scientific reports

Check for updates

OPEN

Two RmlC homologs catalyze dTDP-4-keto-6-deoxy-D-glucose epimerization in *Pseudomonas putida* KT2440

Franziska Koller & Jürgen Lassak✉

L-Rhamnose is an important monosaccharide both as nutrient source and as building block in prokaryotic glycoproteins and glycolipids. Generation of those composite molecules requires activated precursors being provided e. g. in form of nucleotide sugars such as dTDP-β-L-rhamnose (dTDP-L-Rha). dTDP-L-Rha is synthesized in a conserved 4-step reaction which is canonically catalyzed by the enzymes RmlABCD. An intact pathway is especially important for the fitness of pseudomonads, as dTDP-L-Rha is essential for the activation of the polyproline specific translation elongation factor EF-P in these bacteria. Within the scope of this study, we investigated the dTDP-L-Rha-biosynthesis route of *Pseudomonas putida* KT2440 with a focus on the last two steps. Bioinformatic analysis in combination with a screening approach revealed that epimerization of dTDP-4-keto-6-deoxy-D-glucose to dTDP-4-keto-6-deoxy-L-mannose is catalyzed by the two paralogous proteins PP_1782 (RmlC1) and PP_0265 (RmlC2), whereas the reduction to the final product is solely mediated by PP_1784 (RmlD). Thus, we also exclude the distinct RmlD homolog PP_0500 and the genetically linked nucleoside diphosphate-sugar epimerase PP_0501 to be involved in dTDP-L-Rha formation, other than suggested by certain databases. Together our analysis contributes to the molecular understanding how this important nucleotide-sugar is synthesized in pseudomonads.

Rhamnose (Rha) is a naturally occurring sugar being widely distributed among bacteria and plants[1]. Rha is a component of saponins[2], certain bacterial glycans such as rhamnolipids[3] or mycolic acids[4], extracellular polysaccharides[5] and even cytosolic proteins[6] (Fig. 1A). Incorporation of rhamnose into these compounds requires an activated precursor which is provided as a nucleotide sugar. To date, two forms of activated Rha are known to be produced by bacteria: Guanosine diphosphate-α-D-rhamnose (alternative name: 6-deoxy-α-D-mannose) (GDP-Rha)[7] and deoxythymidine-β-L-rhamnose (dTDP-L-Rha)[8]. While GDP-Rha is synthesized from mannose-1-phosphate[7], the pathway for dTDP-L-Rha starts with glucose-1-phosphate (Glc-1P) (Fig. 1B).

Homologs for the synthesis genes of dTDP-L-Rha, *rmlBDAC*, can be identified in gram-positive and gram-negative bacteria[1] and according to their number the pathway consists of four steps (Fig. 1B)[1]. First, a nucleotide transferase RmlA (also named RfbA[12] or RffH[13]) transfers a deoxythymidine monophosphate moiety from deoxythymidine triphosphate to Glc-1P accompanied by the release of pyrophosphate. In the second step, a dehydratase RmlB (also named RfbB[14] or RffG[13]) catalyzes the conversion of dTDP-glucose into dTDP-4-keto-6-deoxy-D-glucose. The third enzyme—an epimerase RmlC (also named RfbC[15])—mediates a double epimerization reaction leading to the formation of dTDP-4-keto-6-deoxy-L-mannose. Fourth, RmlD (also named RfbD)[15] reduces the C4 keto group of the 4-keto-6-deoxy-L-mannose and with this dTDP-L-Rha synthesis is completed. Notably, the pathway was shown to be critical or even essential for viability in the human pathogens *Streptococcus pyogenes*, *S. mutans*[16] and *Mycobacterium tuberculosis*[17]. In the clinically relevant *Pseudomonas aeruginosa*[18], dTDP-L-Rha is important for the synthesis of rhamnolipids[19]. These are bacterial surfactants with a rhamnose moiety as head group and act as a key virulence determinant[20]. Moreover, in about 10% of all bacteria including pseudomonads, a protein monorhamnosylation was described in 2015 which is essential for activation of the polyproline specific translation elongation factor EF-P[6]. Specifically, the glycosyltransferase EarP transfers a rhamnose moiety onto a conserved EF-P arginine residue R32 thereby utilizing dTDP-L-Rha as a precursor[6,21–23]. In the scope of this study, we investigated the dTDP-L-Rha biosynthesis pathway of *P. putida* KT2440 with focus on the epimerization of TDP-4-keto-6-deoxy-D-glucose. *P. putida* strains in general are fast-growing and genetically easily accessible[24].

Department Biology I, Microbiology, Ludwig-Maximilians-Universität München, Planegg/Martinsried, Germany. ✉email: juergen.lassak@lmu.de

