

Journal Pre-proof

Spinning sugars in antigen biosynthesis: characterization of the *Coxiella burnetii* and *Streptomyces griseus* TDP-sugar epimerases

Alice R. Cross, Sumita Roy, Mirella Vivoli Vega, Martin Rejzek, Sergey A. Nepogodiev, Matthew Cliff, Debbie Salmon, Michail N. Isupov, Robert A. Field, Joann L. Prior, Nicholas J. Harmer, on behalf of the GoVV consortium

PII: S0021-9258(22)00343-X

DOI: <https://doi.org/10.1016/j.jbc.2022.101903>

Reference: JBC 101903

To appear in: *Journal of Biological Chemistry*

Received Date: 19 December 2021

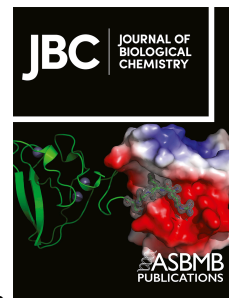
Revised Date: 2 April 2022

Accepted Date: 4 April 2022

Please cite this article as: Cross AR, Roy S, Vivoli Vega M, Rejzek M, Nepogodiev SA, Cliff M, Salmon D, Isupov MN, Field RA, Prior JL, Harmer NJ, on behalf of the GoVV consortium, Spinning sugars in antigen biosynthesis: characterization of the *Coxiella burnetii* and *Streptomyces griseus* TDP-sugar epimerases, *Journal of Biological Chemistry* (2022), doi: <https://doi.org/10.1016/j.jbc.2022.101903>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 THE AUTHORS. Published by Elsevier Inc on behalf of American Society for Biochemistry and Molecular Biology.



Spinning sugars in antigen biosynthesis: characterization of the *Coxiella burnetii* and *Streptomyces griseus* TDP-sugar epimerases

Alice R. Cross^{1,2}, Sumita Roy^{1,2}, Mirella Vivoli Vega^{1,2†}, Martin Rejzek³, Sergey A. Nepogodiev³, Matthew Cliff⁴, Debbie Salmon², Michail N. Isupov², Robert A. Field^{3,4}, Joann L. Prior⁵, Nicholas J. Harmer^{1,2*} on behalf of the GoVV consortium

¹Living Systems Institute, University of Exeter, Stocker Road, Exeter EX4 4QD, U.K.

²Department of Biosciences, University of Exeter, Exeter EX4 4QD, U.K.

³Department of Biological Chemistry, John Innes Centre, Norwich Research Park, Colney Lane, Norwich NR4 7UH, U.K.

⁴Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester M1 7DN, U.K.

⁵Dstl, Porton Down, Salisbury, Wiltshire SP4 0JQ, U.K.

†Present address: School of Biochemistry, University of Bristol, Bristol BS8 1TD, U.K.

*To whom correspondence should be addressed: N.J.Harmer@exeter.ac.uk

Abstract

The sugars streptose and dihydrohydroxystreptose (DHHS) are unique to the bacteria *Streptomyces griseus* and *Coxiella burnetii*, respectively. Streptose forms the central moiety of the antibiotic streptomycin, whilst DHHS is found in the O-antigen of the zoonotic pathogen *C. burnetii*. Biosynthesis of these sugars has been proposed to follow a similar path to that of TDP-rhamnose, catalyzed by the enzymes RmlA, RmlB, RmlC, and RmlD, but the exact mechanism is unclear. Streptose and DHHS biosynthesis unusually requires a ring contraction step that could be performed by orthologues of RmlC or RmlD. Genome sequencing of *S. griseus* and *C. burnetii* has identified StrM and CBU1838 proteins as RmlC orthologues in these respective species. Here, we demonstrate that both enzymes can perform the RmlC 3'',5'' double epimerization activity necessary to support TDP-rhamnose biosynthesis *in vivo*. This is consistent with the ring contraction step being performed on a double epimerized substrate. We further demonstrate that proton exchange is faster at the 3''-position than the 5''-position, in contrast to a previously studied orthologue. We additionally solved the crystal structures of CBU1838 and StrM in complex with TDP, and show that they form an active site highly similar to those of the previously characterized enzymes RmlC, EvaD, and ChmJ. These results support

the hypothesis that streptose and DHHS are biosynthesized using the TDP pathway and that an RmlD paralogue most likely performs ring contraction following double epimerization. This work will support the elucidation of the full pathways for biosynthesis of these unique sugars.

Keywords: Q-fever, epimerase, X-ray crystallography, NMR, GC-MS, enzyme kinetics

Journal Pre-proof

Introduction

The gammaproteobacterium *Coxiella burnetii* evolved recently from tick endosymbionts (1) to become an obligate intracellular pathogen of mammals (2). This bacterium infects a wide range of mammals causing reproductive failures (3,4). It is a significant economic pathogen of sheep, goats and cattle (5), causing spontaneous abortion of pregnancies (6). Humans can be infected by *C. burnetii* either through aerosols or consumption of meat and dairy products (with 1-10 bacteria sufficient to cause disease) (7,8). *C. burnetii* in unpasteurised ruminant milk poses a particular risk (9-11). Humans generally present with a self-limiting febrile illness (12). A small percentage of cases lead to more serious complications such as miscarriage, hepatitis, or endocarditis (2,13,14). *C. burnetii* infection can be effectively treated with doxycycline (2). However, diagnosis is often challenging as *C. burnetii* grows only in highly defined media, requires environment-controlled microaerobic incubators (15), and symptoms of infection are generally non-specific.

From a “One Health” perspective, a vaccine against *C. burnetii* would be the ideal approach to reduce veterinary and human infection (6). Inactivated whole cell vaccines have proved effective in animals and humans (16,17). These are not licensed for human use in the USA, EU or UK as they cause severe reactions in seropositive individuals (18). Current efforts to develop novel vaccines are focused on development of subunit (19) or epitope based vaccines (20,21). A very strong subunit vaccine candidate is the *C. burnetii* polysaccharide O-antigen. A complete O-antigen is required for immune evasion and efficient infection of mammalian cells (22,23). Serial passaging of *C. burnetii* selects for mutants that cease producing some or all their O-antigen (24,25). The most common mutants either lack a polymer of the unusual saccharide virenose (“intermediate”), or the entire O-antigen (26) (Figure 1A). Strains lacking the O-antigen are avirulent in the guinea-pig model of infection (27). The O-antigen is the dominant epitope of wild-type *C. burnetii* (28), suggesting that any vaccine against *C. burnetii* will require the O-antigen (28). This highlights the importance of a deepened understanding of the biosynthesis of this polysaccharide.

The *C. burnetii* O-antigen characteristically contains two C-3 methylated/hydroxymethylated sugars, virenose and dihydrohydroxystreptose (DHHS; Figure 1B) (29,30). Virenose is also found in *Bacillus cereus* and *Streptomyces albaduncus* (31,32). Biosynthetic pathways have been proposed in these two organisms, and a similar pathway has been proposed in *C. burnetii* (33). No equivalent pathway has been proposed for DHHS biosynthesis. The most similar known sugar to DHHS is streptose (Figure 1B), the central saccharide unit of the aminoglycoside antibiotic streptomycin (34), produced by *Streptomyces griseus*. Streptose

differs from DHHS in having no hydroxyl at C-6; and being oxidised at the C-3 hydroxymethyl group. Streptose is synthesised from TDP-glucose, in a manner analogous to the biosynthesis of rhamnose (Figure 1C) (34-37). A reasonable path to DHHS synthesis would involve a C-4 oxidase of TDP-glucose, a 5''-mono or 3'',5''-double epimerase (38), and an enzyme similar to the streptose synthase (37) (Figure 1D,E). Although it remains possible that the epimerase might also perform the ring rearrangement, the only precedent for this reaction was performed by an enzyme from a different structural class (39).

Two operons (*cbu0825* to *cbu0856*, and *cbu1831* to *cbu1838*) have been suggested as coding for DHHS synthesising genes (26). The *cbu1831* to *cbu1838* contains two genes that show high sequence identity to genes in the rhamnose biosynthesis pathway. CBU1834 has over 60% sequence identity to characterised RmlA orthologues (33), and so is highly likely to be a glucose-1-phosphate thymidyltransferase. TDP-glucose has also been proposed to be a precursor of virenose (33), and so is not necessarily characteristic of DHHS biosynthesis. CBU1838 shows 27-53% identity with characterised RmlC orthologues (Table S1). With this sequence identity, this protein is highly likely to be an orthologue of RmlC and act as a TDP-D-xylo-4-hexulose 3,5-epimerase. This activity is not required for virenose biosynthesis. As *C. burnetii* does not produce rhamnose, it is unlikely that it would be required for rhamnose biosynthesis. This leaves DHHS biosynthesis as the only reasonable function of CBU1838.

RmlC is part of a family of epimerases that use a deprotonation/re-protonation mechanism (40). RmlC uses a histidine base to deprotonate the substrate, and a tyrosine acid to re-protonate from the opposite side of the sugar (41-44). The same residues are likely involved in epimerising at both 3'' and 5'' positions: the sugar ring flips between the two reactions to present the second site to be epimerised to the catalytic base and acid (41). These amino acids are highly conserved in the family (Figure 2). RmlC orthologues have been characterised biochemically and structurally from a range of species (Table 1) (38,42,43,45-49). Paralogues that selectively epimerise at either the 3' or 5' carbon have also been characterised (44,50,51). These studies have highlighted likely residues for binding to the TDP carrier (43,52,53). For streptose and DHHS biosynthesis, only epimerisation at the 5'' position is strictly necessary: the C3-C4 bond is broken in the streptose biosynthesis step (35), which will remove the stereocentre at C3. RmlC paralogues for streptose (StrM) and DHHS (CBU1838) biosynthesis have been proposed bioinformatically (33) (e.g. by KEGG (54)); however, such assignments can be inaccurate and so experimental confirmation is desirable (55).

Here, we tested the hypothesis that StrM and CBU1838 are orthologues of RmlC. We expected that they would perform only the epimerisation step, and not perform the structural

rearrangement activity that has yet to be assigned. As both enzymes contain all the catalytic residues proposed for 3'',5''-double epimerisation, we expected that they might perform both transformations. We demonstrate *in vivo* and *in vitro* that these enzymes are competent to perform the double epimerisation. Using NMR, we demonstrate that, surprisingly, proton-deuteron exchange at the 3'' position is faster than at the 5'' position for both enzymes. The structures of CBU1838 and StrM show that they form an active site highly similar to previously characterised enzymes, in keeping with the biochemical data. These data support a 3'', 5''-double epimerised substrate for the streptose ring re-arrangement.

