Molecular mechanism of elongation factor 1A inhibition by a Legionella pneumophila glycosyltransferase

Ramon Hurtado-Guerrero, Tal Zusman, Shalini Pathak, Adel F.M. Ibrahim, Sharon Shepherd, Alan Prescott, Gil Segal, Daan M.F. van Aalten

Biochemical Journal Immediate Publication. Published on 23 Dec 2009 as manuscript BJ20091351

Molecular mechanism of elongation factor 1A inhibition by a *Legionella pneumophila* glycosyltransferase

Ramon Hurtado-Guerrero[1], Tal Zusman[4], Shalini Pathak[1], Adel F. M. Ibrahim[2], Sharon Shepherd[1], Alan Prescott[3], Gil Segal[4] and Daan M. F. van Aalten[1*]

Division of Molecular Microbiology[1], CLS DNA Manipulation Team[2], and Division of Cell Biology and Immunology[3],

College of Life Sciences, University of Dundee, Dundee DD1 5EH, Scotland;

[4]Molecular Microbiology and Biotechnology, Life Sciences, Tel Aviv University, Tel Aviv, Israel

Running head: Mechanisms of a *Legionella pneumophila* glucosyltransferase

[*]Address correspondence to: Daan van Aalten (dmfvanaalten@dundee.ac.uk) or Ramon Hurtado-Guerrero (R.HurtadoGuerrero@dundee.ac.uk)

Keywords: *Legionella pneumophila*, glucosyl transferase, elongation factor 1A, microinjection, mutagenesis, protein structure

Abbreviations: *L. pneumophila* glucosyltransferase (*Lp*GT or Lgt1), Lgt2 (LegC8), Lgt3 (LegC5), toxin A and B from *Clostridium difficile* (*TcdA* and *TcdB*), *L. pneumophila* strain *Philadelphia* 1 ATCC 33152, strain *Lens* and strain *Paris* respectively (abbreviated here to *Lpp*GT, *Lpl*GT and *Lppa*GT respectively), elongation factor 1A (EF1A)

ABSTRACT

Legionnaires' disease is caused by a lethal colonization of alveolar macrophages with the Gram negative bacterium *Legionella pneumophila*. An *L. pneumophila* glucosyltransferase (*Lp*GT or Lgt1) has recently been identified as a virulence factor, shutting down protein synthesis in the human cell by specific glucosylation of elongation factor 1A (EF1A), using an unknown mode of substrate recognition, and retaining glycosyl transfer. We have determined the crystal structure of *Lp*GT in complex with substrates, revealing a GT-A fold with two unusual protruding domains. Through structure-guided mutagenesis of *Lp*GT, several residues essential for binding of the UDP-glucose donor and EF1A acceptor substrates were identified, also affecting *L. pneumophila* virulence as demonstrated by microinjection studies. Together, these data suggest that a positively charged EF1A loop binds in a negatively charged, conserved groove on the *Lp*GT structure, and that two asparagines are essential for catalysis. Furthermore, we show that two further *L. pneumophila* glycosyltransferases, that possess conserved UDP-glucose binding sites and EF1A binding grooves, are, like *Lp*GT, translocated into the macrophage through the Icm/Dot system.

INTRODUCTION

Over the past decade several studies have reported fascinating examples of pathogenic bacteria employing glycosyltransferases secreted into the human cell to control specific signal transduction pathways/cellular processes. For instance, the family GT 44 [1] glycosyltransferases toxin A and B from *Clostridium difficile* (*TcdA* and *TcdB*), cause pseudomembranous colitis and antibiotic-associated diarrhea by monoglucosylating Rho GTPases (such as Rho at Thr37, Rac at Thr35 and Cdc42) leading to the disruption of binding to guanosine nucleotide dissociation inhibitors (GDIs), inhibition of activation of small GTPases by guanine nucleotide exchange factors (GEFs), inhibition of the membrane-cytoplasm cycle and inhibition of the GTPase active form [2-5]. More recently, another bacterial glucosyltransferase, from *Legionella pneumophila* (*Lp*GT, also called Lgt1 or lpg1368), has been suggested to be involved in pathogenesis [6, 7]. *Legionella pneumophila* is an intracellular opportunistic Gram-negative pathogen able to proliferate within human alveolar macrophages [8], and is the causative agent of Legionnaires' disease [9]. This bacterium, once engulfed by macrophages, is able to replicate in vacuoles/phagosomes independently of the classical endolysosomal pathway [8, 10], until nutrient levels decline, leading to activation of the Icm/Dot type IV secretion system that releases several virulence factors [11]. At this stage flagellated bacteria are released and infect new host cells. *Lp*GT, annotated as a family GT88 glucosyltransferase in the CAZY database [1], was discovered to be a virulence factor, glucosylating Ser53 of human elongation factor 1A (hEF1A) by a retaining mechanism [12], leading to inhibition of ribosomal translation and consequently cell death [7]. However, the molecular mechanisms of glycosyl transfer and recognition of hEF1A are not understood.

Here, we report the crystal structure of *Lp*GT, containing a GT-A fold, revealing how the enzyme interacts with UDP-glucose. Through mutagenesis we identify residues important for catalysis and hEF1A recognition, and study the effect of these on virulence through microinjection studies, revealing that a positively charged hEF1A loop is likely to interact with a conserved binding groove on *Lp*GT. We also show that two recently described *Lp*GT apparent homologues (LegC5 or Lgt3/LegC8 or Lgt2) possess a conserved UDP-glucose binding site and hEF1A binding groove and are, like *Lp*GT, translocated through the Icm/Dot machinery, killing mammalian cells through induction of apoptosis.

MATERIAL AND METHODS

Cloning, mutagenesis and purification of Legionella pneumophila glucosyltransferases

Open reading frames lpg1368 and lpl1319 encoding the *Legionella pneumophila* glucosyltransferase 1 (UniProtKB/TrEMBL accessions Q5ZVS2 and Q5WWY0 in *L. pneumophila* strain *Philadelphia* 1 ATCC 33152 and strain *Lens*, respectively, abbreviated here to *Lpp*GT and *Lpl*GT, respectively) were amplified by PCR from genomic DNA extracted from the respective strains using the primers listed in Table I of the Supplementary material, and the PCR products were cloned directly into the bacterial expression vector pGEX6P1 (GE-Healthcare).

Site-directed mutagenesis was carried out following the 'QuikChange' Site-Directed Mutagenesis protocol (Stratagene), using the KOD HotStart DNA polymerase (Novagene). The resulting plasmid, pGEX6P1α*lpgt Philadelphia* strain also referred to here as wild type, was used as template for introducing the following single amino acid changes by site directed mutagenesis: D246A, D248A, D246A-D248A, N293A, E445A, E446A, Y454A, N499A, S519A. FLAG tags were also incorporated by site-directed mutagenesis. All plasmids were verified by sequencing.

The plasmids were transformed into BL21 (DE3) pLysS cells and grown at 37 ^{o}C until reaching an optical density of 0.6 at 600 nm, after which the expression of the protein was induced with 0.2 mM isopropyl β-D-1-thiogalactopyranoside at room temperature for an overnight incubation. The cells were harvested by centrifugation at 3480 g for 30 min and resuspended in a buffer A (25 mM Tris HCl, 250 mM NaCl, 4 mM DTT pH 8.5), containing lysozyme, DNAse and protease inhibitors. The cells were disrupted by a continuous flow cell disruptor (Constant Systems) at a pressure of 30 PSI and cere hthog

Enzymology

Mammalian cell lysates were prepared from HEK293 cells. Cells were harvested and resuspended in cold lysis buffer C (50 mM Tris HCl, 0.1 mM EGTA, 1 mM EDTA, 1% (w/w) Triton X-100, 1 mM Na_3VO_4, 50 mM NaF, 5 mM sodium pyrophosphate, 0.27 M sucrose, 0.1% (v/v) 2-mercaptoethanol, pH 7.5) with protease inhibitor cocktail (Roche). Protein concentrations of the lysates were determined using the Bradford method. Glucosylation reactions were carried out in 20 μl volumes consisting of buffer D (20 mM Tris HCl, 150 mM NaCl, 1 mM $MnCl_2$, pH 7.5), 7 μM recombinant glucosyltransferase, 50 μg of crude cell extract and 2.5 μM UDP-[^{3}H]glucose (American Radiolabeled Chemicals, St. Louis, MO). The mixture was incubated at 37°C for 3 h. The reaction was stopped by the addition of LDS sample buffer and heated at 100 °C for 3 minutes. The samples were then subjected to SDS/PAGE. Gels were stained with 0.1% Coomassie Brilliant Blue R-250 in 40% methanol and 10% acetic acid for 60 minutes and destained overnight with 10% acetic acid and 30% methanol. Incorporation of [^{3}H]glucose in proteins was visualized by treatment with EN^{3}HANCE (Perkin Elmer) for 30 min, followed by fluorography using X-Omat film (Kodak). Densitometry data were quantified using Aida analysis software.