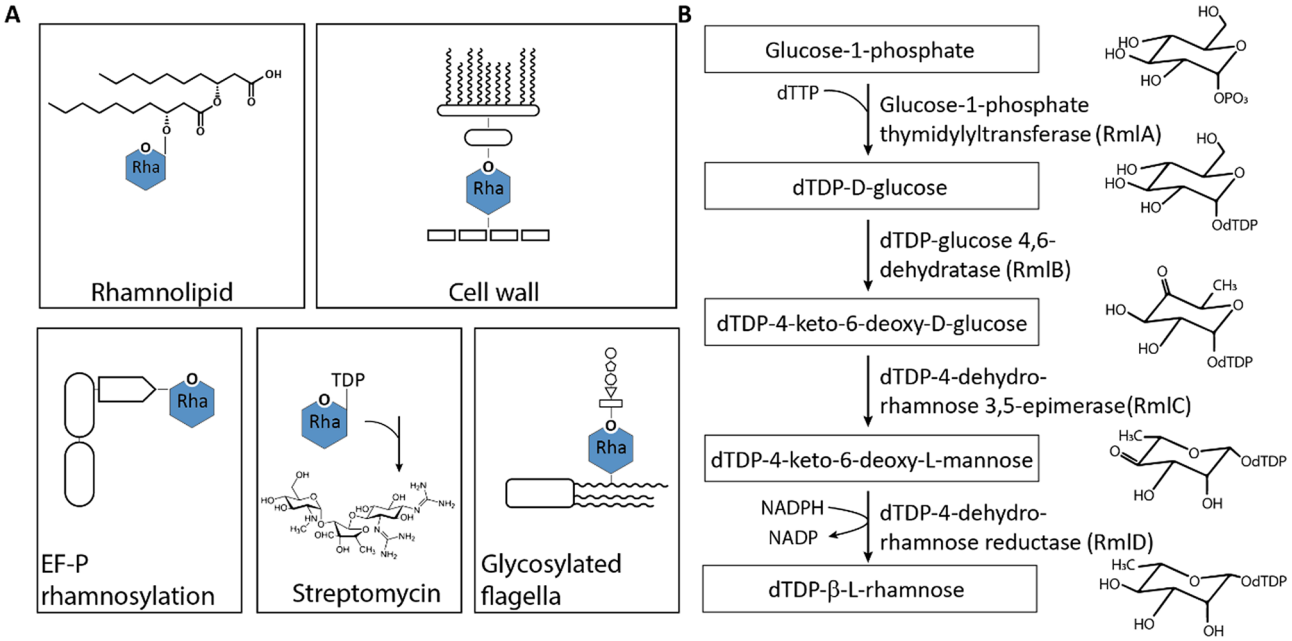

Figure 1. Rhamnose as versatile building block in composite biomolecules. (**A**) L-Rha in bacterial biomolecules. Top left: Rhamnolipids consisting of a rhamnose moiety and a fatty acid tail in *P. aeruginosa*[9]. Top right: Mycobacterial cell wall containing L-Rha as linking sugar between arabinogalactan and peptidoglycan[9]. Bottom left: Rhamnosylation of translation elongation factor EF-P in about 10% of all bacteria[6]. Bottom middle: biosynthesis of Streptomycin inter alia originating from dTDP-L-Rha[10]. Bottom right: Glycosylated flagella with a linking L-Rha moiety in certain *Pseudomonads*[11]. (**B**) dTDP-β-L-rhamnose biosynthesis pathway. Glucose-1-phosphate thymidylyltransferase, the first enzyme of the pathway, transfers a thymidylmonophosphate nucleotide to glucose-1-phosphate, which is further oxidated by dTDP-D-glucose 4,6-dehydratase at the C4 hydroxyl group of the saccharide. The double epimerization reaction at positions C3 and C5 is catalyzed by the dTDP-4-keto-6-deoxy-D-glucose 3,5-epimerase. Finally, the reduction of the C4 keto group by the dTDP-4-keto-6-deoxy-L-mannose reductase leads to dTDP-L-Rha.

They are a paradigm of metabolically versatile microorganisms being able to recycle organic wastes and are key players in the maintenance of environmental quality[24].

Following an unbiased approach and utilizing a restriction based genomic library, we identified the two paralogous proteins PP_1782 (now termed RmlC1) and PP_0265 (now termed RmlC2) as dTDP-4-dehydrorhamnose 3,5-epimerases while the last step namely the reduction to dTDP-L-Rha seems to be solely catalyzed by PP_1784 (RmlD). By contrast, two further candidate genes that were identified by database mining and homology analyses—PP_0500 and PP_0501—are not involved in dTDP-L-Rha biosynthesis. Taken together, our findings contribute to the molecular understanding how dTDP-L-Rha is synthesized in *Pseudomonas putida* KT2440.

Results

A screening system that allows for the discovery of dTDP-L-Rha synthesis genes. To identify genes involved in dTDP-L-Rha biosynthesis, we took advantage of cross functionality of pseudomonal EF-P in *Escherichia coli* and the fact that activation of the translation factor strictly depends on the nucleotide sugar as donor substrate. This cannot necessarily be expected, as the *E. coli* endogenous EF-P significantly differs from its pseudomonal counterpart[25]: although both proteins alleviate ribosome stalling at polyproline stretches[6,26], their modes of activation are phylogenetically unrelated[6,27]. While *E. coli* EF-P (EF-P_{Eco}) strictly depends on (*R*)-β-lysylation[22,28–30] and hydroxylation[31] of a conserved lysine, *Pseudomonas* EF-P (EF-P_{Ppu}) is rhamnosylated at an arginine by the glycosyltransferase EarP (EarP$_{Ppu}$) at the structurally equivalent position[6,21]. Despite these apparent distinct post-translational modifications, a combination of *efp*$_{Ppu}$ and *earP*$_{Ppu}$ from *P. putida* KT2440 can compensate for a lack of *efp* in *E. coli* (Δ*efp*) as long as the endogenous dTDP-L-Rha pathway remains intact (Fig. 2A,C)[6]. Interestingly, loss of any synthesis gene—here exemplified with a Δ*rmlC* strain—does not simply phenocopy Δ*efp* but even results in more severe growth defects, as can be concluded from the corresponding doubling times (Fig. 2B): *E. coli* Δ*efp* cross complemented with *efp*/*earP*$_{Ppu}$ grows twice as fast as the same strain additionally lacking *rmlC* (Δ*efp* Δ*rmlC*). These growth defects are also reflected by the size of the colonies (Fig. 2C). The differences in growth rates provide us with a selection regime to identify dTDP-4-dehydrorhamnose-3,5-epimerase genes from a *P. putida* genomic library.

The library was constructed by partial restriction digestion of the *P. putida* genome with the *dam* and CpG methylation insensitive enzyme *Stu*I (NEB) (Fig. 3). The average fragment size was set to 5 kb to ensure that at least one gene was completely covered (average gene size: 1.132 kbp). These were cloned into *Sma*I linearized pBAD33, which allows for high-level expression by induction of the P$_{BAD}$ promoter with L-arabinose[32].