Journal Pre-proof

Results:**CBU1838 complement RmlC in *E. coli***

We firstly tested whether CBU1838 can complement the *E. coli* RmlC orthologue RfbC. We cultured wild-type MG1655 *E. coli* and *rfbC* mutant bacteria, and *rfbC* mutants complemented with CBU1838. We extracted soluble small molecules from each culture. Wild-type MG1655 showed negligible levels of TDP-glucose and moderate levels of TDP-rhamnose. The *rfbC* mutant showed increased TDP-glucose levels and reduced TDP-rhamnose levels (Figure 3; $p < 0.0001$ and $p < 0.05$ respectively, ANOVA test comparing all tested samples). This is expected for a mutant at this essential step of the pathway. When the *rfbC* mutant was complemented with CBU1838, the wild-type sugar nucleotide levels were restored (Figure 3). In contrast, complementation with the RmlC paralogue EvaD, which preferentially epimerises the 5'' position, was not able to restore wild-type sugar nucleotide levels. These data gave good confidence that CBU1838 performs both 3'' and 5'' epimerisations of TDP-6-deoxy-D-xylo-4-hexulose, as expected.

StrM and CBU1838 complement RmlC in vitro

We next tested whether StrM and CBU1838 can perform the function of *E. coli* RmlC in a coupled assay. We used the test enzyme coupled to RmlB and RmlD in a catalytic cascade to convert TDP-glucose to TDP-rhamnose. Both StrM and CBU1838 complemented RmlC in these assays effectively (Figure 4). Identical concentrations of StrM and RmlC gave a similar effect. A higher concentration of CBU1838 was required to achieve the same effect. This is consistent with the expected function of CBU1838, epimerising a sugar that retains the 6''-OH group. These data further support the concept that StrM and CBU1838 perform both 3'' and 5'' epimerisations of TDP-6-deoxy-D-xylo-4-hexulose.

CBU1838 shows a reduced rate compared to RmlC and StrM, consistent with its modified substrate

We determined the kinetic constants for StrM and RmlC, in comparison to other characterised enzymes. We used purified TDP-6-deoxy-D-xylo-4-hexulose and coupled the test enzymes to RmlD. Data from StrM, RmlC, CBU1838 and EvaD all fitted very well to the Michaelis-Menten equation, with no evidence of cooperativity (Figure S1). StrM and RmlC showed a similar k_{cat} , with StrM having a slightly lower K_M (Table 2). CBU1838 showed a lower k_{cat} and a clearly higher K_M (Table 2). This is consistent with its expected 6''-hydroxy substrate. As expected, the 5''-specific epimerase EvaD showed a substantially lower k_{cat} . We did not observe any activity with the 3''-specific epimerase ChmJ (Figure S2). Previous studies have demonstrated that

ChmJ has no 5''-epimerisation activity (44), and that the coupling enzyme RmlD only reduces double-epimerised products (56,57). The kinetic constants provide further evidence that StrM is likely an orthologue of RmlC. CBU1838 can perform this activity, but its likely native substrate is 6''-hydroxylated.

RmlC, StrM and 1838 epimerise at both the 3'' and 5'' positions

To give further confidence in the activity of CBU1838 and StrM, in a novel assay, we monitored their activity directly using NMR and GC-MS without employing a coupling agent. We performed ¹H-NMR experiments in deuterated buffer. As RmlC uses a deprotonation-protonation mechanism with different proton acceptors and donors, we expected to observe a loss of signal from protons bound to the 5'' and 3'' carbon atoms. We observed solvent exchange at both positions for RmlC, StrM, CBU1838 and EvaD, with relative rates in the same order as the rates observed in a coupled assay (Figure 5A, S3). Surprisingly, the rate of exchange at the 3'' position was significantly faster than the rate at the 5'' position for RmlC, StrM and CBU1838 (Table 3); for EvaD the rate of exchange at the two positions was similar. When the reaction was permitted to proceed to completion, RmlC, StrM, CBU1838, and EvaD all showed almost complete exchange at both 3'' and 5'' positions (Figure 5B). In contrast, the 3'' specific enzyme ChmJ showed almost complete exchange at the 3'' position, with very little exchange observed at the 5'' position. GC-MS analysis of enzyme reactions run to completion confirmed that exchange had occurred at both positions for CBU1838, StrM and RmlC (Figure S4). Consistent with the NMR data and previous work (44), ChmJ showed strong exchange at the 3'' position only. These data strongly support a double epimerase activity for both CBU1838 and StrM.

Sugar backbone rearrangement is not observed with either StrM or CBU1838

In neither NMR nor GC-MS data was there any evidence of a rearrangement of the sugar backbone. The proposed biosynthetic pathways of both StrM and CBU1838 require such a rearrangement, catalysed by either the epimerase or a synthase/reductase for each pathway. Our direct detection of product formation and chemical structure analysis might have allowed us to detect this re-arrangement. However, as the five-membered rings are likely to be substantially disfavoured at equilibrium, the possibility remains that re-arranged structures were present but below the limit of detection.

The structures of StrM and CBU1838 are consistent with their role as double epimerases.

We determined the structures of StrM and CBU1838 by X-ray crystallography, in the presence and absence of TDP (Table S2). Both proteins show an overall dimeric architecture consistent

with previously solved members of this family (Figure 6A). Both proteins show a strand exchange in common with orthologues. A phylogenetic tree (Figure S5) suggests that CBU1838 groups with other 3'', 5'' epimerases, whilst StrM is closer in sequence to the single site epimerases. The sequence identities and structure RMSDs are consistent with this (Table S1). TDP binds to both proteins in a similar manner (Figure 6B,C). TDP is held in place by a network of amino acid side chains^a (F21*, R25*, F28*, E30*, Q48, R61, H64, Y140, K170) from both protomers (starred residues from the domain swapped strands). Additional interactions are formed through bound solvent from N50, Q72, Y134 and E145. These are all strongly conserved across RmlC orthologues (Figure 2). The CBU1838-TDP structure showed a citrate molecule (carried over from crystallisation) bound to the active site. This molecule occupies the same space as the substrate in previously solved structures. The citrate interacts with S52, H64, K74 and Y134 from the protein, and TDP (Figure 6D). These amino acids are well conserved amongst orthologues, and H64/Y134 are the catalytic base and acid proposed from previous studies (41-44). The conformation of these side chains is very similar to that seen in complexes of other enzymes with substrate analogues (Figure S6). These structures give further confidence that CBU1838 and StrM are indeed orthologues of RmlC. They have a very similar fold, show clear and specific binding to TDP, and retain the key catalytic residues in the correct orientation.

^a Residue numbers given are for CBU1838. The equivalent residues in StrM are (in order mentioned on this page) R21*, F24*, E26*, Q45, R57, H60, Y135, and K165; N47, Q68, Y129, and E140; S49 and K70.

Discussion

Streptose and DHHS are monosaccharides produced by a single organism each. Both sugars are proposed to be synthesised through modified TDP-rhamnose pathways. *S. griseus* and *C. burnetii* contain putative orthologues of this biosynthetic pathway. Both organisms contain a single TDP-6-deoxy-D-xylo-4-hexulose 3'',5''-epimerase orthologue (StrM and CBU1838). We sought to determine whether these enzymes perform both epimerisations, and whether either might also perform the sugar ring rearrangement activity. Answering these questions will provide further insight into the biosynthesis of these unusual sugars.

CBU1838 showed robust double epimerase activity both in vitro and in vivo. CBU1838 complemented an *rfbC* (*rmlC*) mutation, restoring flux through the pathway and consuming cellular TDP-glucose and TDP-6-deoxy-D-xylo-4-hexulose (Figure 3). CBU1838 and StrM were competent to replace RmlC in vitro in a biosynthetic pathway (Figure 4). It was notable that substantially more CBU1838 was required to restore activity than RmlC or StrM. Similarly, CBU1838 showed a lower k_{cat} , and higher K_M , than either RmlC or StrM (Table 2). This is consistent with the proposed role of CBU1838, as the DHHS pathway requires a substrate with a 6''-hydroxyl. CBU1838 contains a methionine residue (M123) in place of a phenylalanine conserved among other RmlC orthologues, which would provide space for the hydroxyl. The k_{cat} and K_M values that we obtained for StrM and RmlC are in the same order as those observed for other RmlC paralogues (Table S3). It is notable that our data were collected at 37 °C, whilst the comparative data were collected at 21-25 °C. It is likely that other enzymes would react faster at higher temperatures but might show a higher K_M (58).

Our NMR data show that CBU1838 and StrM have similar activity to *E. coli* RmlC, with CBU1838 again showing a reduced comparative activity. The rate of proton exchange at the 3'' position was faster than that at the 5'' position for all three enzymes (Figure 5; Table 3). The *M. tuberculosis* RmlC showed faster epimerisation at the 5'' position (41). Dong et al. suggested that in the likely catalytic conformation, the 5'' proton is more acidic than the 3'' proton (41). Our experiments measured the loss of proton signal by NMR, whilst the previous study measured the product masses by GC-MS. Furthermore, our study used chemically prepared TDP-6-deoxy-D-xylo-4-hexulose, which has only recently become commercially available. This may explain some differences between our study and previous work. We note that as all the reactions involve proton transfers at all steps (40), there may be a strong kinetic isotope effect (59) on either or both epimerisations. As our methods for differentially measuring the two epimerisations rely on deuterium, this confounding effect cannot be excluded.

The structures of CBU1838 and StrM are highly similar to previously solved structures of paralogues from other species. TDP is held in place by residues that are strongly conserved (Figure 2, 6) (43,48,52). Although we were not able to determine a structure with a TDP-linked sugar, the side chains lining the sugar binding (and catalytic) cavity adopt conformations that are highly similar to those in previous structures (38,41,43-45,48,50,51,53). The proposed catalytic residues, H64 and Y134, occupy a very similar space to the equivalent residues in substrate analogue co-crystal structures (41,53). As their local hydrogen-bonding networks are also conserved, it is likely that they will also be catalytic in these proteins. The presence and conformation of these catalytic residues cannot predict the specificity of the enzymes. Whilst one example of a 5''-specific enzyme showed a mutation that altered the conformation of an important active site side chain (50), 3''-specific epimerases have shown a complete complement of active site residues (44,51). Our development of a functional assay in *E. coli* allowed us to confirm the activity of StrM and CBU1838 *in vivo* (Figure 3); although the 5''-specific enzyme EvaD was expressed to the same levels as the active enzymes, it was not competent to complement *rmIC*. However, this assay did not distinguish between StrM (likely a direct RmIC orthologue) and CBU1838 (which is expected to use the 6''-hydroxylated analogue).