Binding of UDP-glucose to wild type and mutant *Lpp*GT was analyzed by ligand-induced quenching of intrinsic tryptophan fluorescence. Fluorescence measurements were carried out with a Varian Cary Eclipse fluorescence spectrophotometer equipped with a thermostatted cuvette holder equilibrated at 25 °C. Emission spectra were recorded from 300–400 nm upon excitation at 295 nm. Excitation and emission slits were opened to 10 nm and 20 nm, respectively, and the spectra were recorded at a scan speed of 9–60 nm/min. Standard reaction mixtures contained 1 μM of *Lp*GT in 25 mM Tris HCl, 150 mM NaCl, 1 mM $MnCl_2$ (pH 7.5) to a final volume of 1 ml. After preincubation for 10 min at 25°C within the cuvette holder, aliquots of UDP-glucose were added to the mixture (total added volume not exceeding 2% of the total volume). The emission spectrum was recorded after each addition following mixing and 5 min incubation. All of the spectra were corrected for the background emission signal from both the buffer and the unbound UDP-glucose and repeated in triplicate. The equilibrium dissociation constant could be obtained from fitting the fluorescence intensity data to the standard single site binding equation with the software GraFit [13].

Microinjection

HeLa cells [14] were seeded onto 13 mm glass coverslips and allowed to settle overnight. The cells were then microinjected with 30 μM protein containing Texas Red conjugated dextran (to localise the injected cells) in injection buffer E (100 mM glutamic acid, 140 mM KOH, 1 mM $MgSO_4$, 1 mM DTT, pH 7.2 with citric acid [15]) as described previously [16]. The number of injection attempts were counted by the microinjector (Eppendorf) and recorded. Following 48 hours incubation the cells were fixed in 4% paraformaldehyde in PBS and counterstained with DAPI to distinguish apoptotic nuclei. The number of surviving cells (both those still adherent to the coverslip and rounded up cells) were then counted and expressed as a percentage of the number of injection events (Table II). Note that not all injection events were successful and that after 48 hours unperturbed cells would divide such that the final counts could exceed 100%.

Microscopy

HeLa cells were injected as above with FLAG-tagged proteins and following 24 hours incubation the cells were fixed in 4% paraformaldehyde in PBS. The cells were then permeabilised with 1% NP-40 in PBS and blocked with 0.5% fish-skin gelatin in PBS. The FLAG-tagged proteins were then localised with a mouse anti-FLAG antibody (M2-Sigma) and the ribosomes were localised

with a rabbit anti-L26 protein antibody (Sigma). These primary antibodies were in turn labelled with Alexa-488 conjugated goat anti-mouse IgG and Alexa-594 goat ant-rabbit IgG (Molecular Probes) respectively. The cells were imaged on either a Zeiss LSM510 confocal microscope using an alpha Plan-Fluar 100x objective (1.45NA) or a Leica sp2 confocal microscope using a HCX PL APO 63x objective (1.40NA).

Crystallization and structure determination

*Lpp*GT was spin-concentrated to 26 mg/ml and chemically methylated following the protocol described by Walter *et al.* [17]. Crystals were grown by sitting drop experiments at 20 °C, mixing 1 μl of protein, containing 10 mM UDP-glucose and 2 mM $MnCl_2$, with an equal volume of a reservoir solution (20% PEG 3000, 0.2 M NaCl, 0.1 M HEPES, pH 7.5). Under these conditions, crystals appeared within 7–14 days. They were cryoprotected with 20% glycerol, 20% PEG 3000, 0.2 M NaCl, 0.1 M HEPES, pH 7.5 and flash cooled prior to data collection at 100 K.

Following a similar protocol, *Lpl*GT was spin-concentrated to 18 mg/ml and co-crystallised with 10 mM UDP-glucose and 2 mM $MnCl_2$. Crystals were grown in 0.1 M $MgCl_2$, 12% PEG 6000, 0.1 M acetamido iminodiacetic acid, pH 6.5. 2-methyl-2,4-pentanediol (MPD) at 20% in mother liquor was used as a cryoprotectant.

A mercury chloride derivative of *Lpl*GT was generated by soaking experiments in mother liquor supplemented with 100 mM $HgCl_2$, for 3-20 minutes prior to data collection. Data for the native crystals and the heavy atom derivative were collected at beamlines ID23-1/BM14 (ESRF, Grenoble). All data were processed and scaled using the HKL suite [18] and CCP4 software [19], relevant statistics are given in Table I.

By SIRAS methodology, using data from a crystal soaked with $HgCl_2$ as well as a data set on a crystal of the *Lpl*GT cocrystallised with UDP-glucose (Table I), SHELXC/D/E (using the HKL2MAP GUI [20]) identified six sites, yielding phases with a figure of merit of 0.72 to 2.1 Å. An initial model for *Lpl*GT was built by ARP/warp [21] (building 370 residues of the single protein monomer in the asymmetric unit) and improved through cycles of manual model building in Coot [22] and refinement with REFMAC5 [23]. Molecular replacement with this structure was used to generate phases and a starting model for *Lpp*GT, which was refined using similar procedures. Topologies for UDP-glucose, UDP and glucose ligands were generated with PRODRG [24]. The final models were validated with PROCHECK [25] and WHATCHECK [26], model statistics are given in Table I. Coordinates and structure factors have been deposited in the Worldwide Protein Data Bank (wwPDB).

Bacterial strains

The *L. pneumophila* strains used in this study were *L. pneumophila* JR32, a streptomycin-resistant, restriction-negative mutant of *L. pneumophila Philadelphia*-1, which is a wild type strain in terms of intracellular growth [27], and GS3011, an *icm*T deletion mutant [28].

Construction of cyaA fusions

All genes examined were amplified by PCR using a pair of primers (Supplementary Table I) containing suitable restriction sites at the 5' end. The PCR products were subsequently digested with the relevant enzymes and cloned into the pMMB-cyaA-C or pMMB-cyaA-N vectors [29], to generate the plasmids listed in Supplementary Table II, all the inserts were sequenced to verify that no mutations were incorporated during the PCR.

Analyses performed with host cells

Intracellular growth assays in *Acanthamoeba castellanii* and in HL-60-derived macrophages were performed in a similar way to what was previously described [30]. The CyaA translocation assay was performed as described before [31].

RESULTS AND DISCUSSION

LpGT adopts the GT-A fold with a flexible donor binding loop

The *L. pneumophila Philadelphia* glucosyltransferase (*Lpp*GT) was cloned and expressed in *E. coli* as a GST fusion protein and purified by affinity and gel filtration chromatography. *Lpp*GT initially failed to produce diffraction quality crystals. Chemical methylation of solvent exposed lysines led to a protein sample that produced a single, well-diffracting crystal (Table I) for which the phase problem could not be solved. Our attention then shifted to the orthologous protein from *Legionella pneumophila Lens* (*Lpl*GT) that was cloned, expressed and purified using a similar strategy, readily producing well-diffracting crystals. The structure of *Lpl*GT bound to UDP-glucose was solved using a SIRAS experiment with a mercury derivative (Table I and Fig. 1). The phase problem of the *Lpp*GT diffraction data (later found to include ordered UDP and glucose in the active site) was then solved by molecular replacement (Fig. 1). Both structures were refined to high resolution, yielding final models with good R-factors (Table I).