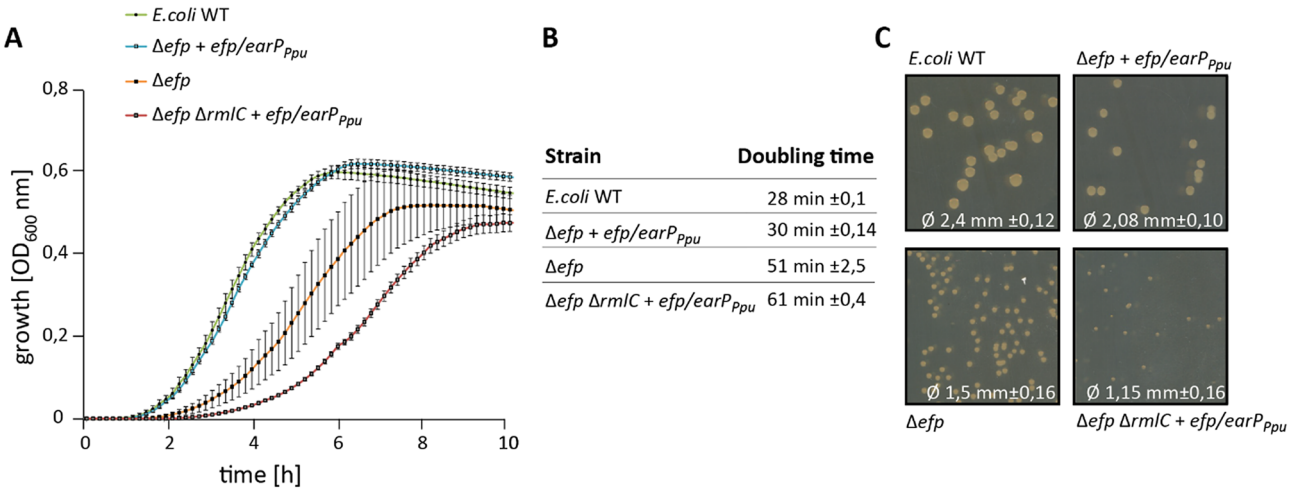

Figure 2. Growth analysis of cross complemented *E.coli* Δ*efp* mutants in dependence of the dTDP-L-Rha pathway. (**A**) *E. coli* MG1655 (*E. coli* WT), *E. coli* MG1655 Δ*efp* expressing pBBR MCS2 *efp*/*earP*$_{Ppu}$ (Δ*efp efp*/*earP*$_{Ppu}$), *E. coli* MG1655 Δ*efp* (Δ*efp*) and *E. coli* MG1655 Δ*efp* Δ*rmlC* expressing pBBR MCS2 *efp*/*earP*$_{Ppu}$ (Δ*efp* Δ*rmlC efp*/*earP*$_{Ppu}$) were grown in LB at 37 °C. Shown is a the mean curve from three independent biological replicates with standard deviations. (**B**) Doubling times of *E. coli* strains listed in A. Doubling times were calculated from growth analysis from three independent biological replicates. (**C**) Comparison of colony size. *E. coli* MG1655 (*E. coli* WT), *E. coli* MG1655 Δ*efp* expressing pBBR MCS2 *efp*/*earP*$_{Ppu}$ (Δ*efp efp*/*earP*$_{Ppu}$), *E. coli* MG1655 Δ*efp* (Δ*efp*) and *E. coli* MG1655 Δ*efp* Δ*rmlC* expressing pBBR MCS2 *efp*/*earP*$_{Ppu}$ (Δ*efp* Δ*rmlC efp*/*earP*$_{Ppu}$) were plated on LB Agar (1.5%). The mean diameter with respective standard deviations is depicted at the bottom. Pictures were taken after o/n growth at 37 °C.

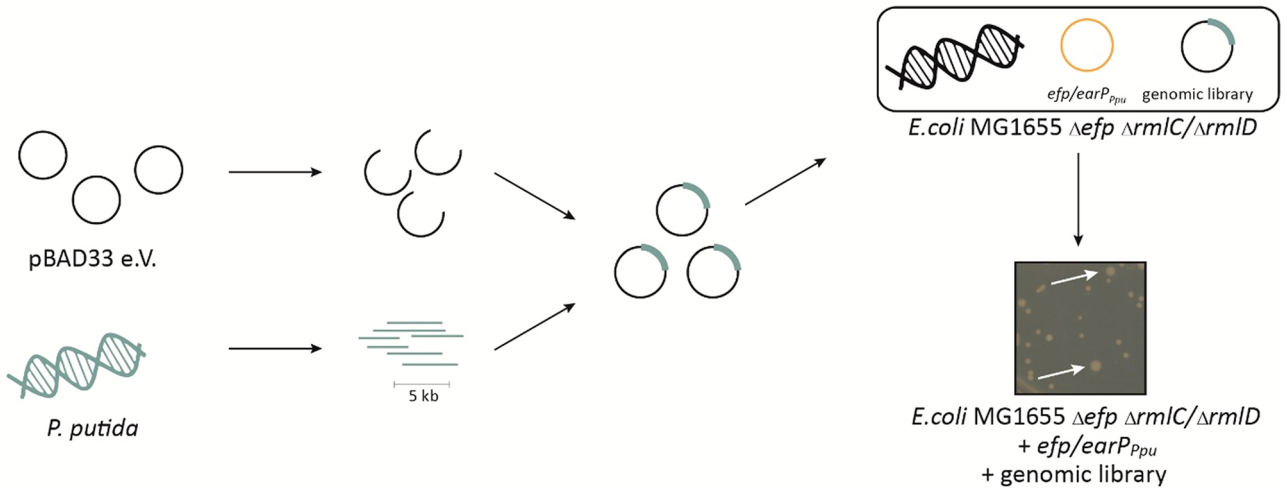

Figure 3. Screening strategy for the identification dTDP-L-Rha biosynthesis genes in *P. putida*. Chromosomal DNA of *P. putida* (green) was fragmented by restriction digestion. Fragments with an average size of 5 kb were then ligated into the arabinose inducible vector (pBAD33). The resulting library was transformed into an *E. coli* Δ*efp* P$_{cadBA}$::*lacZ* reporter strain that concomitantly lacks either *rmlC* or *rmlD* (Δ*rmlC*/ Δ*rmlD*) and additionally encodes *earP*$_{Ppu}$ and *efp*$_{Ppu}$ (orange) of *P. putida* in trans. Genes cross-complementing Δ*rmlC* and Δ*rmlD* recover the impaired growth phenotype and can be selected by size (white arrows).

Transformation of *E. coli* DH10B cells with the library revealed ~ 430,000 clones indicating an around 350-fold coverage of the *P. putida* KT2440 genome (total length 6.18187 Mbp).

Next, we transformed *E. coli* Δ*efp* Δ*rmlC* + *efp*/*earP*$_{Ppu}$ with the library and cultivated the cells in LB (lysogeny broth)[33,34] containing 0.2% L-arabinose. Considering duplication times (Fig. 2B) and genome coverage, we expect *rmlC* copies to accumulate already to a single-digit percentage of the total population within latest two days (= ~ 16 generations with mutant growth phenotype and ~ 32 for wild-type phenotype), even under unfavorable circumstances. Indeed, when plating the second overnight culture on LB agar we obtained colonies of two different sizes. Consequently, we isolated plasmids from 16 large clones and sequencing identified 12 times PP_1782 and four times PP_0265 as the insert. PP_0265 (from now on *rmlC2*/RmlC2) resides next to genes encoding a putative two component signal-transduction system (Fig. 4A). PP_1782 (from now on *rmlC1*/RmlC1) on the