The evidence presented in this and previous studies strongly suggests that both CBU1838 and StrM act as TDP-4-keto-(6-deoxy)-glucose 3'', 5'' double epimerases. The lower rates observed for CBU1838 are consistent with its natural substrate retaining the 6''-hydroxyl group; an activity to generate this substrate has not yet been identified. This suggests that the streptose and DHHS synthase enzymes (Figure 1) accept a substrate that is epimerised at both the 3'' and 5'' positions. Although our data suggest that these enzymes cause proton exchange at the 3'' carbon more efficiently than the 5'' carbon, we cannot exclude that the kinetic isotope effect could be confounding. Our results provide powerful evidence that DHHS biosynthesis in *Coxiella* has convergently evolved to use a similar pathway to streptose biosynthesis in *S. griseus*. The key challenge that this presents is to identify the enzymes that provide the two remaining activities to generate DHHS.

Experimental procedures

Gene cloning and synthesis

The gene fragments for *rmlB*, *rmlC* and *rmlD* were amplified from *E. coli* DH5 α (NEB) genomic DNA (gDNA) by PCR. The *CBU1838* gene fragment was codon-optimised using IDT's tools and ordered as a gBlock from IDT and amplified by PCR. Primer sequences are provided in Table S4. Gene fragments were cloned into pNIC28-Bsa4 (Addgene #26103, gift of Opher Gileadi, SGC Oxford) using ligation-independent cloning (LIC), following published methods (60). Briefly, 16.5 μ g pNIC28-Bsa4 plasmid was digested with 60 U *BsaI* (NEB) in 100 μ L following the manufacturer's protocol, and incubated for 2 h at 50 °C. To generate the LIC cohesive ends in both the plasmid and PCR product, a mixture of 1 μ L water, 5 μ L *BsaI*-digested plasmid/PCR-pure DNA, 2 μ L 5X T4 DNA polymerase buffer (Fermentas), 1 μ L 25 mM dGTP (insert) or dCTP (plasmid), 0.5 μ L 100 mM DTT, 0.5 μ L T4 DNA polymerase (Fermentas) was prepared, and incubated at 22 °C for 30 min, then 20 min at 75 °C. 2 μ L of treated DNA product was added to 1 μ L of treated plasmid. The mixture was incubated at room temperature for 10 min followed by transformation into 5-alpha competent cells (NEB). The *E. coli* RmlD contains an inactivating point mutation. This was corrected by site-directed mutagenesis (SDM) using the QuikChange Lightning kit (Agilent), following the manufacturer's instructions.

TDP-sugar epimerase genes for *Streptomyces griseus strM*, *Amycolatopsis orientalis evaD* and *Streptomyces bikiniensis chmJ* were codon optimised using an in-house Python script (<https://github.com/njharmer/CodonOptimise>) and cloned into pET28b with an additional SUMO tag, pET28b, and pET21a respectively by Twist Bioscience. Plasmid were transformed into the expression strain *E. coli* BL21 (DE3) (Novagen) using kanamycin (50 μ g/mL) for selection. For complementation of *E. coli rfbC* mutants, TDP-sugar epimerase genes for *strM*, *evaD* and *CBU1838* were cloned into the cloning vector pTWIST-A-MC under the control of the lac operon, with a strong ribosome binding site calculated by the RBS Calculator (De Novo DNA). The Genbank files for all plasmids are available in Supplementary Data.

Preparation of electrocompetent *E. coli* mutant cell lines and complementation

Δ *rfbC* *E. coli* strains (Keio collection; Horizon Discovery #OEC4987-200827372) were grown overnight on LB plates with 50 μ g/mL kanamycin. Colonies were inoculated from LB plates into 10 mL fresh LB broth and incubated overnight at 37°C with shaking at 225 rpm. 1 mL of overnight culture was inoculated into 100 mL fresh LB medium with kanamycin and incubated at 37 °C with shaking at 225 rpm until OD600 was approximately 0.6. Cells were harvested by centrifugation (4200 *xg*, 10 min, 4 °C). The cells were resuspended in 40 mL pre-chilled autoclaved Milli-Q water. Cells were harvested by centrifugation as above and resuspended in

40 mL pre-chilled autoclaved 10 % (v/v) glycerol. A minimum of three 10 % glycerol washes were conducted before cells were resuspended in 500 μ L 10 % (v/v) glycerol.

Complementation plasmids (40 ng) were added to 50 μ L electrocompetent cells, vortexed for 30 s and incubated on ice for 10 min. The suspension was transferred to a pre-chilled 0.1 cm electroporation cuvette (Thermo Scientific #5510-11). A single pulse of 1.8 kV was given by a Gene Pulser Xcell[®] Microbial System (Bio-Rad #1652662). Cells were recovered by adding 1000 μ L pre-warmed LB medium. The complemented cell suspension was incubated at 37 °C for 1 hour shaking at 225 rpm. Transformed cells were grown on LB agar plates supplemented with 100 μ g/mL ampicillin overnight at 37 °C.

Detecting expression of complemented *E. coli*

Complemented *E. coli* strain colonies were inoculated into 100 mL bottles containing 10 mL LB medium, 100 μ g/mL ampicillin and incubated overnight at 37 °C with shaking at 225 rpm. 1 mL of each overnight culture was inoculated into individual 500 mL conical flasks containing 100 mL LB medium supplemented with ampicillin and incubated at 37 °C with shaking at 225 rpm. When the OD_{600 nm} reached 0.6, gene expression was induced by adding 200 μ M IPTG. Cultures were incubated at 20 °C with shaking at 225 rpm overnight. 50 mL aliquots of overnight cultures were harvested by centrifugation (4500 xg , 30 min, 4 °C). Cells were lysed using 1 mL BugBuster Master Mix (Millipore #71456-4), following the manufacturer's instructions. The soluble and insoluble fractions were separated by centrifugation at 16 000 xg for 30 min at 4 °C. Samples were separated on a 4-12% ExpressPlus[™] PAGE gel (GenScript #M41212) following the manufacturer's instructions. Samples were transferred to a nitrocellulose membrane (Sartorius Stedim Biotech #11327-41BL) using a Pierce G2 Fast Blotter (Thermo Scientific # 15146375) using the pre-programmed protocol. Antibodies were added using an iBind Western device (ThermoFisher #SLF1000), following the manufacturer's recommendations. The primary and secondary antibodies used were anti-penta His (Qiagen #34660) and goat anti-mouse IgG (LiCOR #926-68070) at a dilution of 1:1000. Blots were imaged using an Odyssey Clx imaging system (LiCOR).

Sample preparation for LC-MS QQQ analysis

Complemented *E. coli* strain colonies were inoculated into 100 mL bottles containing 10 mL LB medium, 100 μ g/mL ampicillin for complemented strains, and incubated overnight at 37 °C with shaking at 225 rpm. 1 mL of each overnight culture was inoculated into individual 500 mL conical flasks containing 100 mL LB medium supplemented with kanamycin and ampicillin as necessary and incubated at 37 °C with shaking at 225 rpm until OD_{600 nm} reached 0.6. Expression was

induced by adding 200 μ M IPTG. Cultures were incubated at 20 °C with shaking at 225 rpm overnight. 50 mL aliquots of overnight cultures were harvested by centrifugation (4500 xg , 30 min, 4 °C). Cellular contents were extracted into 1 mL 50 % (v/v) acetonitrile. Cell suspensions were incubated at 25 °C for 30 min and centrifuged at 20 000 xg for 30 min at 4 °C. The supernatant was filtered using Millex Non-Sterile Low Protein Binding Hydrophilic LCR (PTFE) Membrane (0.45 μ m) filters (Merck #SLLHR04NL) into 1.5 mL MS vials with silicone/PTFE septa and stored at -20 °C for later analysis.

Quantitative LC-MS QQQ Analysis:

Quantitative analysis was performed using an Agilent 6420B triple quadrupole (QQQ) mass spectrometer (Agilent Technologies, Palo Alto, USA) coupled to a 1200 series Rapid Resolution HPLC system. 5 μ L of sample extract was loaded onto an Agilent Poroshell 120 HILIC-Z, 2.7 μ m, 2.1 x 150 mm analytical column (Agilent Technologies #673775-924). For detection using negative ion mode, mobile phase A comprised 90 % (v/v) LC-MS grade acetonitrile with 10 mM ammonium acetate and 5 μ M medronic acid, and mobile phase B was 100% water (LC-MS grade) also with 10 mM ammonium acetate and 5 μ M medronic acid. The following gradient was used: 0 min – 10% B; 6 min – 35% B; 10 – 13 min – 40% B; 15 min – 10% B; followed by 5 minutes re-equilibration time at a flow rate of 0.25 mL min⁻¹ with the column held at 25 °C for the duration. The QQQ source conditions for electrospray ionisation were: 350 °C gas temperature with a drying gas flow rate of 11 L min⁻¹ and a nebuliser pressure of 35 psig. The capillary voltage was 4 kV.

MS Data Analysis:

Data analysis was undertaken using Agilent MassHunter Quantitative Analysis software (version B.07.01, SP1). Data was normalised to the internal standard, ¹³C TDP-glucose, and the optical density of each culture pre-extraction. TDP-glucose, TDP-rhamnose, GDP-mannose, GDP-fucose and UDP-glucose concentrations were calculated using standard calibration curves. For the other compounds (where a standard was not available) normalised peak areas were compared.