The structures of *Lpp*GT/*Lpl*GT reveal three domains: a completely α-helical (α1-α7) N-terminal domain, a central domain containing the double Rossmann fold-like signature typical of the GT-A fold (α8- α15/ β1- β10), with a central β-sheet surrounded by α-helices on both sides. The third domain, termed the “protrusion domain” here, is an unusual α-helical protrusion from the GT-A fold, consisting mainly of α-helices (α16- α30) and two small β-strands (β11- β12). Well defined density for Mg^{2+}/UDP-glucose (*Lpl*GT) and Mn^{2+}/UDP/glucose (*Lpp*GT) was observed in the donor site, almost completely formed by the central subdomain and a long C-terminal loop coming from the third domain (Figs. 1A,4).

A comparison between the *Lpp*GT and *Lpl*GT structures reveals some conformational change within the N-terminal domains (RMS deviation of 0.4 Å for 82 aligned Cα atoms) and differences in order/disorder of a number of regions, such as a loop in the α-helical “protrusion domain” and the C-terminal loop (residues from 509 to 520, Fig. 1A,S1). Although these conformational changes could result from differences in crystal packing (Table I) or bound ligands (*Lpl*GT with Mg^{2+}/UDP-glucose and *Lpp*GT with Mn^{2+}/UDP/glucose), there have been many examples of metal/donor-induced conformational changes of loops/regions in glycosyltransferases that contribute to formation of the acceptor binding site [32]. In the *Lp*GT structures, in particular the flexible C-terminal loop is well positioned to create the acceptor-binding site upon binding to metal and UDP-glucose (Fig. 1A).

LpGT is structurally homologous to C. difficile toxin B and two other Legionella glucosyltransferases

Surprisingly, analysis of the *Lp*GT structures with the DALI server [33] reveals structural homology to *C. difficile* toxin B (*TcdB*, PDB ID 2bvl and 2bvm [34]), *C. novyi* alpha-toxin (PDB ID 2vk9 [35]), and *C. sordellii* lethal toxin (PDB ID 2vkd and 2vkh [35]). While structure-based sequence alignments show very low identities (14-18% with 185-199 aligned residues) (Fig. 2), the structural GT-A cores superpose well (RMSDs of 2.5-3.2 Å, Fig. 1A,1B,1C). For example, *TcdB* is formed by the typical two abutting Rossmann-like folds, which creates the sugar donor-binding site (Fig. 1), and superimposes well with some secondary structures, such as β3- β10, α8, α11- α13 and α15 from *Lp*GT (Fig. 1B,1C).

A recent study has identified two further putative *Legionella* glucosyltransferases, Lgt2 (LegC8) and Lgt3 (LegC5), 72 kDa and 100 kDa proteins, respectively, with additional multiple coil-coil domains [36]. Although sequence alignments show a low level of overall sequence identity (18-28%) (Fig. 2), the crystal structures reveal several regions of high conservation, not only including the UDP-glucose binding site/putative catalytic residues but also a putative acceptor binding groove (Fig. 1D). Compatible with this, a recent study has shown that these three *Legionella* enzymes all possess glycosyltransferase activity against hEF1A and kill eukaryotic cells [36].

Identification of the putative hEF1A docking site

Structural analysis of the known hEF1A glucosylation site (Ser53), using the *S. cerevisiae* E1FA structure (*Sc*EF1A, PDBID 2B7B [37]), reveals this to be located at the tip of a loop between two helices extending about 20 Å from the surface of the protein. Ser53 is flanked by two lysines, conserved between hEF1A and *Sc*EF1A, giving the tip of this loop on overall positive charge (Fig. 3). Interestingly, electrostatic analysis of the conserved putative acceptor binding site in *Lp*GT reveals an overall negative charge (Figs. 1D,3), suggesting electrostatic complementarity between *Lp*GT and the elongation factor glucosylation site. Using the constraints of the location of the UDP-Glc anomeric carbon, the Ser53 glucosylation site and the overall shape of the *Lp*GT/*Sc*EF1A proteins, it is possible to approximately position *Sc*EF1A in the *Lp*GT acceptor site, with qualitative shape-complementarity (Fig. 3). To test this model of interaction, we targeted a key exposed aromatic residue, Tyr454, lining the putative acceptor binding site (Fig.

3), by mutagenesis. Mutation of this residue to alanine led to a reduction in the activity towards hEF1A in HEK293 lysates while not significantly affecting UDP-glucose binding (Fig. 5A and Table II), suggesting approximate identification of an EF1A docking site on *Lp*GT.

LppGT/LplGT bind UDP-glucose through loops from the central and "protrusion" domain

The two crystal structures described here reflect two different states during catalysis (Fig. 4). The structure of *Lpl*GT was solved in complex with UDP-glucose and magnesium, resembling the substrate binding mode. The *Lpp*GT structure was solved in complex with UDP, glucose and manganese, resembling a product complex (Fig. 4). In both structures, the nucleotides occupy the same positions and adopt the same conformations, while a shift in position is observed for the glucose (1.6 Å max. atomic shift). The nucleotide is located between three loops: α12- α13, α4- α8 and the C-terminal loop (disordered in the *Lpl*GT structure, where only residues 518-523 are visible, Figs. 1A,4). The uracil ring is sandwiched between Trp139 (β4- α8 loop) and Pro225 (located at the beginning of α13) by hydrophobic stacking interactions (Fig. 4). With the exception of these two hydrophobic contacts, the majority of the interactions occur through hydrogen bonds with backbone residues from Ile138, Trp139, Phe140, Ile142, Ile247 and Ser519 (Fig. 4). Moreover, the phosphate group oxygens are recognised by hydrogen bonds with Ser519 and Trp520 side chains in the *Lpl*GT structure; and, as a result of the shift in position of the glucose, Asn499, Asn514 and Trp520 side chains and the Leu518 backbone in the *Lpp*GT structure (Fig.2). Although Ser519 hydrogen bonds the UDP α-phosphate, mutation of this residue affected neither UDP-glucose binding nor activity (Fig. 4 and Table II). A structural comparison with *TcdB* shows that two conserved residues, Trp102 and Trp520 in toxin B (Fig. 4), are involved in binding to uridine and phosphate groups, and are present in all GT44 *Clostridium* and GT88 *Legionella* enzymes, forming a common sequence signature of these two large family of enzymes. Mutagenesis has shown that Trp102 and Trp520 *TcdB* are important for UDP-glucose binding and catalysis [38, 39].

The DxD motif is essential for catalysis but not for donor binding

The active site metals of GT-A fold glycosyltransferases are known to have two roles: to induce a conformational change in a flexible loop and to stabilise a transition state during catalysis, with the help of two key aspartic acids in the "DxD motif" [32, 40, 41]. The phosphate group oxygens, Asp248 of the DxD motif and two ordered water molecules appear to pentagonally coordinate

the metal (Mg^{2+}) in the *Lpl*GT structure, whereas the *Lpp*GT reveals a hexagonally coordinated Mn^{2+} using an additional water molecule, but with Asp246 instead of Asp248 (Fig. 4). This is accompanied by a shift of 1.8 Å in the position of the metal. The importance of the DxD motif in *Lp*GT has been previously investigated using the D246N or D246N/D248N mutants, resulting in a decrease in glucosyltransfer activity against hEF1A [7]. We have further investigated this to dissect the effects on donor binding and activity. The D248A mutant showed no significant reduction of overall activity, but binding of UDP-glucose was an order of magnitude weaker as measured by tryptophan fluorescence (Fig. 5A and Table II), in agreement with the previously proposed donor binding role of the aspartic acids in the DxD motif [42, 43]. However, the D246A mutant showed a significant reduction in activity compared to the *Lpp*GT wild type enzyme, with only a moderate reduction in donor binding (Fig. 5A and Table II). Strikingly the D246A-D248A mutant had no detectable activity and reduced binding of UDP-glucose by an order of magnitude (Fig. 5A and Table II). Thus, the aspartic acids in the DxD may have roles in stabilization of the transition state and activity, as previously proposed by Qasba *et al.* [32], in addition to being important, but not essential, for binding of UDP-glucose [42-44].