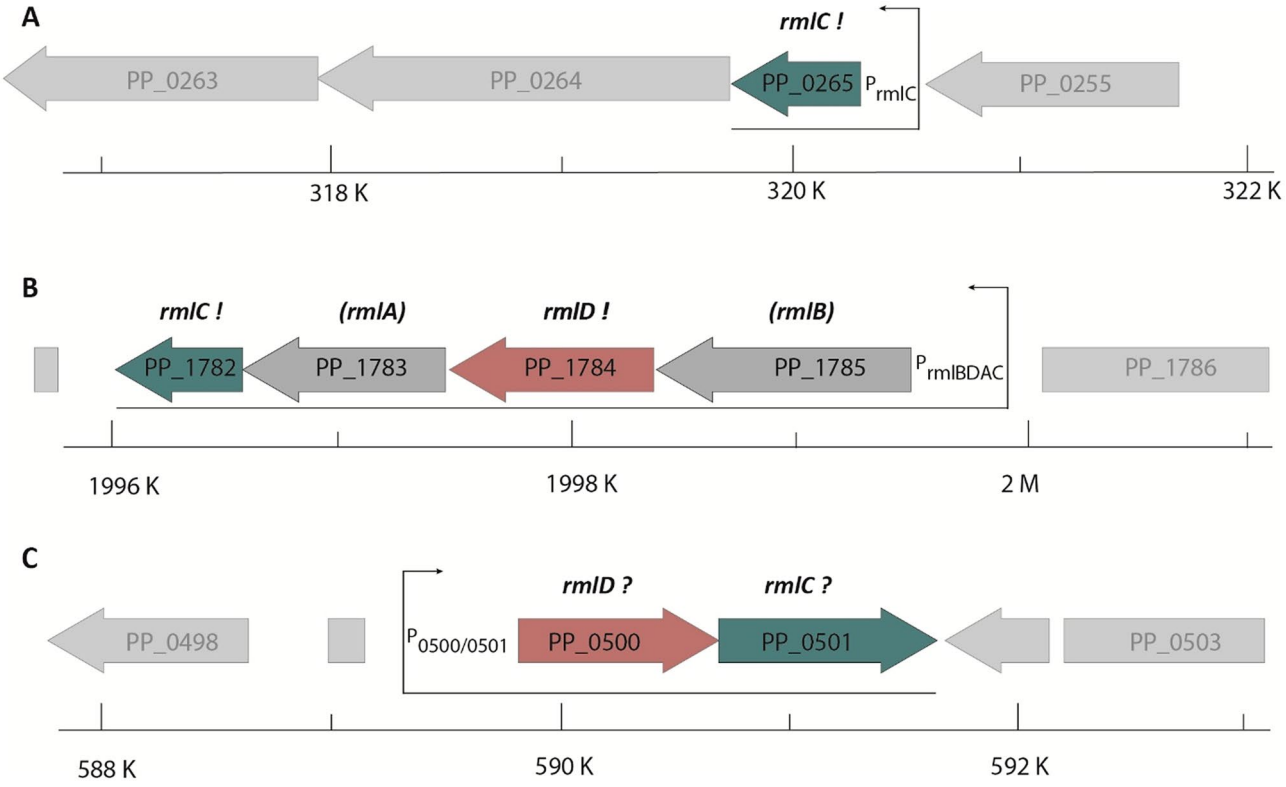

Figure 4. Genomic organization of *rmlC* and *rmlD* candidate genes in *P. putida*. (**A**) PP_0265 gene region, (**B**) PP_1782_PP_1784 gene region. (**C**) PP_0500 and PP_0501 gene region. Putative (?) or validated (!) homologs/analogs of *rmlC* and *rmlD* are shown in green and red respectively. Bottom: position within *P. putida* genome. Arrows indicate monocistrons. The scale indicates the position within the *P. putida* genome.

other hand is the last of four genes in a putative dTDP-L-Rha biosynthesis operon PP_1785-PP_1782 (Fig. 4B). To substantiate our hypothesis on the dTDP-L-Rha biosynthetic operon, we conducted a second library screen with *E. coli* cells now lacking *rmlD* in addition to *efp* (Δ*efp* Δ*rmlD* + *efp*/*earP*$_{Ppu}$) instead of *rmlC*. With this strain we exclusively enriched clones harboring a copy of PP_1784 (from now on *rmlD*/RmlD), a homolog of *E. coli* RmlD. Thus, we provide experimental evidence that PP_1785-PP_1782 form a *rmlBDAC1* operon in *P. putida* KT2440 and further identified a second gene encoding for an dTDP-4-dehydrorhamnose 3,5-epimerases—RmlC2.

PP_0265/PP_1782 and PP_1784 are dTDP-4-dehydrorhamnose 3,5-epimerases and dTDP-4-dehydrorhamnose reductase, respectively. Our library screen was complemented by database mining and a homology search. In addition to *rmlC1* and *rmlC2,* we found PP_0501 being annotated as nucleoside diphosphate sugar epimerase of unknown specificity and as such might function as further dTDP-4-dehydrorhamnose 3,5-epimerase (String[35,36], Pfam[37], Uniprot[38], Metacyc[39] database). However, while RmlC1 & RmlC2 are highly homologous to each other (64% identity), PP_0501 shares no similarities at the sequence level. Nonetheless and in addition to its annotated function PP_0501 forms an operon with a putative dTDP-4-dehydrorhamnose reductase gene, PP_0500[35,36,40] (Fig. 4C). This protein, on the contrary, shares similarities with RmlD both at the sequence level (29% identity) as well as structurally.

To test the putative role of PP_0500 and PP_0501 in dTDP-L-Rha biosynthesis we made again benefit of EarP mediated activation of *P. putida* EF-P and its functionality in *E. coli.* Hence, we cloned the two genes into pBAD33 simultaneously adding a His_6-tag coding sequence for immunodetection in order to ensure proper protein production (Fig. 5). *rmlC1*, *rmlC2* and *rmlD* were also included in the study. The resulting plasmids pBAD33-*rmlC1*, pBAD33-*rmlC2*, pBAD33-PP_0501 as well as pBAD33-*rmlD* and pBAD33-PP_0500 were introduced into *E. coli* Δ*efp* Δ*rmlC* + *efp*/*earP*$_{Ppu}$ and Δ*efp* Δ*rmlD* + *efp*/*earP*$_{Ppu}$, respectively. Of note, these are reporter strains in which EF-P functionality is coupled to LacZ expression (Fig. 5A). Whereas β-galactosidase activity is low in cell with an incomplete dTDP-L-Rha biosynthesis pathway, introduction of either *rmlC1, rmlC2* (Fig. 5B) or *rmlD* (Fig. 5C) into the respective mutant strains led to a significant increase. By contrast, neither PP_0500 nor PP_0501 were able to rescue the Δ*efp*$_{Eco}$ mutant phenotype.

In parallel we analyzed the rhamnosylation status of EF-P$_{Ppu}$ utilizing anti-rhamnosylarginine specific antibodies (anti-Arg^{Rha})[21,41,42]. Immunodetection of EF-P$_{Ppu}$ rhamnosylation matched with the reporter expression levels on the one hand confirming the enzymatic activities of RmlC1, RmlC2 and RmlD as dTDP-4-dehydrorhamnose 3,5-epimerase and dTDP-4-dehydrorhamnose reductase, respectively. On the other hand, they falsify speculation and database annotations that attribute PP_0500 and PP_0501 a function in dTDP-L-Rha biosynthesis (String[35,36], Pfam[37], Uniprot[38], Metacyc[39] database).

