC (3ryk; Shornikov et al., 2017), and RfbD (3sc6; Law et al., 2017)), and will describe the structure of the second enzyme dTDP-α-D-glucose 4,6-dehydratase (RfbB) here. RfbB produces dTDP-4-keto-6-deoxyglucose from dTDP-α-D-glucose using an NAD^+ coenzyme via a series of oxidation, dehydration, and reduction steps. First, NAD^+ oxidizes the glucosyl C4′ of dTDP-glucose. Second, water is eliminated from the glucosyl C5′ and C6′ of the newly formed intermediate. Finally, NADH reduces this intermediate at the glucosyl C6′ position to form the product dTDP-4-keto-6-deoxyglucose (Allard et al., 2002; Gabriel and Lindquist, 1968; Gross et al., 2000; Hegeman et al., 2001). The third enzyme of the dTDP-L-rhamnose pathway (RfbC) can then use this product to proceed with biosynthesis of dTDP-L-rhamnose. Alternatively, dTDP-4-keto-6-deoxyglucose can be used by enzymes of other metabolic pathways since this sugar is a precursor to nearly every known deoxy-sugar biosynthetic pathway (Liu and Thorson, 1994), including polyketide biosynthetic pathways.

* Corresponding author at: San Francisco State University, 1600 Holloway Ave., San Francisco, CA 94132, USA.
E-mail address: mkuhn@sfsu.edu (M.L. Kuhn).

https://doi.org/10.1016/j.jsb.2018.01.006
Received 11 December 2017; Received in revised form 8 January 2018; Accepted 9 January 2018

2. Protein expression and purification

The RfbB and RfbC genes from *Bacillus anthracis* str. Ames were cloned into the pMCSG19c ampicillin resistant vector using previously described procedures (Kwon and Peterson, 2014). The genes in this vector are expressed with an N-terminal maltose binding protein (MBP) tag followed by a tobacco vein mottling virus (TVMV) protease cleavage site, a polyhistidine tag, tobacco etch virus (TEV) protease cleavage site, and the protein of interest. The TVMV protease is constitutively expressed by an integrated gene on the plasmid, which removes the MBP tag prior to purification with Ni-NTA affinity chromatography. Each plasmid was separately transformed into kanamycin resistant *E. coli* BL21 (DE3) Magic cells (Dieckman et al., 2002) for protein expression. The RfbB protein was grown in 1.5 L of M9 medium in the presence of SeMet as described previously (Makowska-Grzyska et al., 2012) and the RfbC protein was grown and expressed separately in 1.5 L of Terrific Broth. Cells were grown to an OD_{600nm} of 0.6–0.8 at 310.15 K and then cultures were cooled on ice. Protein expression was induced using 0.5 mM IPTG at room temperature overnight with shaking. The next day cells from both cultures (RfbB SeMet and RfbC native) were harvested by centrifugation, resuspended in 100 mL of lysis buffer (10 mM Tris HCl pH 8.3, 500 mM NaCl, 10% glycerol, 0.01% Igepal CA630, and 5 mM beta-mercaptoethanol (BME)), and sonicated together. Both RfbB and RfbC proteins were purified at the same time on the same 5 mL Ni-NTA affinity column and concentrated using previously described procedures (Kuhn et al., 2013). The polyhistidine tags were not removed from the proteins. When we previously expressed and purified the proteins separately, the average yield of RfbB and RfbC was approximately 45 mg/L and 25 mg/L, respectively (Shornikov et al., 2017). Therefore, the relative ratio of RfbB:RfbC proteins in the mixture used for crystallization was estimated as approximately 2:1. As a result, the purified protein used for crystallization was not necessarily a stable complex of RfbB and RfbC in equal molar ratios.