Expression and purification of TDP-sugar epimerases:

RmlB, RmlC, RmlD, StrM, EvaD and ChmJ were expressed in 500 mL of ZYM-5052 auto-induction media supplemented with 100 μ g/mL kanamycin following the methods of (61). Each flask was inoculated with 10 mL of an overnight culture and grown at 37 °C with agitation at 200 rpm until OD₆₀₀ reached 0.6. Cultures were further incubated at 20 °C for 18 hours with agitation. Cells were harvested by centrifugation at 4500 xg for 30 min at 4 °C. The pellet was resuspended in binding buffer (20 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole, pH 8.0) and lysed by sonication

(SONIC Vibra cell VCX130). The sample was clarified by centrifugation (24 000 xg for 30 min at 4 °C). The soluble fraction was purified using an ÄKTAexpress chromatography system (Cytiva). The sample was purified using a 1 mL HisTrap crude column (Cytiva). After loading sample, the column was washed with binding buffer, and the protein step eluted into binding buffer supplemented with 250 mM. The eluate was purified over a Superdex 200 pg 16/600 size-exclusion column (Cytiva) and eluted isocratically into 10 mM HEPES, 500 mM NaCl, pH 7.5 (for coupled assays and crystallisation) or 10 mM Na_2HPO_4 , 500 mM NaCl, pH 7.5 for NMR studies. The eluted protein was concentrated using a Vivaspin centrifugal concentrator (Generon) and stored at -20 °C with 20 % (v/v) glycerol for enzymatic assays; or stored at -80°C in small aliquots without any glycerol for crystallization or NMR.

Biochemical assays:

The TDP-rhamnose biosynthesis assay coupled the product of RmlC paralogues to RmlD. In brief, reactions were performed in 96-well flat-bottomed plates (Greiner #655001) in a total volume of 200 μL . For initial studies, reactions consisted of 50 mM HEPES pH 7.5, 20 mM MgCl_2 , 350 μM NADPH, 4 μM NAD^+ , 0.4 μM RmlB, 2 μM RmlC paralogue and 2 μM RmlD. Reactions were initiated by addition of 500 μM TDP-glucose (Carbosynth #MT04383). Reactions were monitored by measurement of the absorbance at 340 nm over 30 min in an Infinite M200 PRO plate reader (Tecan) incubating at 30 °C. For determination of kinetic parameters, RmlB and NAD^+ were removed and TDP-glucose was replaced with TDP-6-deoxy-D-xylo-4-hexulose (Carbosynth #NT29846) added at a concentration range between 0-3 mM (*EcRmlC* / *SgStrM*) or 0-9 mM (CBU1838 or *AoEvaD*). Three experimental replicates were performed for all reactions. Data were fitted to the Michaelis-Menten equation using Graphpad v. 8.1.2.

Deuterium-incorporation analysis by ^1H NMR:

Reaction mixtures contained 10 mM MgCl_2 , 2.38 mM TDP-6-deoxy-D-xylo-4-hexulose (Carbosynth) and 8 μM NAD^+ in a final volume of 250 μL deuterated 100 mM phosphate buffer pD 7.0. A “zero time” spectrum was recorded without any epimerase. Epimerases (38 nM *EcRmlC*, 46 nM *SgStrM*, 177 nM CBU1838, or 530 nM *AoEvaD*) were added to the above reaction mixture to a final volume of 250 μL .

All data were collected on a Bruker Neo 600 MHz spectrometer equipped with TCI cryoprobe. Standard Bruker proton experiments were carried out at 298 K (25 °C). Periodic ^1H NMR experiments (64 scans) were carried out in 3 mm NMR tubes. ^1H NMR spectra were recorded every 10 min over a 90 min period, set using the Topspin multi_zgvd command. Later intervals were set manually to record at 2 h, 4 h, 8 h, 18 h and 24 h, as appropriate. TopSpin 4.0.6 was used for data processing using command ‘efp’ followed by automatic phase correction (‘apk’)

and automatic baseline correction ('abs'). The residual water peak at δ 4.7058 ppm was used as reference for spectra calibration. A set of selected diagnostic signals from no-enzyme control samples was integrated for calibration of the relative abundance of each signal. An integral of H-6 thymidine peak (^1H) was used for peak intensity calibration of the rest of integrals of diagnostic signals and GraphPad Prism v8.1.2 was used to analyse the data obtained.

Deuterium-incorporation analysis GC-MS

0.5 mg TDP-6-deoxy-D-xylo-4-hexulose (Carbosynth) was incubated in deuterated buffer alone, or with the addition of either 1 μM of EcRmlB, EcRmlC, SgStrM, CBU1838 or AoEvaD in 250 μL final volume. After 18 h, enzymes were inactivated and removal by ethanol precipitation followed by centrifugation at 14,000 $\times g$ for 10 min. Sodium borohydride (~ 2 mg) was added to each sample for the reduction of keto-moety followed by incubation at room temperature for 2 h, with shaking at 120 rpm. The reaction was terminated with few drops of glacial acetic acid, followed by 20 μL 10 % (v/v) acetic acid in methanol. Solvents were evaporated under reduced pressure using GeneVac, for 1 h at 35 $^\circ\text{C}$ and 25 μg myo-inositol was added to the sample as an internal standard. The sugar-nucleotide bond was hydrolysed by using 250 μL 2 M trifluoroacetic acid (TFA) followed by incubation at 100 $^\circ\text{C}$ for 1.5 h. After cooling the samples, solvents were evaporated again and 100 μL isopropanol was added before a final evaporation in order to remove traces of TFA for 1 h.

A second reduction was achieved by addition of 250 μL 1 M ammonia, containing 10 mg/mL sodium borohydride for 1 h and then 250 μL acetic acid-methanol (1:9, v/v) was added. The samples were dried under reduced pressure (GeneVac, 1h at 40 $^\circ\text{C}$) and the addition of acetic acid-methanol followed by drying was repeated once again. This was followed by addition of 250 μL methanol (without acetic acid) twice with evaporating the solvent off each time (20 min each). The resulting sample was acetylated by incubation at 120 $^\circ\text{C}$ for 20 min in a mixture of 100 μL acetic anhydride and 100 μL pyridine. After this, approximately 200 μL toluene was added, and the mixture was concentrated under reduced pressure. The addition of toluene and evaporation was repeated at room temperature.

The alditol acetates were separated by partition between equal volume of methylene chloride and water. After partition the methylene chloride layer was left to evaporate slowly at room temperature. Finally, samples were dissolved in acetone and analysed by gas chromatography-mass spectrometry (GC-MS).

Crystallization:

Protein concentrations used for crystallization were 5.85 mg/mL (*SgStrM*) and 5.2 mg/mL or 7.8 mg/mL (CBU1838). Crystals were grown using the microbatch method using an Oryx8 crystallization robot (Douglas Instruments). Crystals were grown in hydrophobic plates (Douglas Instruments #VB-Silver-1/1) and covered in a 1:1 mix of paraffin oil and silicone oil (Molecular Dimensions). Crystals were grown either as a 1:1 mixture of protein and mother liquor, or as a 3:2:1 mixture of protein, mother liquor and seed stocks where seeds had been prepared from previous crystals. The successful co-crystallisation conditions, substrates soaking conditions, and cryoprotectants used are detailed in Table S5.

X-ray data collection and structure determination:

Data were collected at Diamond Light Source (Didcot, UK) at 100 K using Pilatus 6M-F detectors (Table S2). All data were processed using XDS (62). CBU1838 data were subjected to anisotropic ellipsoidal truncation using the STARANISO server (63). The resolution limit for StrM datasets was set at a point where $CC_{\frac{1}{2}}$ was above 0.3 in the highest resolution shell. CBU1838 data were processed to include regions where the mean I/σ was systematically above 1.2. Further data processing and structural studies was carried out using the CCP4 program package (64) and the CCP4i2 interface (65). Molecular replacement was used to provide initial phases for datasets using Morda (66) and MolRep (67). The best models for CBU1838 and StrM were 2IXI (41) (54% identity) and 4HMZ (44) (52% identity) respectively. Density fitting was performed using Coot (68) and refined with REFMAC5 (69). Model validation was carried out with internal modules of CCP4i2 and Coot, employing MolProbity calculations (70).

Acknowledgements

The authors thank Simone de Rose (University of Exeter) for assistance with crystallisation.

SR, MVV, RF and NJH were funded by BBSRC grant BB/N001591/1. ARC and NJH were funded by BBSRC grant BB/M016404/1 and by Dstl grant DSTLX-1000098217. Work at the John Innes Centre (MR, SAN, RAF) was supported by the UK BBSRC Institute Strategic Program on Molecules from Nature -Products and Pathways [BBS/E/J/000PR9790] and the John Innes Foundation, and the InnovateUK: IBCatalyst (Grant BB/M02903411).

Author contributions

RAF, JP and NJH designed the study and obtained funding. ARC, SR, MVV, MR, SAN, MC and DS collected data. ARC, SR, MVV, MR, MC, DS, MI, RAF and NJH analysed data. ARC and NJH wrote the initial manuscript. All authors contributed to revision of the manuscript.

Data availability

All data underpinning this work are publicly available. Structure coordinates and structure factor files are deposited with the Protein Data Bank (accession numbers: 7PA1, 7P7I, 7P7W, 7P9L, 7P9P and 7P9Y). Enzymatic and mass spectrometry data are available as Supplementary Files or from Open Research Exeter (doi: to be confirmed on acceptance).