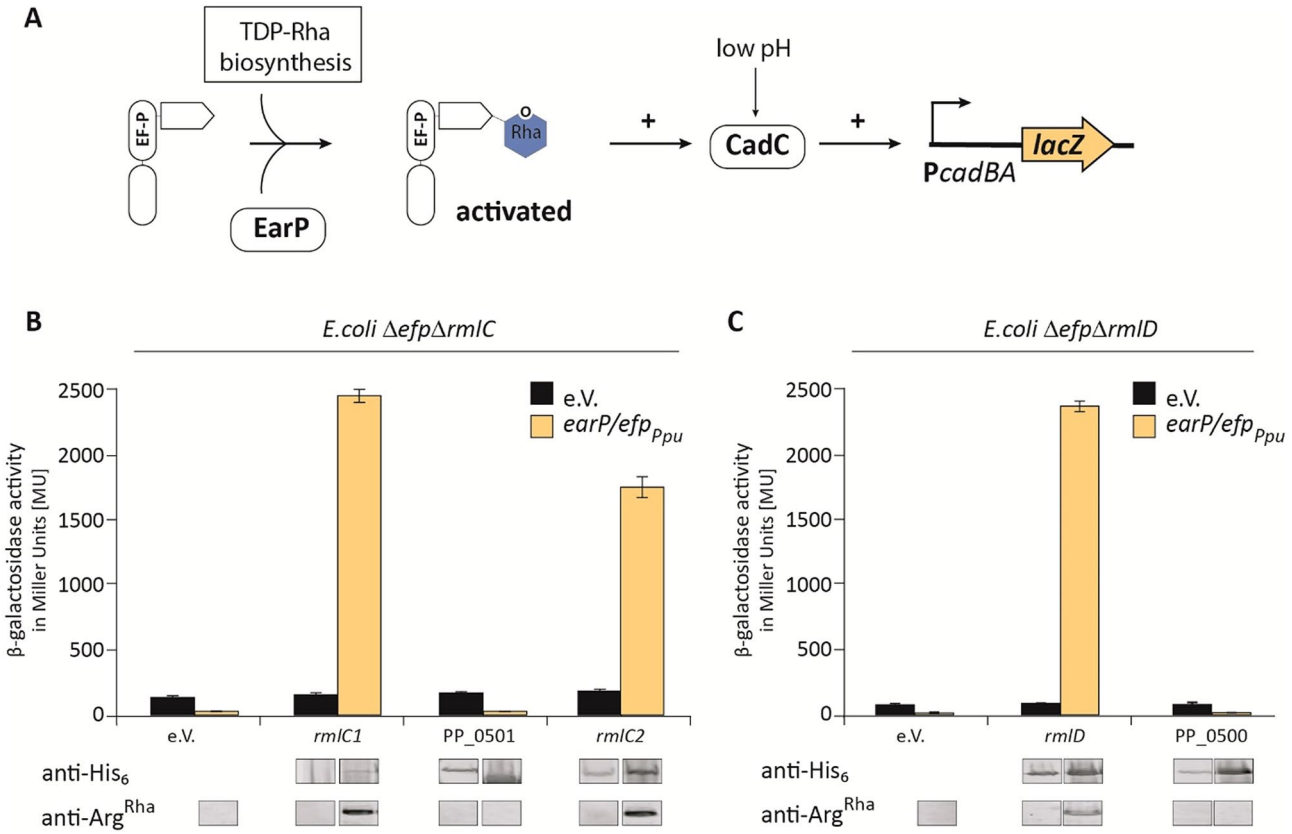

Figure 5. Analysis of in vivo activities of activated EF-P in dTDP-L-Rha biosynthesis deletion strains. (**A**) β-Galactosidase reporter assay. The assay is based on the lysine decarboxylase acid stress response of *E. coli*, the CadABC module[43]. At low pH, the transcriptional activator CadC activates the promoter of its two downstream genes (P_{cadBA}) thereby inducing the expression of *lacZ* in an *E. coli* MG1655 P_{cadBA}::*lacZ* strain. Proper translation of CadC is dependent on the presence of EF-P which is activated by mono-rhamnosylation, a reaction catalyzed by the glycosyltransferase EarP using dTDP-L-Rha (blue) as substrate. Thus β-galactosidase activity can be taken as an indirect readout for functional dTDP-L-Rha biosynthesis. (**B**,**C**) Functionalities of RmlC1, RmlC2, RmlD, PP_0500 and PP_0501 were determined by measuring the β-galactosidase activities of *E. coli* MG1655 P_{cadBA}::*lacZ* Δ*efp* Δ*rmlC* (**B**)/Δ*rmlD* (**C**) with heterologous expression of a candidate gene from the pBAD33 vector. The empty vector (e.V.) was included as negative control. Additionally, all strains encoded the *earP/efp*$_{Ppu}$ operon in trans, being encoded from pBBR MCS2 vectors (grey bars) expressed from the native promoter. Again, the corresponding empty vector served as control (black bars). All strains were grown o/n in LB pH 5.8 and activity is given in Miller Units (MU). Means of three independent measurements are shown. Standard deviations from three independent experiments were determined. Bottom: Western blot analysis of o/n cultures *E. coli* depicted in (**B**) and (**C**). Rhamnosylated *EF-P*$_{Ppu}$ (EF-P^{Rha}) was detected using 0.25 µg/ml anti-Arg^{Rha}. Expression of candidate genes was verified using 0.1 µg/ml anti-His_6. Full-length Western Blots and corresponding SDS-gels are depicted in Fig. S3.

Discussion

In the scope of this study, we have investigated the dTDP-L-Rha pathway of *P. putida* KT2440 with a focus on the epimerization of dTDP-4-keto-6-deoxy-D-glucose. Combining an unbiased approach and utilizing a genomic library, we identified two paralogous proteins RmlC1 and RmlC2. Duplication of *rmlC* is not restricted to *P. putida* KT2440 but certain other pseudomonads such as *P. monteilii*, *P. fulva*, *P. plecoglossicida* or *P. asiatica* harbor also two gene copies. In fact, functional redundancy in the dTDP-L-Rha biosynthesis pathway is nothing unusual. As an example, the two enzymes RffH and RffG of *E. coli* are paralogous to RmlA and RmlB, respectively[13]. Such duplications may be useful, e.g., to compensate for bottleneck reactions in the dTDP-L-Rha biosynthesis[44]. Such bottlenecks can occur at different stages as the pathway is not only utilized to ultimately generate dTDP-L-Rha. Specifically, dTDP-4-keto-6-deoxy-D-glucose is also a precursor of dTDP-3-acetamido-α-D-fucose[45] and TDP-D-viosamine[46] which are found as part of the glycan pattern in *P. syringae*[47]. Similarly, the two paralogs RmlC1 and RmlC2 in *P. putida* KT2440 might serve as starting point of similar but so far unknown reactions. Moreover, gene duplications open the gate for regulated expression in turn allowing the precise adjustment of the desired ratio of distinct NDP-sugars depending on parts of the dTDP-L-Rha biosynthesis pathway. It would also allow for the accumulation of educts or products of the preceding reactions such as dTDP-glucose and Glc-1P. Notably, whereas *rmlC1* is part of an operon in which presumably the full dTDP-L-Rha pathway is encoded, the *rmlC2* resides in the vicinity of two genes encoding a two-component system (TCS) of thus far unknown

	Feature/genotype	References
Plasmid		
pBAD33	CamR-cassette, p15A origin, araC coding sequence, ara operator	[32]
pBBR1MCS2	KanR-cassette, pBBR origin of replication, *oriT*	[50]
pBAD33_*rmlC1*	CamR-cassette, arabinose inducible expression of RmlC1	This study
pBAD33_*rmlD*	CamR-cassette, arabinose inducible expression of RmlD	This study
pBAD33_PP_0265	CamR-cassette, arabinose inducible expression of PP_0265	This study
pBAD33_PP_0500	CamR-cassette, arabinose inducible expression of PP_0500	This study
pBAD33_PP_0501	CamR-cassette, arabinose inducible expression of PP_0501	This study
pBBR1MCS2_*earP_efp*	KanR-cassette, *earP* and *efp* including the P_{earP} native operon promoter	[6,21]
Strain		
E. coli DH5αλpir	F-λ-endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF) U169, hsdR17(rK– mK+)	[51]
E. coli DH10B	F– *mcrA* Δ(*mrr-hsdRMS-mcrBC*) φ80*lacZ*ΔM15 Δ*lacX74 recA1 endA1 araD139* Δ (*ara-leu*)7697 *galU galK* λ– *rpsL*(Str^R) *nupG*	[52]
E. coli MG1655	K-12 F– λ– ilvG– rfb-50 rph-1	[53]
E. coli PcadBA::*lacZ* Δ*efp*	MG1655 PcadBA::*lacZ* Δ(cadBA) Δ*efp*	[26]
E. coli PcadBA::*lacZ* Δ*efp* Δ*rmlC*	MG1655 PcadBA::*lacZ* Δ(cadBA) Δ*efp* Δ*rmlC*	[26]
E.coli PcadBA::*lacZ* Δ*efp* Δ*rmlD*	MG1655 PcadBA::*lacZ* Δ(cadBA) Δ*efp* Δ*rmlD*	[26]

Table 1. Plasmids and strains used in this study.

function. Based on the predicated domain composition, this specific TCS presumably transduces external signals into gene transcription. One might therefore speculate on regulated expression of *rmlC2* according to the environmental conditions.