In the substrate complex, glucose is hydrogen bonded to Asp230, Arg233 and Asp246 while in the product binding mode structure Asn293 substitutes Asp246. These changes in hydrogen bonding and the position of the DxD motif may reflect the shift in ligand position and the identity of the metal (Fig. 4). Interestingly, two of the glucose-interacting residues (Asp230/Arg233) are conserved in *TcdB* (Asp270/Arg273) (Fig. 4. Mutation of these residues in *TcdB* showed significant effects on binding to UDP-glucose and activity [39].

The LpGT active site confirms a retaining mechanism with two catalytic asparagines

TcdB belongs to a large family of GT-A glycosyltransferases with retaining character [45]. Among this family, the better known include *Neisseria meningitidis* galactosyltransferase (Lgtc) [42, 46-48], bovine α–1,3-galactosyltransferase (3GalT) [49], the two enzymes responsible for the formation of blood type A and B, α–1,3-N-acetylgalactosaminyltransferase (GTA) and α1,3-galactosyltransferase (GTB) [50-52], respectively, UDP GalNAc:polypeptide α-N-acetylgalactosaminyltransferases (ppGalNAcTs) [53-55] and mannosylglycerate synthase (MGS) [56]. For these retaining enzymes, a glutamine or glutamic acid residue has been proposed as the catalytic nucleophile involved in a $D_N{}^*A_{Nss}$ ion pair mechanism [45]. Comparison of the *Lp*GT, *C. difficile* toxin B and other retaining glycosyltransferase active sites

suggests that *Lp*GT may follow a retaining mechanism (Fig. 4), in agreement with recent work that has established a retaining mechanism for this enzyme by NMR spectroscopy [12]. The structures reveal not only conservation of the DxD motif and glucose binding residues, but also a key asparagine (Asn293), positioned ~4 Å away from the anomeric carbon, similarly to Gln189 in LgtC and Asn384 in *TcdB* that have been proposed to be involved in a back-side nucleophilic push [42, 45]. We mutated all *Lp*GT residues close to the anomeric carbon, including Asn293 and Asn499 and also Glu445 and Glu446, two residues positioned such that they could possibly act as a catalytic base if *Lp*GT would employ an inverting mechanism instead (Fig. 4. As expected, none of these residues affected binding of UDP-glucose (Table II), also showing that these mutant proteins were properly folded. Furthermore, mutation of the two glutamates did not affect activity of *Lp*GT towards hE1FA in cell lysates (Fig. 5A). Strikingly, however, mutation of either Asn293 or Asn499 resulted in mutant enzymes without any detectable glucosyltransfer activity (Fig. 5A). Thus, these mutations confirm that *Lp*GT follows a retaining catalytic mechanism as proposed for *C. difficile* toxin and other glycosyltransferases [42, 45-48]. Given the position of the two asparagines, Asn293 could act as the weak nucleophile involved in pushing the anomeric carbon during the Sn_i-like mechanism (through a $D_N{*}A_{Nss}$ ion pair) and Asn499 may play an essential role stabilising the transition state and the leaving group.

Inactive LpGT mutants are impaired in induction of HeLa cell apoptosis

Although it is known that the toxicity of *Lp*GT stems from its glucosylation of hEF1A, the precise mechanism of cell death is as yet unclear. Immunofluorescence microscopy analysis of HeLa cells microinjected with wild type *Lpp*GT suggests that the cells die with typical hallmarks of apoptosis, such as clumped DNA in the nuclei, plasma membrane blebbing and dead cells in phagosomes of adjoining cells (Fig. 5B and supplementary Fig. S2). More studies will be required to address how these enzymes, in addition to disruption of ribosome translation by their glucosyltransferase activity, induce apoptosis in mammalian cells.

Subsequently, we repeated the microinjection experiments using the *Lpp*GT mutants generated as part of this work. While all the mutants and wild type *Lpp*GT showed some detectable levels of cell death within 48 hours as had been reported previously [7], all the mutants, with exception of the D246A and D248A single mutants (from the DxD motif), showed a significant reduction in cell death compared to the wild type (Fig. 5B and Table II). The fact that the catalytically inactive mutants protect against cell death but not growth suggests that the toxicity of *Lpp*GT does not

entirely depend on its glucosyltransferase activity. It is possible that the inactive mutants are still able to disrupt protein-protein complexes or deplete hEF1A levels.

LppGT, Lgt2 and Lgt3 are exported through the Icm/Dot machinery

Although *Lpp*GT, Lgt2 and Lgt3 have been shown to kill mammalian cells through electroporation [7, 36] or microinjection studies (this study), it is as yet unclear how these enzymes are exported from *Legionella* into the macrophages. It is worth noting that a range of known *Legionella* virulence factors are substrates for the Icm/Dot system used by the bacterium to secrete virulence factors into the macrophage cytosol [10, 57, 58]. We tested whether *Lpp*GT, LegC5 and LegC8 were substrates for the Icm/Dot system by N-terminally fusing these enzymes to an inactive *Bordetella pertussis* adenylate cyclase toxin (CyaA), and stably expressing these in *L. pneumophila* JR32. These stable transfectants were used to infect HL-60-derived macrophages, and cAMP levels were determined to follow translocation [30]. In agreement with a previous study, Lgt3 was translocated into the macrophage [59, 60] and similar results were obtained for Lgt2 (Fig. 3C). Surprisingly, however, no significant cAMP levels were detected for *Lpp*GT. The majority of the *Legionella pneumophila* Icm/Dot substrates are known to contain a C-terminal signal sequence [61]. Interestingly, the only crystal structure of a type IV effector, RalF [62], shows a C-terminal disordered region with the last twenty residues as the signal sequence [62, 63], and secondary structure predictions suggest the Lgt3/Lgt2 C-termini contain a similar number of disordered amino acids. However, the *Lpp*GT structure shows that the C-terminus is well ordered, indeed, forms part of the active site (Fig.1A). Thus, we decided to repeat the secretion experiment with a C-terminal fusion of *Lpp*GT with CyaA. Strikingly, this led to translocation into macrophages, suggesting that, unusually, *Lp*GT contains its type IV signal sequence at the N-terminus (Fig. 1A and 5C). The N-terminus of *Lpp*GT is a-helical with only the first ten residues being completely disordered. It is not clear whether these residues are necessary and sufficient as a type IV secretion signal or whether a structural motif may be required for recognition by the Icm/Dot machinery. With the exception of *Lp*GT, only two other previous studies report cases of effector proteins carrying N-terminal type IV signal peptides [60, 64]. Taken together, these results are compatible with *Lpp*GT, Lgt2 and Lgt3 acting as virulence factors, ejected by *Legionella pneumophila* into the macrophage and contributing to cell death through inhibition of protein synthesis.

Considering the differences in sequence and length between *Lp*GT, Lgt2 and Lgt3, it is possible that these enzymes show differences in localisation in mammalian cells. To study this, we

expressed C-terminally FLAG-tagged versions of these enzymes in *E. coli* (retaining full activity, data not shown) and microinjected these into HeLa cells. Despite the significant differences in sequence and length (Fig.2), all three were diffusely distributed throughout the cell, showing some co-localisation with the ribosomal L26 protein, in agreement with hE1FA being one of the substrates of these enzymes (Fig. 5D).

CONCLUSIONS

Legionella is an accidental intracellular infectious bacterium. During the life cycle in the host macrophage, the bacterium needs to tightly control host cell processes to allow optimal replication without (initially) killing the host. At the end of the replication stage, apoptosis is induced and bacteria are released. Numerous effectors have been described, and some cases of redundancy have been reported [58, 65, 66]. Here we have studied a family of redundant glucosyltransferases, which kill mammalian host macrophages by apoptosis through post-translational modification of hEF1A and other, as yet unknown, mechanisms independently of their catalytic activity, as suggested by our microinjection studies with inactive mutants. Our data suggest redundancy in localisation and activity and reveals the *Lp*GT, Lgt2 and Lgt3 are all Icm/Dot effectors, consistent with the notion that they are virulence factors.