References

1. Duron, O., Noel, V., McCoy, K. D., Bonazzi, M., Sidi-Boumedine, K., Morel, O., Vavre, F., Zenner, L., Jourdain, E., Durand, P., Arnathau, C., Renaud, F., Trape, J. F., Biguezoton, A. S., Cremaschi, J., Dietrich, M., Leger, E., Appelgren, A., Dupraz, M., Gomez-Diaz, E., Diatta, G., Dayo, G. K., Adakal, H., Zoungrana, S., Vial, L., and Chevillon, C. (2015) The Recent Evolution of a Maternally-Inherited Endosymbiont of Ticks Led to the Emergence of the Q Fever Pathogen, *Coxiella burnetii*. *Plos Pathog* **11**
2. Eldin, C., Melenotte, C., Mediannikov, O., Ghigo, E., Million, M., Edouard, S., Mege, J. L., Maurin, M., and Raoult, D. (2017) From Q Fever to *Coxiella burnetii* Infection: a Paradigm Change. *Clin Microbiol Rev* **30**, 115-190
3. Duncan, C., Kersh, G. J., Spraker, T., Patyk, K. A., Fitzpatrick, K. A., Massung, R. F., and Gelatt, T. (2012) *Coxiella burnetii* in Northern Fur Seal (*Callorhinus ursinus*) Placentas from St. Paul Island, Alaska. *Vector-Borne Zoonot* **12**, 192-195
4. Gonzalez-Barrio, D., and Ruiz-Fons, F. (2019) *Coxiella burnetii* in wild mammals: A systematic review. *Transbound Emerg Dis* **66**, 662-671
5. Mori, M., and Roest, H. J. (2018) Farming, Q fever and public health: agricultural practices and beyond. *Arch Public Health* **76**
6. Cross, A. R., Baldwin, V. M., Roy, S., Essex-Lopresti, A. E., Prior, J. L., and Harmer, N. J. (2019) Zoonoses under our noses. *Microbes Infect* **21**, 10-19
7. Clark, N. J., and Magalhaes, R. J. S. (2018) Airborne geographical dispersal of Q fever from livestock holdings to human communities: a systematic review and critical appraisal of evidence. *Bmc Infect Dis* **18**
8. Tigertt, W. D., Benenson, A. S., and Gochenour, W. S. (1961) Airborne Q Fever. *Bacteriol Rev* **25**, 285-+
9. Anastacio, S., Carolino, N., Sidi-Boumedine, K., and da Silva, G. J. (2016) Q Fever Dairy Herd Status Determination Based on Serological and Molecular Analysis of Bulk Tank Milk. *Transbound Emerg Dis* **63**, e293-300
10. Muskens, J., van Engelen, E., van Maanen, C., Bartels, C., and Lam, T. J. (2011) Prevalence of *Coxiella burnetii* infection in Dutch dairy herds based on testing bulk tank milk and individual samples by PCR and ELISA. *Vet Rec* **168**, 79
11. Ryan, E. D., Wrigley, K., Hallinan, A., McGrath, G., and Clegg, T. A. (2018) Antibodies to *Coxiella burnetii* in Irish bulk tank milk samples. *Vet Rec* **182**, 550
12. Raoult, D., Marrie, T. J., and Mege, J. L. (2005) Natural history and pathophysiology of Q fever. *Lancet Infect Dis* **5**, 219-226
13. Million, M., and Raoult, D. (2017) No Such Thing as Chronic Q Fever. *Emerg Infect Dis* **23**, 856-857
14. Seddon, O., Ashrafi, R., Duggan, J., Rees, R., Tan, C., Williams, J., Carson, G., and Healy, B. (2016) Seroprevalence of Q Fever in Patients Undergoing Heart Valve Replacement Surgery. *J Heart Valve Dis* **25**, 375-379

15. Boden, K., Wolf, K., Hermann, B., and Frangoulidis, D. (2015) First isolation of *Coxiella burnetii* from clinical material by cell-free medium (ACCM2). *Eur J Clin Microbiol* **34**, 1017-1022
16. Taurel, A. F., Guatteo, R., Joly, A., and Beaudeau, F. (2012) Effectiveness of vaccination and antibiotics to control *Coxiella burnetii* shedding around calving in dairy cows. *Vet Microbiol* **159**, 432-437
17. Zuckerman, J. N., Van Damme, P., Van Herck, K., and Loscher, T. (2003) Vaccination options for last-minute travellers in need of travel-related prophylaxis against hepatitis A and B and typhoid fever: a practical guide. *Travel Med Infect Dis* **1**, 219-226
18. Kazar, J., Brezina, R., Palanova, A., Tvrda, B., and Schramek, S. (1982) Immunogenicity and reactogenicity of a Q fever chemovaccine in persons professionally exposed to Q fever in Czechoslovakia. *Bull World Health Organ* **60**, 389-394
19. Samuel, J. E., Van Schaik, E. J., and Gregory, A. E. (2019) VACCINE FOR IMMUNIZATION AGAINST Q-FEVER (Office, U. P. ed., USA
20. Reeves, P. M., Paul, S. R., Sluder, A. E., Brauns, T. A., and Poznansky, M. C. (2017) Q-vaxcelerate: A distributed development approach for a new *Coxiella burnetii* vaccine. *Hum Vacc Immunother* **13**, 2977-2981
21. Scholzen, A., Richard, G., Moise, L., Baeten, L. A., Reeves, P. M., Martin, W. D., Brauns, T. A., Boyle, C. M., Raju Paul, S., Bucala, R., Bowen, R. A., Garritsen, A., De Groot, A. S., Sluder, A. E., and Poznansky, M. C. (2019) Promiscuous *Coxiella burnetii* CD4 Epitope Clusters Associated With Human Recall Responses Are Candidates for a Novel T-Cell Targeted Multi-Epitope Q Fever Vaccine. *Front Immunol* **10**, 207
22. Downs, C. M. (1968) Phagocytosis of *coxiella burnetii*, phase I and phase II by peritoneal monocytes from normal and immune guinea pigs and mice. *Zentralbl Bakteriol Orig* **206**, 329-343
23. van Schaik, E. J., Chen, C., Mertens, K., Weber, M. M., and Samuel, J. E. (2013) Molecular pathogenesis of the obligate intracellular bacterium *Coxiella burnetii*. *Nat Rev Microbiol* **11**, 561-573
24. Hackstadt, T. (1986) Antigenic variation in the phase I lipopolysaccharide of *Coxiella burnetii* isolates. *Infect Immun* **52**, 337-340
25. Kersh, G. J., Oliver, L. D., Self, J. S., Fitzpatrick, K. A., and Massung, R. F. (2011) Virulence of pathogenic *Coxiella burnetii* strains after growth in the absence of host cells. *Vector Borne Zoonotic Dis* **11**, 1433-1438
26. Beare, P. A., Jeffrey, B. M., Long, C. M., Martens, C. M., and Heinzen, R. A. (2018) Genetic mechanisms of *Coxiella burnetii* lipopolysaccharide phase variation. *Plos Pathog* **14**, e1006922
27. Moos, A., and Hackstadt, T. (1987) Comparative virulence of intra- and interstrain lipopolysaccharide variants of *Coxiella burnetii* in the guinea pig model. *Infect Immun* **55**, 1144-1150
28. Beare, P. A., Unsworth, N., Andoh, M., Voth, D. E., Omsland, A., Gilk, S. D., Williams, K. P., Sobral, B. W., Kupko, J. J., Porcella, S. F., Samuel, J. E., and Heinzen, R. A. (2009) Comparative Genomics Reveal Extensive Transposon-Mediated Genomic Plasticity and Diversity among Potential Effector Proteins within the Genus *Coxiella*. *Infection and Immunity* **77**, 642-656
29. Amano, K., Williams, J. C., Missler, S. R., and Reinhold, V. N. (1987) Structure and biological relationships of *Coxiella burnetii* lipopolysaccharides. *J Biol Chem* **262**, 4740-4747
30. Schramek, S., Radziejewskalebrecht, J., and Mayer, H. (1985) 3-C-Branched Aldoses in Lipopolysaccharide of Phase-I *Coxiella-Burnetii* and Their Role as Immunodominant Factors. *Eur J Biochem* **148**, 455-461
31. Kharel, M. K., Nybo, S. E., Shepherd, M. D., and Rohr, J. (2010) Cloning and characterization of the ravidomycin and chrysomycin biosynthetic gene clusters. *ChemBiochem* **11**, 523-532

32. Li, Z., Mukherjee, T., Bowler, K., Namdari, S., Snow, Z., Prestridge, S., Carlton, A., and Bar-Peled, M. (2017) A four-gene operon in *Bacillus cereus* produces two rare spore-decorating sugars. *J Biol Chem* **292**, 7636-7650
33. Flores-Ramirez, G., Janecek, S., Miernyk, J. A., and Skultety, L. (2012) In silico biosynthesis of virenose, a methylated deoxy-sugar unique to *Coxiella burnetii* lipopolysaccharide. *Proteome Sci* **10**, 67
34. Flatt, P. M., and Mahmud, T. (2007) Biosynthesis of aminocyclitol-aminoglycoside antibiotics and related compounds. *Nat Prod Rep* **24**, 358-392
35. Bruton, J., and Horner, W. H. (1966) Biosynthesis of streptomycin. 3. Origin of the carbon atoms of streptose. *J Biol Chem* **241**, 3142-3146
36. Giraud, M. F., and Naismith, J. H. (2000) The rhamnose pathway. *Curr Opin Struc Biol* **10**, 687-696
37. Wahl, H. P., and Grisebach, H. (1979) Biosynthesis of streptomycin. dTDP-dihydrostreptose synthase from *Streptomyces griseus* and dTDP-4-keto-L-rhamnose 3,5-epimerase from *S. griseus* and *Escherichia coli* Y10. *Biochim Biophys Acta* **568**, 243-252
38. Giraud, M. F., Leonard, G. A., Field, R. A., Berlind, C., and Naismith, J. H. (2000) RmlC, the third enzyme of dTDP-L-rhamnose pathway, is a new class of epimerase. *Nat Struct Biol* **7**, 398-402
39. Savino, S., Borg, A. J. E., Dennig, A., Pfeiffer, M., de Giorgi, F., Weber, H., Dubey, K. D., Rovira, C., Mattevi, A., and Nidetzky, B. (2019) Deciphering the enzymatic mechanism of sugar ring contraction in UDP-apiose biosynthesis. *Nat Catal* **2**, 1115-1123
40. Van Overtveldt, S., Verhaeghe, T., Joosten, H. J., van den Bergh, T., Beerens, K., and Desmet, T. (2015) A structural classification of carbohydrate epimerases: From mechanistic insights to practical applications. *Biotechnol Adv* **33**, 1814-1828
41. 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., and Naismith, J. H. (2007) RmlC, a C3' and C5' carbohydrate epimerase, appears to operate via an intermediate with an unusual twist boat conformation. *J Mol Biol* **365**, 146-159
42. Huddleston, J. P., Anderson, T. K., Girardi, N. M., Thoden, J. B., Taylor, Z., Holden, H. M., and Raushel, F. M. (2021) Biosynthesis of d-glycero-l-gluco-Heptose in the Capsular Polysaccharides of *Campylobacter jejuni*. *Biochemistry* **60**, 1552-1563
43. Jakimowicz, P., Tello, M., Meyers, C. L., Walsh, C. T., Buttner, M. J., Field, R. A., and Lawson, D. M. (2006) The 1.6-Å resolution crystal structure of NovW: a 4-keto-6-deoxy sugar epimerase from the novobiocin biosynthetic gene cluster of *Streptomyces spheroides*. *Proteins* **63**, 261-265
44. Kubiak, R. L., Phillips, R. K., Zmudka, M. W., Ahn, M. R., Maka, E. M., Pyeatt, G. L., Roggensack, S. J., and Holden, H. M. (2012) Structural and functional studies on a 3'-epimerase involved in the biosynthesis of dTDP-6-deoxy-D-allose. *Biochemistry* **51**, 9375-9383
45. Christendat, D., Saridakis, V., Dharamsi, A., Bochkarev, A., Pai, E. F., Arrowsmith, C. H., and Edwards, A. M. (2000) Crystal structure of dTDP-4-keto-6-deoxy-D-hexulose 3,5-epimerase from *Methanobacterium thermoautotrophicum* complexed with dTDP. *J Biol Chem* **275**, 24608-24612
46. Kantardjieff, K. A., Kim, C. Y., Naranjo, C., Waldo, G. S., Lakin, T., Segelke, B. W., Zemla, A., Park, M. S., Terwilliger, T. C., and Rupp, B. (2004) *Mycobacterium tuberculosis* RmlC epimerase (Rv3465): a promising drug-target structure in the rhamnose pathway. *Acta Crystallogr D Biol Crystallogr* **60**, 895-902
47. Li, Z. Z., Riegert, A. S., Goneau, M. F., Cunningham, A. M., Vinogradov, E., Li, J., Schoenhofen, I. C., Thoden, J. B., Holden, H. M., and Gilbert, M. (2017) Characterization of the dTDP-Fuc3N and dTDP-Qui3N biosynthetic pathways in *Campylobacter jejuni* 81116. *Glycobiology* **27**, 358-369