While our genomic library revealed two RmlC paralogs in *P. putida* database mining indicated a further enzyme with similar activity PP_0501. However, our in vivo rhamnosylation assay disproved the initial hypothesis. Notably, the UDP-*N*-acetylglucosamine C4-epimerase PelX from *P. protegens* Pf-5 is structurally the closest homolog (identity 67%)[48]. PelX is involved in the biosynthesis of the GalNAc-rich bacterial polysaccharidepolysaccharide Pel, that is essential for pellicle biofilm formation[48,49]. One can hence hypothesize, that PP_0501 and the adjacent putative reductase PP_0500 might be involved in that pathway, instead.

Material and methods

Bacterial strains and growth condition. All strains and plasmids used in this study are listed and described in Table 1. *E. coli* cells were grown in Miller modified Lysogeny Broth (LB)[33,34] at 37 °C aerobically under agitation, if not indicated otherwise. LB agar plates contained 1.5% agar. Mean diameters were measured from 20 colonies from 2 different LB agar plates from of the respective strain after incubation at 37 °C for 16 h. Growth measurements were conducted in 96 well plates. Therefore, 200 µl LB was inoculated with o/n cultures at an OD_{600} 0.001. OD_{600} was monitored in 10-min intervals for 12 h in a Tecan Spark with 240 rpm at 37 °C. The medium was supplemented with antibiotics at the following concentrations: 50 µg/ml kanamycin sulfate and 30 µg/ml chloramphenicol. Plasmids carrying the P_{BAD} promoter[32] were induced with L-arabinose at a final concentration of 0.2% (w/v).

Molecular biology methods. Oligonucleotides used in this study are listed and described in the Supplementary Table S1. Plasmid DNA was isolated using the Hi Yield Plasmid Mini Kit from Süd Laborbedarf according to manufacturer's instructions. DNA fragments were purified from agarose gels using the Hi Yield Gel/PCR DNA fragment extraction kit from Süd Laborbedarf. All restriction enzymes, DNA modifying enzymes and the Q5 high fidelity DNA polymerase for PCR amplification were purchased from New England BioLabs and used according to manufacturer's instructions.

Genomic library. The genomic DNA (gDNA) was isolated from 50 ml o/n culture of *P putida* KT2440 according to the protocol described in reference[54]. Further purification was achieved using Phase Lock Gel (QuantaBio) with Phenol–Chloroform. After the centrifugation, isopropanol precipitation was repeated. The pellet was resuspended in water, the final amount was 60 µg DNA.

Plasmid DNA was purified as described in "Molecular biology methods" from 12 ml *E. coli* DH5α cells. The plasmid DNA was diluted in water, the final amount was 10 µg DNA.

The library was constructed using *Sma*I (pBAD33 vector) and *Stu*I (gDNA) for digestion resulting in an average size of 5 kb per insert (Bionexus, Inc.). After ligation, the plasmids were transformed into *E. coli* DH10 B (Lucigen). Quality control was done by restriction digest of library clones with *BamH*I. All restriction enzymes were produced by New England Biolabs, Frankfurt. The library was reisolated from *E. coli* DH10B as described in "Molecular biology methods" and transferred into corresponding reporter strains.

Bioinformatic tools. The multiple sequence alignment was generated using NCBI BLAST[55,56] and Clustal Omega[57]. Candidate homologues were identified and analysed using String[35,36], Pfam[37], Uniprot[38], Metacyc[39] databases. Protein structures were predicted using Phyrre2[58]. Illustrations were generated with UCSF Chimera[59].

B-Galactosidase assay. *E. coli* MG1655 P_{cadBA}::*lacZ* Δ*efp* Δ*rmlC*/Δ*rmlD* expressing lacZ under the control of the cadBA promoter were grown in buffered LB (pH 5.8) overnight (o/n) and harvested by centrifugation. β-Galactosidase activities were determined as described in reference in biological triplicates and are given in Miller units (MU)[60]. Standard deviations from three independent experiments were determined.

SDS-PAGE and western blotting. Electrophoretic separation of proteins was carried out using 12.5% SDS-PAGE as described by Laemmli[61]. Separated proteins were visualized in gel using 0.5% (vol/vol) 2-2-2-trichloroethanol[62] and detected within a Gel Doc EZ gel documentation system (Bio-Rad). The proteins were transferred onto a nitrocellulose membrane by vertical Western blotting (4 °C). Antigens were detected using 0.1 g/ml anti-His_6 tag (Abcam, Inc.) or 0.25 g/ml of anti-Arg^{Rha}[41]. Primary antibodies (rabbit) were the targeted using 0.1 µg/ml anti-rabbit IgG (IRDye 680RD) (donkey) antibodies (Abcam). Target proteins were visualized via Odyssey CLx Imaging System (LI-COR, Inc).