It is not clear why *L. pneumophila* would inject these three proteins with similar activities and localisation patterns into the host cell. Studies carried out by Belyi *et al.* show that *Lp*GT and Lgt2 are produced at the beginning of the stationary phase of its growth curve, and Lgt3 only at the early stage of the growth curve [36]. Thus, it is possible that *Legionella* only makes these proteins in the slowly replicating and non-dividing stages, either at the beginning or at the end of host infection. There are many other examples of *Legionella* secrecting redundant virulence factors, such as DrrA/SidM [58] with GEF and GDF activity, LidA [65] that binds to Rab1-GTP and LepB [66] with GTPase activating activity (GAP), all of them involved in the human Rab1 cycle. Similarly, *Lpp*GT, Lgt2 and Lgt3 may form a redundant mammalian killer toxin family, which may be relevant to start new infections and produce host death by apoptosis in order to infect new cells. In order to address how important these enzymes are in the *Legionella* infection cycle, studies with single, double and triple knockouts should be carried out.

Our structural studies show that *Lp*GT is a metal-dependent enzyme possessing a GT-A fold, and support the notion that Lgt2 and Lgt3 are also active glycosyltransferases, possessing conserved UDP-glucose and acceptor binding sites. Structural and mutation analyses suggest that a negatively charged binding groove in *Lp*GT may recognise a EF1A positive charged loop carrying the acceptor serine. Mutagenesis studies on several amino acids in the active site suggested *Lp*GT may employ a retaining mechanism involving two catalytic residues, Asn293, acting as the weak nucleophile, and Asn499, stabilising the transition state and UDP leaving group. It appears that the DxD motif is important in catalysis, yet not essential for donor binding.

In conclusion *Lp*GT, Lgt2 and Lgt3 are a redundant set of virulence factors forming a glucosyltransferase family with conservation of a putative EF1A binding site and a retaining mechanism. This work will form the basis for future studies towards the (protein) substrate specificity of these enzymes and the development of chemical biological probes for further cell biological studies of these virulence factors.

REFERENCES

1 Coutinho, P. M., Henrissat, B. (1999) Carbohydrate-active enzymes: an integrated database approach. In Recent Advances in Carbohydrate Bioengineering (Gilbert, H. J., Davies, G., Henrissat, B. and Svensson, B. , ed.), pp. 3-12, The Royal Society of Chemistry, Cambridge

2 Just, I., Selzer, J., Wilm, M., von Eichel-Streiber, C., Mann, M. and Aktories, K. (1995) Glucosylation of Rho proteins by Clostridium difficile toxin B. Nature **375**, 500-503

3 Schirmer, J. and Aktories, K. (2004) Large clostridial cytotoxins: cellular biology of Rho/Ras-glucosylating toxins. Biochim Biophys Acta **1673**, 66-74

4 Lyerly, D., Wilkins, TD. (1995) Clostridium difficile. Raven Press, New York

5 Jank, T., Giesemann, T. and Aktories, K. (2007) Rho-glucosylating Clostridium difficile toxins A and B: new insights into structure and function. Glycobiology **17**, 15R-22R

6 Belyi, I., Popoff, M. R. and Cianciotto, N. P. (2003) Purification and characterization of a UDP-glucosyltransferase produced by Legionella pneumophila. Infect Immun **71**, 181-186

7 Belyi, Y., Niggeweg, R., Opitz, B., Vogelsgesang, M., Hippenstiel, S., Wilm, M. and Aktories, K. (2006) Legionella pneumophila glucosyltransferase inhibits host elongation factor 1A. Proc Natl Acad Sci U S A **103**, 16953-16958

8 Albert-Weissenberger, C., Cazalet, C. and Buchrieser, C. (2007) Legionella pneumophila - a human pathogen that co-evolved with fresh water protozoa. Cell Mol Life Sci **64**, 432-448

9 Kumpers, P., Tiede, A., Kirschner, P., Girke, J., Ganser, A. and Peest, D. (2008) Legionnaires' disease in immunocompromised patients: a case report of Legionella longbeachae pneumonia and review of the literature. J Med Microbiol **57**, 384-387

10 Ninio, S. and Roy, C. R. (2007) Effector proteins translocated by Legionella pneumophila: strength in numbers. Trends Microbiol **15**, 372-380

11 Shin, S. and Roy, C. R. (2008) Host cell processes that influence the intracellular survival of Legionella pneumophila. Cell Microbiol **10**, 1209-1220

12 Belyi, Y., Stahl, M., Sovkova, I., Kaden, P., Luy, B. and Aktories, K. (2009) Region of elongation factor 1A1 involved in substrate recognition by Legionella pneumophila glucosyltransferase Lgt1: identification of Lgt1 as a retaining glucosyltransferase. J Biol Chem **284**, 20167-20174

13 Leatherbarrow, R. (2001) GraFit Version 5. Erithacus Software Ltd., Horley, UK

14 Neumann, B., Held, M., Liebel, U., Erfle, H., Rogers, P., Pepperkok, R. and Ellenberg, J. (2006) High-throughput RNAi screening by time-lapse imaging of live human cells. Nat Methods **3**, 385-390

15 Izant, J. G., Weatherbee, J. A. and McIntosh, J. R. (1983) A microtubule-associated protein antigen unique to mitotic spindle microtubules in PtK1 cells. J Cell Biol **96**, 424-434

16 Prescott, A. R., Dowrick, P. G. and Warn, R. M. (1992) Stable and slow-turning-over microtubules characterize the processes of motile epithelial cells treated with scatter factor. J Cell Sci **102 (Pt 1)**, 103-112

17 Walter, T. S., Meier, C., Assenberg, R., Au, K. F., Ren, J., Verma, A., Nettleship, J. E., Owens, R. J., Stuart, D. I. and Grimes, J. M. (2006) Lysine methylation as a routine rescue strategy for protein crystallization. Structure **14**, 1617-1622

18 Otwinowski, Z., Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods in Enzymology **276**, 307-326
19 COLLABORATIVE COMPUTATIONAL PROJECT, N. (1994) The CCP4 Suite: Programs for Protein Crystallography. Acta Cryst. **D50**, 760-763
20 Pape, D., Seil, R., Kohn, D. and Schneider, G. (2004) Imaging of early stages of osteonecrosis of the knee. Orthop Clin North Am **35**, 293-303, viii
21 Perrakis, A., Morris, R. and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement. Nat Struct Biol **6**, 458-463
22 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr **60**, 2126-2132
23 Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr **53**, 240-255
24 Schuttelkopf, A. W. and van Aalten, D. M. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr D Biol Crystallogr **60**, 1355-1363
25 Laskowski, R. A., Moss, D. S. and Thornton, J. M. (1993) Main-chain bond lengths and bond angles in protein structures. J Mol Biol **231**, 1049-1067
26 Hooft, R. W., Vriend, G., Sander, C. and Abola, E. E. (1996) Errors in protein structures. Nature **381**, 272
27 Sadosky, A. B., Wiater, L. A. and Shuman, H. A. (1993) Identification of Legionella pneumophila genes required for growth within and killing of human macrophages. Infect Immun **61**, 5361-5373
28 Zusman, T., Yerushalmi, G. and Segal, G. (2003) Functional similarities between the icm/dot pathogenesis systems of Coxiella burnetii and Legionella pneumophila. Infect Immun **71**, 3714-3723
29 Zusman, T., Aloni, G., Halperin, E., Kotzer, H., Degtyar, E., Feldman, M. and Segal, G. (2007) The response regulator PmrA is a major regulator of the icm/dot type IV secretion system in Legionella pneumophila and Coxiella burnetii. Mol Microbiol **63**, 1508-1523
30 Segal, G. and Shuman, H. A. (1999) Legionella pneumophila utilizes the same genes to multiply within Acanthamoeba castellanii and human macrophages. Infect Immun **67**, 2117-2124
31 Altman, E. and Segal, G. (2008) The response regulator CpxR directly regulates expression of several Legionella pneumophila icm/dot components as well as new translocated substrates. J Bacteriol **190**, 1985-1996
32 Qasba, P. K., Ramakrishnan, B. and Boeggeman, E. (2005) Substrate-induced conformational changes in glycosyltransferases. Trends Biochem Sci **30**, 53-62
33 Holm, L. and Sander, C. (1995) Dali: a network tool for protein structure comparison. Trends Biochem Sci **20**, 478-480
34 Reinert, D. J., Jank, T., Aktories, K. and Schulz, G. E. (2005) Structural basis for the function of Clostridium difficile toxin B. J Mol Biol **351**, 973-981
35 Ziegler, M. O., Jank, T., Aktories, K. and Schulz, G. E. (2008) Conformational changes and reaction of clostridial glycosylating toxins. J Mol Biol **377**, 1346-1356
36 Belyi, Y., Tabakova, I., Stahl, M. and Aktories, K. (2008) Lgt: a family of cytotoxic glucosyltransferases produced by Legionella pneumophila. J Bacteriol **190**, 3026-3035