48. Shornikov, A., Tran, H., Macias, J., Halavaty, A. S., Minasov, G., Anderson, W. F., and Kuhn, M. L. (2017) Structure of the Bacillus anthracis dTDP-L-rhamnose-biosynthetic enzyme dTDP-4-dehydrorhamnose 3,5-epimerase (RfbC). *Acta Crystallogr F Struct Biol Commun* **73**, 664-671
49. Thoden, J. B., Vinogradov, E., Gilbert, M., Salinger, A. J., and Holden, H. M. (2015) Bacterial Sugar 3,4-Ketoisomerases: Structural Insight into Product Stereochemistry. *Biochemistry* **54**, 4495-4506
50. Merkel, A. B., Major, L. L., Errey, J. C., Burkart, M. D., Field, R. A., Walsh, C. T., and Naismith, J. H. (2004) The position of a key tyrosine in dTDP-4-Keto-6-deoxy-D-glucose-5-epimerase (EvaD) alters the substrate profile for this RmlC-like enzyme. *J Biol Chem* **279**, 32684-32691
51. Salinger, A. J., Brown, H. A., Thoden, J. B., and Holden, H. M. (2015) Biochemical studies on WbcA, a sugar epimerase from Yersinia enterocolitica. *Protein Sci* **24**, 1633-1639
52. Dhaked, D. K., Bala Divya, M., and Guruprasad, L. (2019) A structural and functional perspective on the enzymes of Mycobacterium tuberculosis involved in the L-rhamnose biosynthesis pathway. *Prog Biophys Mol Biol* **145**, 52-64
53. Dong, C., Major, L. L., Allen, A., Blankenfeldt, W., Maskell, D., and Naismith, J. H. (2003) High-resolution structures of RmlC from Streptococcus suis in complex with substrate analogs locate the active site of this class of enzyme. *Structure* **11**, 715-723
54. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research* **44**, D457-D462
55. Koller, F., and Lassak, J. (2021) Two RmlC homologs catalyze dTDP-4-keto-6-deoxy-D-glucose epimerization in Pseudomonas putida KT2440. *Sci Rep* **11**, 11991
56. Blankenfeldt, W., Kerr, I. D., Giraud, M. F., McMiken, H. J., Leonard, G., Whitfield, C., Messner, P., Graninger, M., and Naismith, J. H. (2002) Variation on a theme of SDR. dTDP-6-deoxy-L-lyxo-4-hexulose reductase (RmlD) shows a new Mg²⁺-dependent dimerization mode. *Structure* **10**, 773-786
57. Graninger, M., Nidetzky, B., Heinrichs, D. E., Whitfield, C., and Messner, P. (1999) Characterization of dTDP-4-dehydrorhamnose 3,5-epimerase and dTDP-4-dehydrorhamnose reductase, required for dTDP-L-rhamnose biosynthesis in Salmonella enterica serovar Typhimurium LT2. *J Biol Chem* **274**, 25069-25077
58. Arcus, V. L., Prentice, E. J., Hobbs, J. K., Mulholland, A. J., Van der Kamp, M. W., Pudney, C. R., Parker, E. J., and Schipper, L. A. (2016) On the Temperature Dependence of Enzyme-Catalyzed Rates. *Biochemistry* **55**, 1681-1688
59. Anderson, V. E. (1992) Isotope effects on enzyme-catalyzed reactions. *Curr Opin Struct Biol* **2**, 757-764
60. Savitsky, P., Bray, J., Cooper, C. D., Marsden, B. D., Mahajan, P., Burgess-Brown, N. A., and Gileadi, O. (2010) High-throughput production of human proteins for crystallization: the SGC experience. *J Struct Biol* **172**, 3-13
61. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-234
62. Kabsch, W. (2010) Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132
63. Tickle, I. J., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., and Bricogne, G. (2018) STARANISO. Global Phasing Ltd.
64. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242
65. Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P. R., Jenkins, H. T., Keegan, R., Krissinel, E., Stevenson, K., Lebedev, A., McNicholas, S. J., Nicholls, R. A.,

- Noble, M., Pannu, N. S., Roth, C., Sheldrick, G., Skubak, P., Turkenburg, J., Uski, V., von Delft, F., Waterman, D., Wilson, K., Winn, M., and Wojdyr, M. (2018) CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr D Struct Biol* **74**, 68-84
66. Vagin, A., and Lebedev, A. (2015) MoRDa, an automatic molecular replacement pipeline. *Acta Crystallogr A* **71**, S19-S19
67. Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**, 22-25
68. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501
69. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-367
70. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21
71. Webb, B., and Sali, A. (2016) Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Protein Sci* **86**, 2 9 1-2 9 37
72. Gouet, P., Robert, X., and Courcelle, E. (2003) ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* **31**, 3320-3323

Figure Legends

Figure 1. Saccharide structures in *Coxiella burnetii* and related sugars. **A.** Overview of the *C. burnetii* lipopolysaccharide and O-antigen. Wild-type *C. burnetii* makes a smooth LPS with the inner and outer cores and a full O-antigen. Mutants cause either an “intermediate” phenotype without virenose in the O-antigen, or a “rough” phenotype lacking the outer core and O-antigen. The linkage of the outer core and O-antigen is not known. Key abbreviations: KDO: 2-keto-3-deoxy-D-mannoctanoic acid; Man: mannose; GlcNAc: *N*-acetyl-glucosamine; Glc: glucose; GalNAc: *N*-acetyl-galactosamine; Vir: virenose; DHHS: dihydrohydroxystreptose. Based on (26). **B, C.** The unusual monosaccharides virenose and dihydrohydroxystreptose (DHHS) from the *C. burnetii* O-antigen (**B**) and the related sugar streptose (**C**). The differences between DHHS and streptose are highlighted with red circles. **D, E.** Proposed biosynthetic pathways for TDP-linked streptose (**D**) and DHHS (**E**). The equivalent *E. coli* enzymes for each step are shown in blue in between the two panels; proposed enzymes for the *S. griseus* (34) and *C. burnetii* (this study) pathways, where known, are shown in black. Chemical structures drawn using BIOVIA Draw v16.1.

Figure 2: Structural alignment of RmlC paralogs. Alignment generated using Modeller v9.21 (71) from the PDB ID codes 2IXC (41), 1OI6 (50) and 4HMZ (44) (*MtRmlC*, *AoEvaD* and *SbChmJ*, respectively), and the PDB files from CBU1838 and SgStrM solved in this study. Chain A was used in each case. Secondary structure elements for *MtRmlC* and *SgStrM* are noted above and below the sequence alignments, respectively. The strand exchange between chains is highlighted with a blue box. The residues that interact with TDP and citrate, and which are likely the catalytic acid and base are indicated with black arrows, sky blue dotted arrows and pink dashed arrows respectively. Figure generated using ESPript v3.0222 (72); a red box/white character relates strict identity, a black bold character relates similarity in a group, and filled in yellow characters relates similarity across groups.

Figure 3: CBU1838 complements *rmlC* in vivo. MG1665 *E. coli*, the Keio collection *rmlC* mutant, and complements of this mutant with *CBU1838* and *AoEvaD* were grown to stationary phase in LB. Cells were harvested and metabolites extracted in 50% (v/v) acetonitrile. Metabolite concentrations were determined by HPLC-MS using a triple-quad instrument. Concentrations were determined by comparison to standards at defined concentrations and normalised to an internal ¹³C labelled TDP-glucose sample added to all samples. Five experimental replicates were taken for each data point. Bars represent the mean. All points for

each compound were compared by ANOVA followed by Tukey's multiple comparison test using Graphpad Prism v.9.2.0. *: $p < 0.05$; ***: $p < 0.0005$; ****: $p < 0.0001$

Figure 4: StrM and CBU1838 complement RmlC *in vitro*. The turnover of TDP-glucose to TDP-rhamnose by RmlB, RmlC and RmlD was monitored by following the reduction of NADPH at 340 nm. 200 μ M each NADPH and TDP-glucose were incubated with 0.4 μ M RmlB and RmlD, and as appropriate 0.4 μ M RmlC, 0.1 μ M StrM, or 1.6 μ M CBU1838. StrM and CBU1838 were able to complement RmlC in this assay, albeit at a reduced rate for CBU1838. Three experimental replicates were taken for each condition.

Figure 5: Rate of RmlC paralogues monitored by NMR. RmlC paralogues were incubated with purified TDP-6-deoxy-D-xylo-4-hexulose. **A.** The loss of proton signal at the 3'' and 5'' positions in the sugar was monitored by NMR. The rate was monitored every 5-10 minutes over 90 minutes to calculate an initial rate. **B.** Samples were incubated over 18 hours to allow the reaction to proceed to completion. RmlC, StrM and CBU1838 showed almost complete conversion to product in both 3'' and 5'' positions. ChmJ in contrast showed conversion only at the 3'' position. Note that the substrate showed almost no loss of signal over the course of the experiment. Error bars represent the standard error in the mean.