Received: 11 February 2021; Accepted: 26 May 2021
Published online: 07 June 2021

References

1. Giraud, M. F. & Naismith, J. H. The rhamnose pathway. *Curr. Opin. Struct. Biol.* **10**, 687–696. https://doi.org/10.1016/s0959-440x(00)00145-7 (2000).
2. Moses, T., Papadopoulou, K. K. & Osbourn, A. Metabolic and functional diversity of saponins, biosynthetic intermediates and semi-synthetic derivatives. *Crit. Rev. Biochem. Mol. Biol.* **49**, 439–462. https://doi.org/10.3109/10409238.2014.953628 (2014).
3. Soberón-Chávez, G., González-Valdez, A., Soto-Aceves, M. P. & Cocotl-Yañez, M. Rhamnolipids produced by *Pseudomonas*: From molecular genetics to the market. *Microb. Biotechnol.* **14**, 136–146. https://doi.org/10.1111/1751-7915.13700 (2021).
4. Ma, Y. *et al.* Drug targeting *Mycobacterium tuberculosis* cell wall synthesis: Genetics of dTDP-rhamnose synthetic enzymes and development of a microtiter plate-based screen for inhibitors of conversion of dTDP-glucose to dTDP-rhamnose. *Antimicrob. Agents Chemother.* **45**, 1407–1416. https://doi.org/10.1128/aac.45.5.1407-1416.2001 (2001).
5. Gao, M., D'Haeze, W., De Rycke, R., Wolucka, B. & Holsters, M. Knockout of an azorhizobial dTDP-L-rhamnose synthase affects lipopolysaccharide and extracellular polysaccharide production and disables symbiosis with *Sesbania rostrata*. *Mol. Plant Microbe Interact.* **14**, 857–866. https://doi.org/10.1094/mpmi.2001.14.7.857 (2001).
6. Lassak, J. *et al.* Arginine-rhamnosylation as new strategy to activate translation elongation factor P. *Nat. Chem. Biol.* **11**, 266–270. https://doi.org/10.1038/nchembio.1751 (2015).
7. Kneidinger, B. *et al.* Identification of two GDP-6-deoxy-D-lyxo-4-hexulose reductases synthesizing GDP-D-rhamnose in *Aneurinibacillus thermoaerophilus* L420–91T. *J. Biol. Chem.* **276**, 5577–5583. https://doi.org/10.1074/jbc.M010027200 (2001).
8. Shibaev, V. N. Biosynthesis of bacterial polysaccharide chains composed of repeating units. *Adv. Carbohydr. Chem. Biochem.* **44**, 277–339. https://doi.org/10.1016/s0065-2318(08)60080-3 (1986).
9. Wild, M., Caro, A. D., Hernández, A. L., Miller, R. M. & Soberón-Chávez, G. Selection and partial characterization of a *Pseudomonas aeruginosa* mono-rhamnolipid deficient mutant. *FEMS Microbiol. Lett.* **153**, 279–285. https://doi.org/10.1111/j.1574-6968.1997.tb12586.x (1997).
10. Wahl, H. P. & Grisebach, H. 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. https://doi.org/10.1016/0005-2744(79)90291-2 (1979).
11. Schirm, M. *et al.* Structural and genetic characterization of glycosylation of type a flagellin in *Pseudomonas aeruginosa*. *J. Bacteriol.* **186**, 2523–2531. https://doi.org/10.1128/jb.186.9.2523-2531.2004 (2004).
12. Blankenfeldt, W., Asuncion, M., Lam, J. S. & Naismith, J. H. The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidylyltransferase (RmlA). *EMBO J.* **19**, 6652–6663. https://doi.org/10.1093/emboj/19.24.6652 (2000).
13. Marolda, C. L. & Valvano, M. A. Genetic analysis of the dTDP-rhamnose biosynthesis region of the *Escherichia coli* VW187 (O7:K1) *rfb* gene cluster: Identification of functional homologs of *rfbB* and *rfbA* in the *rff* cluster and correct location of the *rffE* gene. *J. Bacteriol.* **177**, 5539–5546. https://doi.org/10.1128/jb.177.19.5539-5546.1995 (1995).
14. Allard, S. T., Giraud, M. F., Whitfield, C., Messner, P. & Naismith, J. H. The purification, crystallization and structural elucidation of dTDP-D-glucose 4,6-dehydratase (RmlB), the second enzyme of the dTDP-L-rhamnose synthesis pathway from *Salmonella enterica* serovar typhimurium. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 222–225. https://doi.org/10.1107/s0907444999016200 (2000).
15. Graninger, M., Nidetzky, B., Heinrichs, D. E., Whitfield, C. & Messner, P. 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. https://doi.org/10.1074/jbc.274.35.25069 (1999).
16. van der Beek, S. L. *et al.* Streptococcal dTDP-L-rhamnose biosynthesis enzymes: Functional characterization and lead compound identification. *Mol. Microbiol.* **111**, 951–964. https://doi.org/10.1111/mmi.14197 (2019).
17. Ma, Y., Pan, F. & McNeil, M. Formation of dTDP-rhamnose is essential for growth of mycobacteria. *J. Bacteriol.* **184**, 3392–3395. https://doi.org/10.1128/jb.184.12.3392-3395.2002 (2002).
18. Alhede, M., Bjarnsholt, T., Givskov, M. & Alhede, M. *Pseudomonas aeruginosa* biofilms: Mechanisms of immune evasion. *Adv. Appl. Microbiol.* **86**, 1–40. https://doi.org/10.1016/b978-0-12-800262-9.00001-9 (2014).
19. Aguirre-Ramírez, M., Medina, G., González-Valdez, A., Grosso-Becerra, V. & Soberón-Chávez, G. The *Pseudomonas aeruginosa* rmlBDAC operon, encoding dTDP-L-rhamnose biosynthetic enzymes, is regulated by the quorum-sensing transcriptional regulator RhlR and the alternative sigma factor σS. *Microbiology (Reading, England)* **158**, 908–916. https://doi.org/10.1099/mic.0.054726-0 (2012).
20. Zulianello, L. *et al.* Rhamnolipids are virulence factors that promote early infiltration of primary human airway epithelia by *Pseudomonas aeruginosa*. *Infect. Immun.* **74**, 3134–3147. https://doi.org/10.1128/iai.01772-05 (2006).
21. Krafczyk, R. *et al.* Structural basis for EarP-mediated arginine glycosylation of translation elongation factor EF-P. *MBio* https://doi.org/10.1128/mBio.01412-17 (2017).