37 Andersen, G. R., Pedersen, L., Valente, L., Chatterjee, I., Kinzy, T. G., Kjeldgaard, M. and Nyborg, J. (2000) Structural basis for nucleotide exchange and competition with tRNA in the yeast elongation factor complex eEF1A:eEF1Balpha. Mol Cell **6**, 1261-1266

38 Busch, C., Hofmann, F., Gerhard, R. and Aktories, K. (2000) Involvement of a conserved tryptophan residue in the UDP-glucose binding of large clostridial cytotoxin glycosyltransferases. J Biol Chem **275**, 13228-13234

39 Jank, T., Giesemann, T. and Aktories, K. (2007) Clostridium difficile glucosyltransferase toxin B-essential amino acids for substrate binding. J Biol Chem **282**, 35222-35231

40 Ramakrishnan, B. and Qasba, P. K. (2001) Crystal structure of lactose synthase reveals a large conformational change in its catalytic component, the beta1,4-galactosyltransferase-I. J Mol Biol **310**, 205-218

41 Charnock, S. J. and Davies, G. J. (1999) Structure of the nucleotide-diphospho-sugar transferase, SpsA from Bacillus subtilis, in native and nucleotide-complexed forms. Biochemistry **38**, 6380-6385

42 Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G. and Strynadka, N. C. (2001) Crystal structure of the retaining galactosyltransferase LgtC from Neisseria meningitidis in complex with donor and acceptor sugar analogs. Nat Struct Biol **8**, 166-175

43 Busch, C., Hofmann, F., Selzer, J., Munro, S., Jeckel, D. and Aktories, K. (1998) A common motif of eukaryotic glycosyltransferases is essential for the enzyme activity of large clostridial cytotoxins. J Biol Chem **273**, 19566-19572

44 Zhang, Y., Malinovskii, V. A., Fiedler, T. J. and Brew, K. (1999) Role of a conserved acidic cluster in bovine beta1,4 galactosyltransferase-1 probed by mutagenesis of a bacterially expressed recombinant enzyme. Glycobiology **9**, 815-822

45 Lairson, L. L., Henrissat, B., Davies, G. J. and Withers, S. G. (2008) Glycosyltransferases: Structures, Functions, and Mechanisms. Annu Rev Biochem **77**, 521-555

46 Lairson, L. L., Chiu, C. P., Ly, H. D., He, S., Wakarchuk, W. W., Strynadka, N. C. and Withers, S. G. (2004) Intermediate trapping on a mutant retaining alpha-galactosyltransferase identifies an unexpected aspartate residue. J Biol Chem **279**, 28339-28344

47 Zhang, Y., Swaminathan, G. J., Deshpande, A., Boix, E., Natesh, R., Xie, Z., Acharya, K. R. and Brew, K. (2003) Roles of individual enzyme-substrate interactions by alpha-1,3-galactosyltransferase in catalysis and specificity. Biochemistry **42**, 13512-13521

48 Jamaluddin, H., Tumbale, P., Withers, S. G., Acharya, K. R. and Brew, K. (2007) Conformational changes induced by binding UDP-2F-galactose to alpha-1,3 galactosyltransferase- implications for catalysis. J Mol Biol **369**, 1270-1281

49 Gastinel, L. N., Bignon, C., Misra, A. K., Hindsgaul, O., Shaper, J. H. and Joziasse, D. H. (2001) Bovine alpha1,3-galactosyltransferase catalytic domain structure and its relationship with ABO histo-blood group and glycosphingolipid glycosyltransferases. Embo J **20**, 638-649

50 Yamamoto, F., Clausen, H., White, T., Marken, J. and Hakomori, S. (1990) Molecular genetic basis of the histo-blood group ABO system. Nature **345**, 229-233

51 Patenaude, S. I., Seto, N. O., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M. and Evans, S. V. (2002) The structural basis for specificity in human ABO(H) blood group biosynthesis. Nat Struct Biol **9**, 685-690

52 Lee, H. J., Barry, C. H., Borisova, S. N., Seto, N. O., Zheng, R. B., Blancher, A., Evans, S. V. and Palcic, M. M. (2005) Structural basis for the inactivity of human blood group O2 glycosyltransferase. J Biol Chem **280**, 525-529

53 Fritz, T. A., Hurley, J. H., Trinh, L. B., Shiloach, J. and Tabak, L. A. (2004) The beginnings of mucin biosynthesis: the crystal structure of UDP-GalNAc:polypeptide alpha-N-acetylgalactosaminyltransferase-T1. Proc Natl Acad Sci U S A **101**, 15307-15312

54 Kubota, T., Shiba, T., Sugioka, S., Furukawa, S., Sawaki, H., Kato, R., Wakatsuki, S. and Narimatsu, H. (2006) Structural basis of carbohydrate transfer activity by human UDP-GalNAc: polypeptide alpha-N-acetylgalactosaminyltransferase (pp-GalNAc-T10). J Mol Biol **359**, 708-727

55 Fritz, T. A., Raman, J. and Tabak, L. A. (2006) Dynamic association between the catalytic and lectin domains of human UDP-GalNAc:polypeptide alpha-N-acetylgalactosaminyltransferase-2. J Biol Chem **281**, 8613-8619

56 Flint, J., Taylor, E., Yang, M., Bolam, D. N., Tailford, L. E., Martinez-Fleites, C., Dodson, E. J., Davis, B. G., Gilbert, H. J. and Davies, G. J. (2005) Structural dissection and high-throughput screening of mannosylglycerate synthase. Nat Struct Mol Biol **12**, 608-614

57 Shohdy, N., Efe, J. A., Emr, S. D. and Shuman, H. A. (2005) Pathogen effector protein screening in yeast identifies Legionella factors that interfere with membrane trafficking. Proc Natl Acad Sci U S A **102**, 4866-4871

58 Murata, T., Delprato, A., Ingmundson, A., Toomre, D. K., Lambright, D. G. and Roy, C. R. (2006) The Legionella pneumophila effector protein DrrA is a Rab1 guanine nucleotide-exchange factor. Nat Cell Biol **8**, 971-977

59 De Buck, E., Anne, J. and Lammertyn, E. (2007) The role of protein secretion systems in the virulence of the intracellular pathogen Legionella pneumophila. Microbiology **153**, 3948-3953

60 de Felipe, K. S., Pampou, S., Jovanovic, O. S., Pericone, C. D., Ye, S. F., Kalachikov, S. and Shuman, H. A. (2005) Evidence for acquisition of Legionella type IV secretion substrates via interdomain horizontal gene transfer. J Bacteriol **187**, 7716-7726

61 Cambronne, E. D. and Roy, C. R. (2007) The Legionella pneumophila IcmSW complex interacts with multiple Dot/Icm effectors to facilitate type IV translocation. PLoS Pathog **3**, e188

62 Nagai, H., Cambronne, E. D., Kagan, J. C., Amor, J. C., Kahn, R. A. and Roy, C. R. (2005) A C-terminal translocation signal required for Dot/Icm-dependent delivery of the Legionella RalF protein to host cells. Proc Natl Acad Sci U S A **102**, 826-831

63 Nagai, H., Kagan, J. C., Zhu, X., Kahn, R. A. and Roy, C. R. (2002) A bacterial guanine nucleotide exchange factor activates ARF on Legionella phagosomes. Science **295**, 679-682

64 Chen, J., Reyes, M., Clarke, M. and Shuman, H. A. (2007) Host cell-dependent secretion and translocation of the LepA and LepB effectors of Legionella pneumophila. Cell Microbiol **9**, 1660-1671

65 Machner, M. P. and Isberg, R. R. (2006) Targeting of host Rab GTPase function by the intravacuolar pathogen Legionella pneumophila. Dev Cell **11**, 47-56

66 Ingmundson, A., Delprato, A., Lambright, D. G. and Roy, C. R. (2007) Legionella pneumophila proteins that regulate Rab1 membrane cycling. Nature **450**, 365-369

ACKNOWLEDGEMENTS

We thank the European Synchrotron Radiation Facility, Grenoble, for beam time. This work was supported by a Wellcome Trust Senior Research Fellowship and an MRC Programme Grant to DvA. Coordinates and structure factors of all structures have been deposited with the PDB (entries 2WZF and 2WZG).