3. Crystallization

RfbB and RfbC catalyze two sequential steps of the dTDP-L-rhamnose biosynthetic pathway and in some metabolic pathways sequential genes can produce multi-protein complexes (Marsh et al., 2013). Since we were unable to obtain crystals of the RfbB protein even in the presence of ligands, we screened RfbB for crystals in the presence of RfbC because it is the next enzyme in the biosynthetic pathway and may help stabilize RfbB for crystallization (See results and discussion for full explanation of how we arrived at this approach). The total RfbB and RfbC protein concentration used for crystallization was 7.5 mg/mL in 10 mM Tris HCl pH 8.3, 500 mM NaCl, 5 mM BME, 10% glycerol, 5 mM $MgCl_2$, and 1 mM NAD^+. One μL of protein was added to 1 μL of reservoir solution in a 96-well microplate made for sitting drop vapor diffusion. Crystals of the RfbB protein grew in the condition containing 0.1 M Tris HCl pH 8.0, 1.56 M ammonium sulfate, and 8.5% (w/v) PEG3350.

4. Data collection and processing

Crystals of RfbB were soaked in a cryo-protectant solution, which was made by mixing in a 1:1 ratio 3.6 M ammonium sulfate and 50% sucrose, and flash-cooled in liquid nitrogen. Data were collected on the 21ID-G beamline of the Life Sciences Collaborative Access Team (LS-CAT) at the Advanced Photon Source at Argonne National Laboratory. Data were indexed, scaled, and integrated with HKL-2000 (Otwinowski and Minor, 1997). The data collection and processing statistics are presented in Table 1.

Table 1
Data collection, processing, and structure refinement statistics, and the model quality of RfbB.

Data collection	
Space group	*I*422
a, *b*, *c* (Å)	165.56, 165.56, 292.12
α, *β*, γ (°)	90.00, 90.00, 90.00
Resolution range (Å)	30.00–2.90 (2.95–2.90)
No. of measured reflections	422,805
No. of unique reflections	44,858 (2,208)
Completeness (%)	100.0 (99.9)
Redundancy	9.4 (9.7)
$\langle I/\sigma(I)\rangle$	13.0 (2.9)
R_{merge}	0.15 (0.77)
Overall B factor from Wilson plot (Å^2)	70.3
Refinement	
Resolution range (Å)	29.95–2.91 (2.98–2.91)
Completeness (%)	99.7 (99.6)
No. of reflections, working set	42,594 (3,072)
No. of reflections, test set	2264 (160)
R_{cryst}	0.167 (0.269)
R_{free}	0.204 (0.328)
No. of non-H-atoms	
Protein (No. of molecules)	10,037 (4)
Ion	2 (NI); 70 (SO4)
Ligand	176 (NAD); 46 (SUC)
Water	76
R.m.s. deviations	
Bonds (Å)	0.010
Angles (°)	1.667
Average B factors (Å^2)	
Protein	66.5
Ion	116.8 (NI); 112.3 (SO4)
Ligand	60.6 (NAD); 111.6 (SUC)
Water	59.6
TLS bodies	4 for each of the four chains
Ramachandran plot[†]	
Favored regions (%)	97.0
Allowed regions (%)	3.0
Outliers (%)	0.0

[†] Statistics are based on MolProbity (available on the web (http://molprobity.biochem.duke.edu), Davis et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Research 35: W375-W383). Values in parentheses are for highest-resolution shell.

5. Structure solution and refinement

The RfbB structure was determined using Phaser (McCoy et al., 2007) from the CCP4 package (Winn et al., 2011). The structure with PDB code 1r66 (51% sequence identity to RfbB; Allard et al., 2004) was used as a search model for molecular replacement and the initial model was rebuilt with ARP/wARP (Morris et al., 2003), manually corrected in Coot (Emsley and Cowtan, 2004; Emsley et al., 2010), and refined with REFMAC v5.7 (Murshudov et al., 2011). At this point NAD^+, sucrose, sulfate and a nickel ion (coordination number 4, which is more frequent for nickel; the nickel was likely from purification since we used nickel affinity chromatography) were fitted in the well defined electron density maps and the model was refined for several cycles using Refmac and corrected in Coot. During the final stages of refinement, we applied the translation-libration-screw (TLS) group correction. The TLS groups were defined using the TLSMD server (Painter and Merritt, 2006; http://skuld.bmsc.washington.edu/~tlsmd/). The quality of the final model was assessed 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/). Structure refinement statistics and the model quality are shown in Table 1. Chains A and C are missing residues 86–94, 273, 274, and 322; Chains B and D are missing residues 86–94, 272–274, and 322. Coordinates and supporting experimental data for the RfbB structure have been deposited into the Protein Data Bank under the accession code 6bi4.