22. Yanagisawa, T., Sumida, T., Ishii, R., Takemoto, C. & Yokoyama, S. A paralog of lysyl-tRNA synthetase aminoacylates a conserved lysine residue in translation elongation factor P. *Nat. Struct. Mol. Biol.* **17**, 1136–1143. https://doi.org/10.1038/nsmb.1889 (2010).
23. Rajkovic, A. *et al.* Cyclic Rhamnosylated Elongation Factor P Establishes Antibiotic Resistance in *Pseudomonas aeruginosa*. *MBio* **6**, e00823. https://doi.org/10.1128/mBio.00823-15 (2015).
24. Martins dos Santos, V. A. P., Timmis, K. N., Tümmler, B. & Weinel, C. In *Pseudomonas: Volume 1 Genomics Life Style and Molecular Architecture* (ed. Ramos, J.-L.) 77–112 (Springer US, 2004).
25. Choi, S. & Choe, J. Crystal structure of elongation factor P from *Pseudomonas aeruginosa* at 1.75 Å resolution. *Proteins* **79**, 1688–1693. https://doi.org/10.1002/prot.22992 (2011).
26. Ude, S. *et al.* Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science* **339**, 82–85. https://doi.org/10.1126/science.1228985 (2013).
27. Lassak, J., Wilson, D. N. & Jung, K. Stall no more at polyproline stretches with the translation elongation factors EF-P and IF-5A. *Mol. Microbiol.* **99**, 219–235. https://doi.org/10.1111/mmi.13233 (2016).
28. Navarre, W. W. *et al.* PoxA, yjeK, and elongation factor P coordinately modulate virulence and drug resistance in *Salmonella enterica*. *Mol. Cell.* **39**, 209–221. https://doi.org/10.1016/j.molcel.2010.06.021 (2010).
29. Gilreath, M. S. *et al.* β-Lysine discrimination by lysyl-tRNA synthetase. *FEBS Lett.* **585**, 3284–3288. https://doi.org/10.1016/j.febslet.2011.09.008 (2011).
30. Pfab, M. *et al.* Synthetic post-translational modifications of elongation factor P using the ligase EpmA. *FEBS J.* **288**, 663–677. https://doi.org/10.1111/febs.15346 (2021).
31. Peil, L. *et al.* Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 15265–15270. https://doi.org/10.1073/pnas.1310642110 (2013).
32. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J. Bacteriol.* **177**, 4121–4130. https://doi.org/10.1128/jb.177.14.4121-4130.1995 (1995).
33. Bertani, G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* **62**, 293–300. https://doi.org/10.1128/jb.62.3.293-300.1951 (1951).
34. Bertani, G. Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *J. Bacteriol.* **186**, 595–600. https://doi.org/10.1128/jb.186.3.595-600.2004 (2004).
35. Snel, B., Lehmann, G., Bork, P. & Huynen, M. A. STRING: A web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res.* **28**, 3442–3444. https://doi.org/10.1093/nar/28.18.3442 (2000).
36. Szklarczyk, D. *et al.* STRING v11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613. https://doi.org/10.1093/nar/gky1131 (2019).
37. El-Gebali, S. *et al.* The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427–D432. https://doi.org/10.1093/nar/gky995 (2019).
38. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515. https://doi.org/10.1093/nar/gky1049 (2019).
39. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes—A 2019 update. *Nucleic Acids Res.* **48**, D445–D453. https://doi.org/10.1093/nar/gkz862 (2020).
40. Seo, J., Brencic, A. & Darwin, A. J. Analysis of secretin-induced stress in *Pseudomonas aeruginosa* suggests prevention rather than response and identifies a novel protein involved in secretin function. *J. Bacteriol.* **191**, 898–908. https://doi.org/10.1128/jb.01443-08 (2009).
41. Li, X. *et al.* Resolving the α-glycosidic linkage of arginine-rhamnosylated translation elongation factor P triggers generation of the first Arg(Rha) specific antibody. *Chem. Sci.* **7**, 6995–7001. https://doi.org/10.1039/c6sc02889f (2016).
42. Gast, D. *et al.* A set of rhamnosylation-specific antibodies enables detection of novel protein glycosylations in bacteria. *Org. Biomol. Chem.* **18**, 6823–6828. https://doi.org/10.1039/d0ob01289k (2020).
43. Tetsch, L., Koller, C., Haneburger, I. & Jung, K. The membrane-integrated transcriptional activator CadC of *Escherichia coli* senses lysine indirectly via the interaction with the lysine permease LysP. *Mol. Microbiol.* **67**, 570–583. https://doi.org/10.1111/j.1365-2958.2007.06070.x (2008).
44. Brandis, G. & Hughes, D. The SNAP hypothesis: Chromosomal rearrangements could emerge from positive selection during niche adaptation. *PLoS Genet.* **16**, e1008615. https://doi.org/10.1371/journal.pgen.1008615 (2020).
45. Pföstl, A. *et al.* Biosynthesis of dTDP-3-acetamido-3,6-dideoxy-alpha-D-glucose. *Biochem. J.* **410**, 187–194. https://doi.org/10.1042/bj20071044 (2008).
46. Pandey, R. P., Parajuli, P., Gurung, R. B. & Sohng, J. K. Donor specificity of YjiC glycosyltransferase determines the conjugation of cytosolic NDP-sugar in *in vivo* glycosylation reactions. *Enzyme Microb. Technol.* **91**, 26–33. https://doi.org/10.1016/j.enzmictec.2016.05.006 (2016).
47. Yamamoto, M. *et al.* Identification of genes involved in the glycosylation of modified viosamine of flagellins in *Pseudomonas syringae* by mass spectrometry. *Genes* **2**, 788–803. https://doi.org/10.3390/genes2040788 (2011).
48. Marmont, L. S. *et al.* PelX is a UDP-*N*-acetylglucosamine C4-epimerase involved in Pel polysaccharide-dependent biofilm formation. *J. Biol. Chem.* **295**, 11949–11962. https://doi.org/10.1074/jbc.RA120.014555 (2020).
49. Friedman, L. & Kolter, R. Genes involved in matrix formation in *Pseudomonas aeruginosa* PA14 biofilms. *Mol. Microbiol.* **51**, 675–690. https://doi.org/10.1046/j.1365-2958.2003.03877.x (2004).
50. Kovach, M. E. *et al.* Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**, 175–176. https://doi.org/10.1016/0378-1119(95)00584-1 (1995).
51. Macinga, D. R., Parojcic, M. M. & Rather, P. N. Identification and analysis of aarP, a transcriptional activator of the 2′-*N*-acetyltransferase in *Providencia stuartii*. *J. Bacteriol.* **177**, 3407–3413. https://doi.org/10.1128/jb.177.12.3407-3413.1995 (1995).
52. Durfee, T. *et al.* The complete genome sequence of *Escherichia coli* DH10B: Insights into the biology of a laboratory workhorse. *J. Bacteriol.* **190**, 2597–2606. https://doi.org/10.1128/jb.01695-07 (2008).
53. Guyer, M. S., Reed, R. R., Steitz, J. A. & Low, K. B. Identification of a sex-factor-affinity site in *E. coli* as gamma delta. *Cold Spring. Harb. Symp. Quant. Biol.* **45**, 135–140. https://doi.org/10.1101/sqb.1981.045.01.022 (1981).
54. Pospiech, A. & Neumann, B. A versatile quick-prep of genomic DNA from gram-positive bacteria. *Trends Genet.* **11**, 217–218. https://doi.org/10.1016/s0168-9525(00)89052-6 (1995).
55. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **44**, D7–D19. https://doi.org/10.1093/nar/gkv1290 (2016).
56. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410. https://doi.org/10.1016/s0022-2836(05)80360-2 (1990).
57. Sievers, F. & Higgins, D. G. Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol. Biol.* **1079**, 105–116. https://doi.org/10.1007/978-1-62703-646-7_6 (2014).
58. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858. https://doi.org/10.1038/nprot.2015.053 (2015).
59. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612. https://doi.org/10.1002/jcc.20084 (2004).
60. Miller, J. H. *A Short Course in Bacterial Genetics. A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria* (Cold Spring Harbor, 1992).

61. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685. https://doi.org/10.1038/227680a0 (1970).
62. Ladner, C. L., Yang, J., Turner, R. J. & Edwards, R. A. Visible fluorescent detection of proteins in polyacrylamide gels without staining. *Anal. Biochem.* **326**, 13–20. https://doi.org/10.1016/j.ab.2003.10.047 (2004).

Acknowledgements
This work was funded by the Deutsche Forschungsgemeinschaft research grant LA 3658/1-1 and Research Training Group GRK2062/1 (Molecular Principles of Synthetic Biology). We thank Ralph Krafczyk and Kirsten Jung for fruitful discussions. Further we thank Lis Winter, Maximilian Dorok and Martin Erhard for their contributing results.

Author contributions
All experiments were performed by F.K. F.K. and J.L. designed the study. The manuscript was written by F.K. and J.L.

Funding
Open Access funding enabled and organized by Projekt DEAL.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-91421-x.

Correspondence and requests for materials should be addressed to J.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.