FIGURE LEGENDS

Fig. 1. **Structure of *Lp*GT.**

(**A and B**) Ribbon diagram of *Lpp*GT crystal structure **(A)** in complex with UDP, glucose and manganese, and comparison with *TcdB* (PDB ID 2bvl and 2bvm [34], **(B)**). Secondary structures are represented in red and green colour for α-helixes of *Lpp*GT and toxin B N-terminal domain, brown and blue colour for α-helixes and β-strands in central domain of *Lpp*GT and toxin B, and pink and olive colour for α-helixes and β-strands of the C-terminal "protrusion domain" of *Lpp*GT and toxin B, respectively. UDP and glucose are shown in green sticks and manganese as a pink sphere. The two aspartic acids (D246 and D248 in *Lpp*GT) are shown as cyan sticks. Arrows indicate flexible regions in both crystal structures.

(C) Superposition of *Lpp*GT (grey) and *Tcd*B (brown). UDP and glucose are shown in green and blue sticks in *Lpp*GT and *Tcd*B, respectively.

(D) Surface representation of the LppGT enzyme, coloured by sequence conservation (from red (100% identity)) to grey (<50% identity)) with Lgt2 and Lgt3.

Fig. 2. Multiple sequence alignment of the GT88 family members (*LppGT*, *Lppa**GT***, *Lpl*GT, *Lgt2 and Lgt3*) and *Tcd*B. Secondary structure elements from the *LppGT* structure are dark-grey for the N-terminal domain, black for the central domain and light-grey for the "potrusion domain". Conserved catalytic glutamine residues are indicated with a black circle, the aspartic residues of the DxD motif are marked with a black star, amino acids interacting with UDP-glucose by direct hydrogen bonds or hydrophobic stacking interactions are pointed out by a black triangle, and amino acids interacting with UDP-glucose by indirect hydrogen bonds through water molecules are highlighted by a grey triangle.

Fig. 3. (Right panel) Electrostatic surface representation of *Lpp*GT and *Sc*EF1A (PDBID 2B7B [37]), showing a negative binding groove, which may interact with the positively charged loop on *Sc*EF1A that carries the acceptor serine (Ser53). Tyr454 in the putative binding groove on *Lpp*GT and the acceptor serine Ser53 on *Sc*EF1A are indicated by arrows. (Left panel) Zoom in of the putative interaction site between *Lpp*GT and *Sc*EF1A. The surface of *Lpp*GT and *Sc*EF1A are represented as brown and green respectively. Tyr454 and Ser53 are shown as sticks in pink

colour while charged residues are in yellow (including some other residues forming the loop in which Ser53 is localized, such as Gly52 and Phe54).

Fig. 4. **Active site of *Lp*GT bound to UDP-glucose and metals.**
Stereo figures of the *Lpl*GT, *Lpp*GT and *TcdB* active sites. *Lpl*GT is shown in complex with UDP-glucose and magnesium. *Lpp*GT and *TcdB* crystal structures are shown in complex with UDP, glucose and manganese. The amino acids are shown in grey sticks. The ligands and metals are shown in green sticks and pink spheres, respectively. Protein-ligand hydrogen bonds are shown as dotted green lines. The distances between Asn293 and Asn499 to the anomeric carbon are shown as thin lines. Unbiased (i.e. before inclusion of any inhibitor model) $|F_o| - |F_c|$, ϕ_{calc} electron density maps are shown at 2.5 σ.

Fig.5. **Site directed mutagenesis, microinjection, translocation and localisation studies.**

(A) Upper panel, autoradiography of wild type and mutant enzymes incubated with HEK293 lysates and UDP-[^{3}H]glucose. Lower panel, quenching of intrinsic *Lpp*GT tryptophan fluorescence measured at increasing concentrations of UDP-glucose. All data points represent the average of three measurements ± standard deviation. The K_d for UDP-glucose was determined by fitting fluorescence intensity data against free UDP-glucose concentration (insert graph) - see Table II for fitted K_d for wild type and mutants.

(B) HeLa cells microinjected with wild type, D246A/D248A, N293A mutants and GST as control at a protein concentration of 30 μM in the injection needle. Cells were co-injected with Texas Red-conjugated dextran and incubated for 48 h and counterstained with DAPI. Representative fields are shown for the cells after 48 hours incubation. Wild type protein showed few surviving red (Texas Red dextran) cells and of these most were rounded up cells compared to GST protein and double mutant. N293A was slightly protective compared to the wild type enzyme.

(C) Translocation experiments. All three enzymes were translocated through the Icm/Dot machinery (brown bars) compared to a mutated Δ*icmT Legionella* strain (white bars) as measured by cAMP concentrations.

(D) Microinjected FLAG-tagged *Lpp*GT protein was injected into HeLa cells and incubated for 24 h. The FLAG-tagged protein (green) was diffusely distributed throughout the cell and showed

rare co-localisation with the counterstain ribosomal L26 protein (red). A detail of the centre of the top panels is shown in the bottom panels. Scale bars 10 μm for top panels, 5 μm for bottom panels.

Table I

Data collection and refinement statistics. Values in parentheses refer to the highest resolution shell. Ramachandran plot statistics were determined with PROCHECK [25].