Figure 6: StrM and CBU1838 show expected interactions with ligands. **A:** Ribbon diagram of the structure of StrM (left) and CBU1838 (right). Both proteins form the domain-swapped dimer characteristic of the RmlC proteins. The red lines indicate the dimer interfaces. The domain-swapped strands are indicated by black arrowheads. Proteins are shown in cartoon format, coloured blue: protomer 1, N-terminus to red: protomer 2, C-terminus. **B,C:** Structures of both proteins complexed to TDP show a network of interactions conserved with other RmlC enzymes. Interaction maps for TDP with both enzymes were calculated using LigPlus v. 1.4.5. Structures are shown with protein backbone shown as cartoon, with TDP and interacting residues shown as sticks. Colours: nitrogen, blue; oxygen, red; phosphorus, orange; protomer 1 carbon, green; protomer 2 carbon, yellow; TDP carbon, rose. Hydrogen bonds are shown as red dashes. **D:** A citrate molecule occupies the CBU1838 substrate sugar site and makes interactions with key residues. Left: LigPlus summary of interactions. Right: Structural representation of the citrate. The two conserved catalytic residues H60 and Y130 flank the citrate. Representation as in B, with citrate carbons shown in skyblue. Panels A, B and D made using PyMOL v. 2.3.4.

Tables

Table 1: Summary of existing structures of RmlC paralogues. Key published structures of proteins with experimentally determined activities.

^aSpecies abbreviations: Mtu: *Mycobacterium tuberculosis*; Pa: *Pseudomonas aeruginosa*; Ss: *Streptococcus suis*; Mth: *Methanobacterium thermoautotrophicum*; Sty: *Salmonella typhimurium*; Ba: *Bacillus anthracis*; Ao: *Amycolatopsis orientalis*; Sb: *Streptomyces bikiniensis*; Sn: *Streptomyces niveus*; Ye: *Yersinia enterocolitica*; At: *Aneurinibacillus thermoaerophilus*; Tt: *Thermoanaerobacterium thermosaccharolyticum*; Cj: *Campylobacter jejuni*.

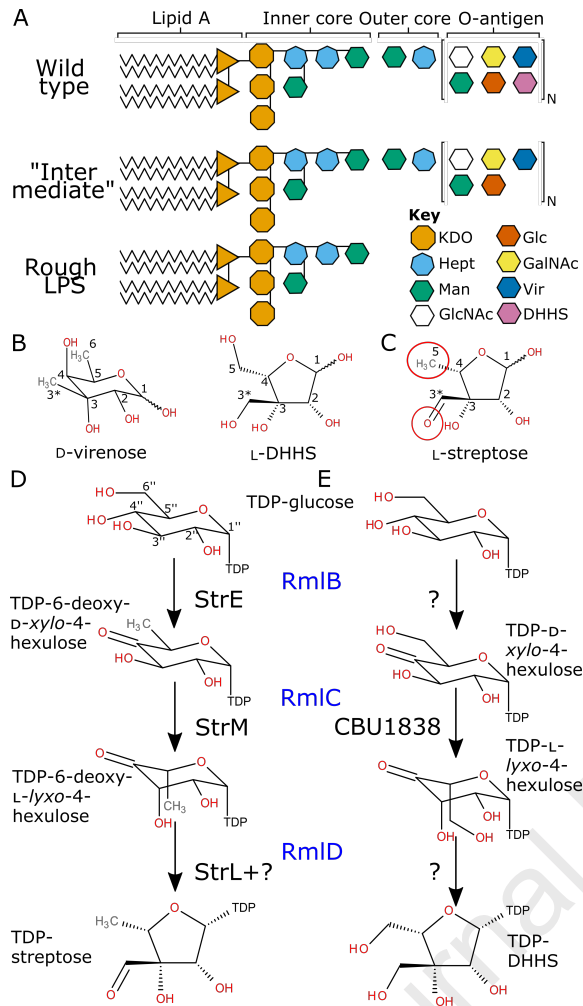
Reaction	Enzyme activity	EC number	Common name	PDB ID	Species ^a	Ref.
1	TDP-6-deoxy-D-xylo-4-hexulose 3'',5''-epimerase	5.1.3.13	RmlC	2IXC/1UPI	Mtu	(41,46)
				2IXH	Pa	(41)
				1NYW	Ss	(53)
				1EP0	Mth	(45)
				1DZR	Sty	(38)
				3RYK	Ba	(48)
2	TDP-6-deoxy-D-xylo-4-hexulose 5''-epimerase	N/A	EvaD	1OI6	Ao	(50)
3	TDP-6-deoxy-D-xylo-4-hexulose 3''-epimerase	5.1.3.27	ChmJ	4HNO	Sb	(44)
			NovW	2C0Z	Sn	(43)
4	CDP-6-deoxy-D-xylo-4-hexulose epimerase	N/A	WbcA	5BUV	Ye	(51)
5	TDP-6-deoxy-3,4,keto-hexulose isomerase	5.3.2.4	FdtA	2PA7	At	(49)
			QdtA	4ZU7	Tt	
			WlaRB	5TPU	Cj	(47)
6	GDP-D-glycero-4-keto-D-lyxo-heptose-3'',5''-epimerase	5.1.3	Cj1430	7M15	Cj	(42)

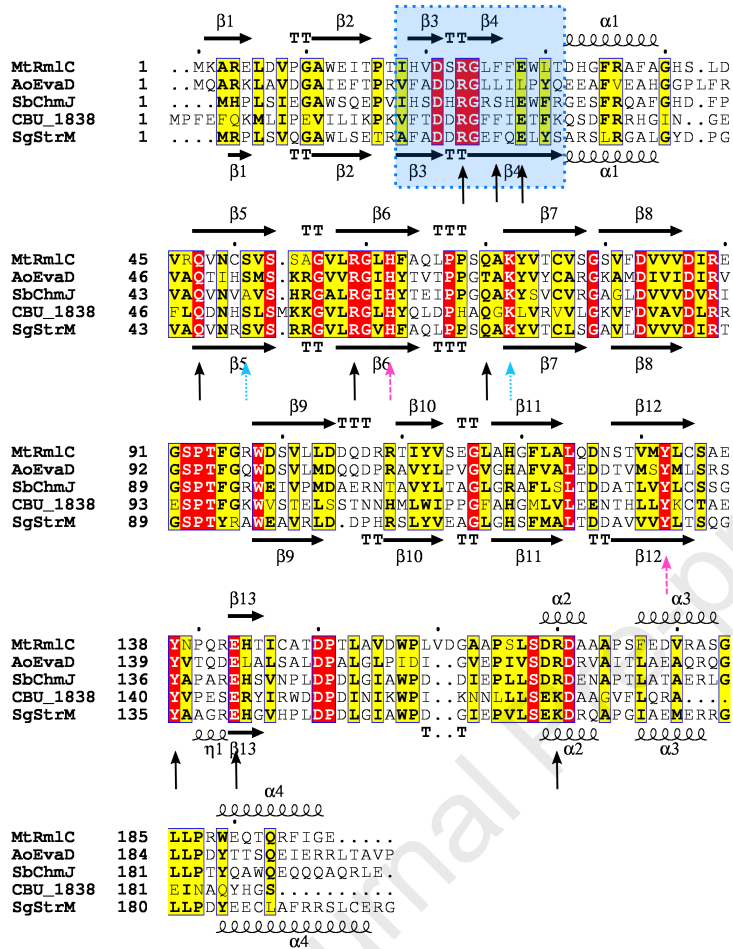
Table 2: Kinetic constants for *EcRmlC*, *SgStrM*, *CbCBU1838* and *AoEvaD*. Kinetic data were determined using commercial TDP-6-deoxy-D-xylo-4-hexulose, with the reaction coupled to *EcRmlD* to allow continuous measurement of the oxidation of NADPH. Data were fitted to the Michaelis-Menten equation using Graphpad v. 8.0 and the fit quality determined from the R² value.

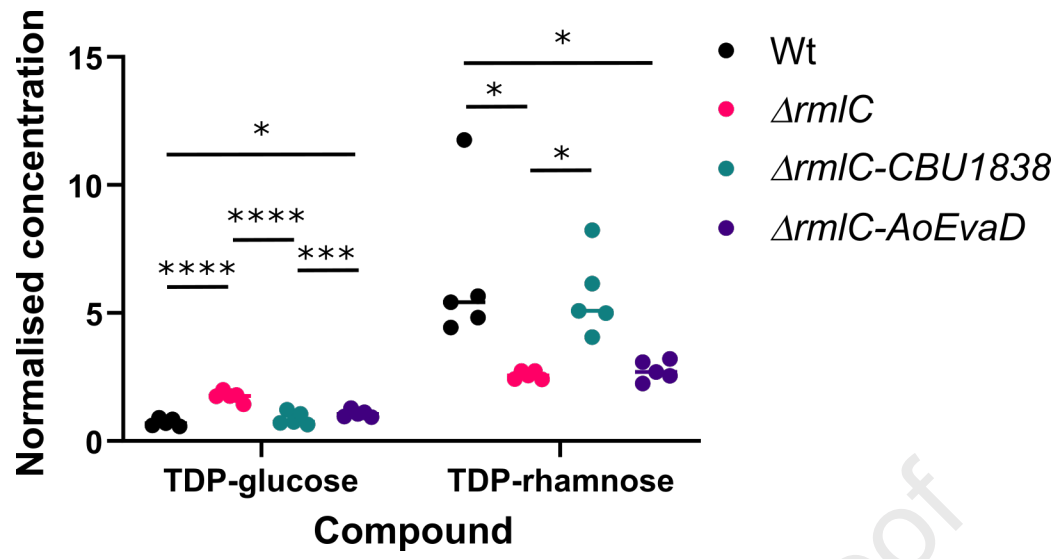
Epimerase	k_{cat} (s ⁻¹)	K_M (μM)	k_{cat}/K_M (s ⁻¹ mM ⁻¹)	R ²
<i>EcRmlC</i>	5.5 ± 0.10	830 ± 40	6.6 ± 0.2	0.996
<i>SgStrM</i>	5.4 ± 0.1	600 ± 40	9.1 ± 0.4	0.988
<i>CBU1838</i>	1.75 ± 0.05	2,000 ± 200	0.88 ± 0.04	0.986
<i>AoEvaD</i>	0.39 ± 0.02	632 ± 98	0.63 ± 0.07	0.921