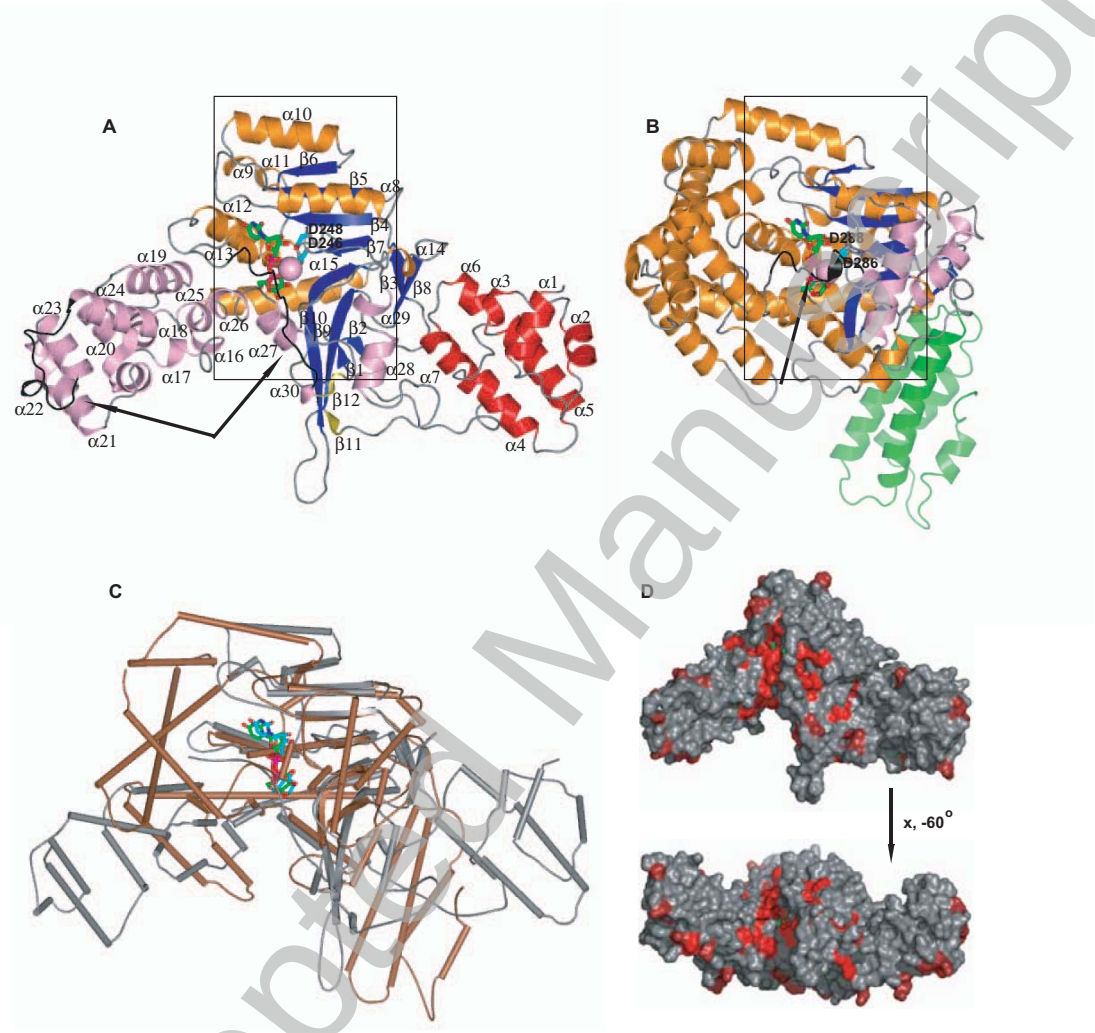
	*Lpl*GT $HgCl_2$ derivative	*Lpl*GT + UDP-glucose	*Lpp*GT + UDP + glucose
Space group	H3	H3	$P2_1$
Wavelength (Å)	1.003	1.282	1.892
Resolution (Å)	20.0–2.10	20.0–1.9	20.0–2.1
Cell dimensions	*a* = 122.3 Å, *b* = 122.3 Å, *c* = 103.1 Å	*a* = 122.8 Å, *b* = 122.8 Å, *c* = 103.4 Å	*a* = 51.0 Å, *b* = 104.6 Å, *c* = 53.3 Å
Unique reflections	33622	45918	31842
Completeness	0.999 (0.991)	0.992 (0.969)	0.971 (0.918)
R_{sym}	0.099 (0.520)	0.073 (0.50)	0.125 (0.402)
$I/\sigma(I)$	26.0 (5.7)	29 (5.8)	23.2 (5.5)
Redundancy	12.8 (11.8)	10.3 (7.2)	6.5 (3.6)
R_{work} / R_{free}		0.199 / 0.261	0.189 / 0.254
RMSD from ideal geometry, bonds (Å)		0.012	0.008
RMSD from ideal geometry, angles (°)		1.4	1.2
B-factor RMSD ($Å^2$) (backbone bonds)		0.74	0.56
<*B*> protein ($Å^2$)		35.4	27.3
<*B*> UDP-glucose ($Å^2$)		29.4	
<*B*> UDP ($Å^2$)			19.4
<*B*> glucose ($Å^2$)			26.9
<*B*> solvent ($Å^2$)		40.6	33.5
Ramachandran plot: Most favoured (%) Additionally allowed (%) Generously allowed (%)		97.7 2.3 0.0	96.2 2.9 0.21
PDB ID		2WZG	2WZF

TABLE II

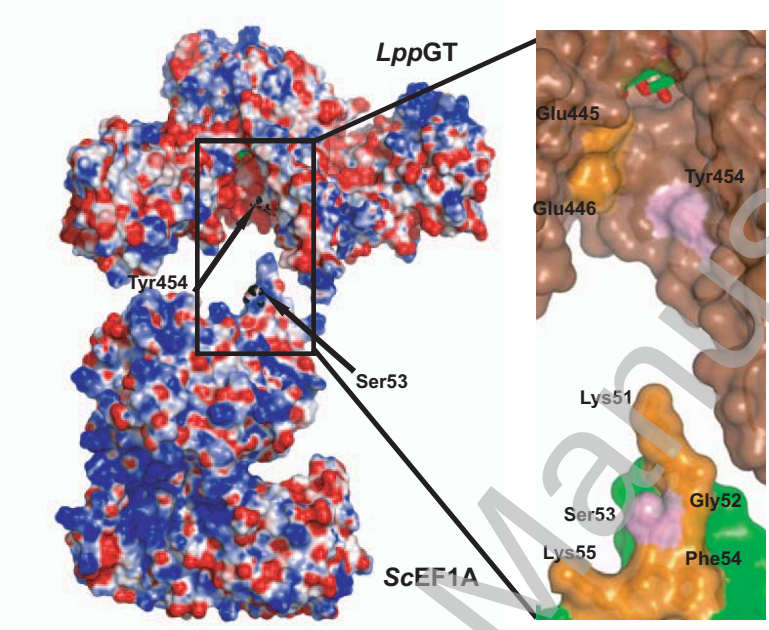
Activity and UDP-glucose binding of single and double mutants compared to the wild type enzyme, *Lpp*GT. Percentage of surviving and intact HeLa cells counted after 48 h. Intact cells were not apoptotic and rounded. The K_d for UDP-glucose was determined by fitting fluorescence intensity data, obtained by tryptophan fluorescence experiments, against free UDP-glucose concentration (see Fig.3A). The activity and K_d experiments represent three independent experiments while microinjection studies represent two independent experiments. Microinjected cells were counted after 48 h.

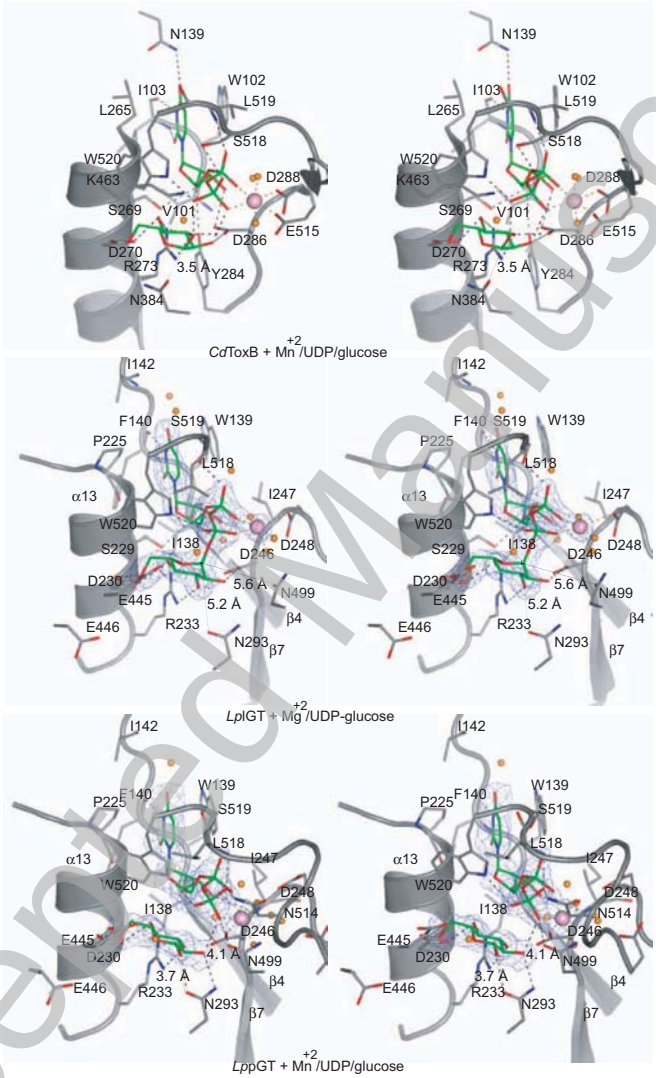
	% activity	K_d (nM)	% surviving cells	% intact cells
Wild-type (*Lpp*GT)	100	100 ± 29	33	10
D246A	48 ± 3	400 ± 80	38	18
D248A	92 ± 3	990 ± 90	36	18
D246A-D248A	nd	830 ± 100	100	95
N293A	nd	72 ± 30	93	90
E445A	86 ± 7	41 ± 25	-	-
E446A	78 ± 10	130 ± 40	-	-
Y454A	43 ± 2	60 ± 30	-	-
N499A	nd	90 ± 20	110	110
S519A	87 ± 4	100 ± 43	-	-
GST	-	-	305	300
Buffer	-	-	247	245

Biochemical Journal Immediate Publication. Published on 23 Dec 2009 as manuscript BJ20091351



Biochemical Journal Immediate Publication. Published on 23 Dec 2009 as manuscript BJ20091351





Biochemical Journal Immediate Publication. Published on 23 Dec 2009 as manuscript BJ20091351