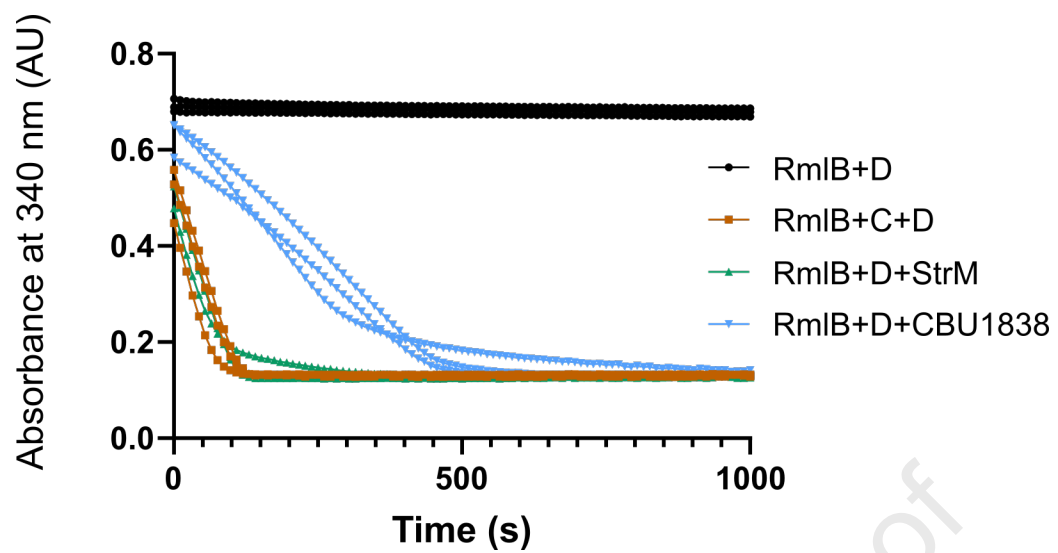
Table 3: Rate of hydrogen-deuterium exchange catalysed by RmlC paralogues. RmlC paralogues were incubated with commercial TDP-6-deoxy-D-xylo-4-hexulose. The loss of proton signal at the 3'' and 5'' positions in the sugar was monitored by NMR. The rate was monitored for up to 90 minutes to calculate an initial rate.

Epimerase	Rate of H-D exchange (s ⁻¹)	
	3''	5''
RmlC	1.86 ± 0.08	0.8 ± 0.1
StrM	1.26 ± 0.04	0.28 ± 0.09
CBU1838	0.87 ± 0.02	0.12 ± 0.03
AoEvaD	0.31 ± 0.01	0.28 ± 0.01
ChmJ	0.48 ± 0.07	0.2 ± 0.1

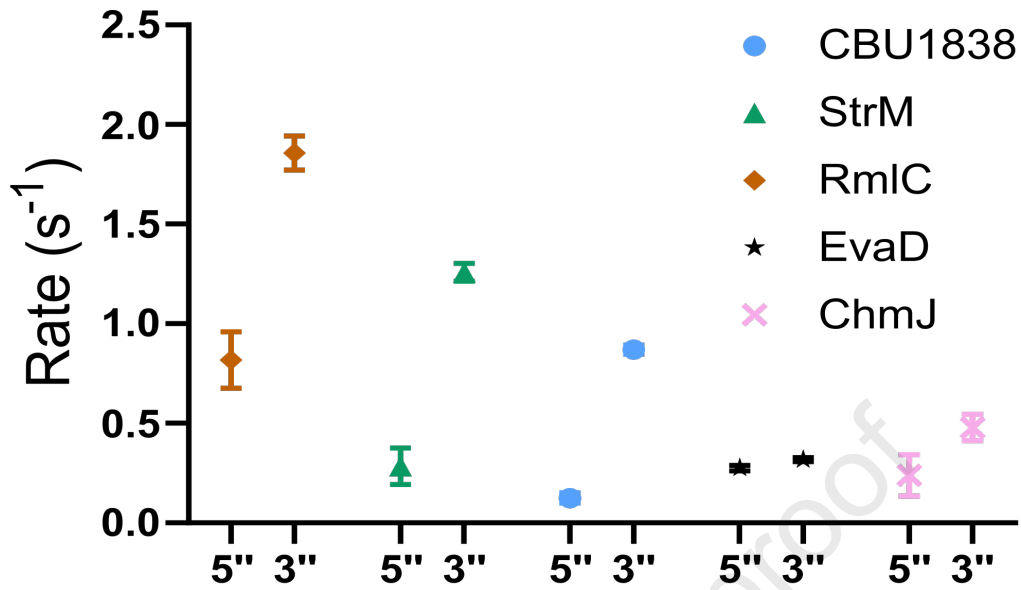




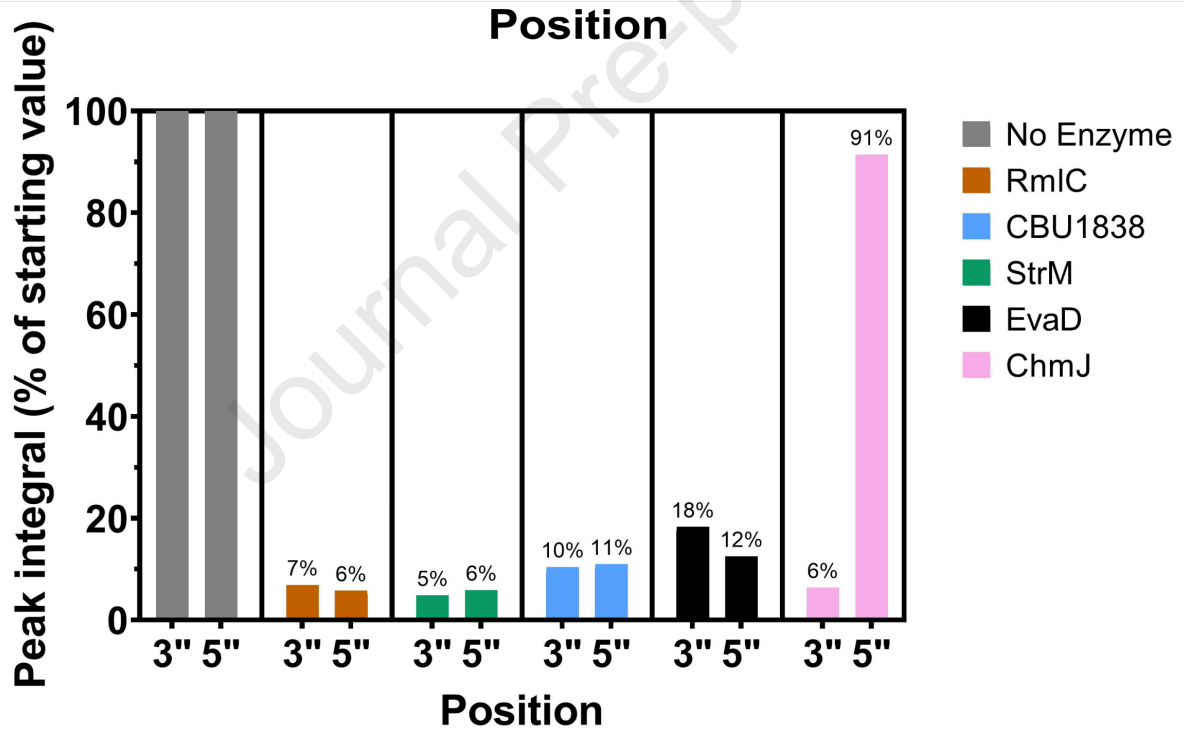


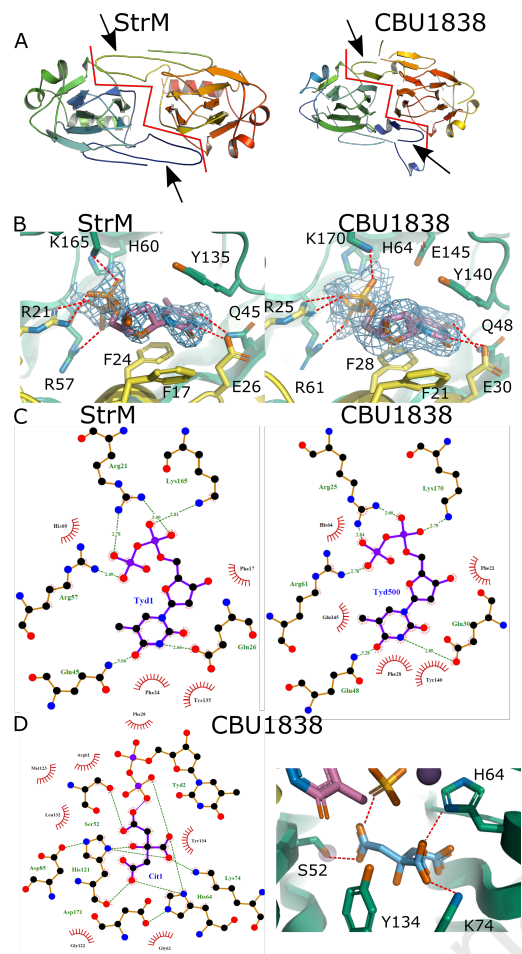


A



B





Spinning sugars in antigen biosynthesis: a direct study of the *Coxiella burnetii* and *Streptomyces griseus* TDP-sugar epimerases

Cross et al.

Author CRediT statement

Alice Cross: Investigation, Formal analysis, Writing – Original Draft, Review & Editing

Sumita Roy: Investigation, Formal analysis, Writing – Review & Editing

Mirella Vivoli Vega: Investigation, Formal analysis, Writing – Review & Editing

Martin Rejzek: Investigation, Formal analysis, Writing – Review & Editing

Sergey Nepogodiev: Investigation, Formal analysis, Writing – Review & Editing

Matthew Cliff: Investigation, Formal analysis, Writing – Review & Editing

Debbie Salmon: Investigation, Formal analysis, Writing – Review & Editing

Michail Isupov: Formal analysis, Data curation, Writing – Review & Editing

Robert A. Field: Writing – Review & Editing, Formal analysis, Supervision, Funding acquisition

Joann L. Prior: Writing – Review & Editing, Supervision, Funding acquisition

Nicholas Harmer: Formal analysis, Data curation, Writing – Original Draft, Review & Editing, Supervision, Funding acquisition

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Nicholas Harmer reports financial support was provided by Biotechnology and Biological Sciences Research Council. Nicholas Harmer reports financial support was provided by UK Research and Innovation. Nicholas Harmer reports financial support was provided by Dstl. Robert Field reports financial support was provided by Biotechnology and Biological Sciences Research Council. Robert Field reports financial support was provided by the John Innes Foundation. Robert Field reports financial support was provided by Innovate UK.