

# Structural Basis of the Action of Glucosyltransferase Lgt1 from *Legionella pneumophila*

Wei Lü<sup>1</sup>, Juan Du<sup>1</sup>, Michael Stahl<sup>2</sup>, Tina Tzivelekidis<sup>2</sup>, Yury Belyi<sup>3</sup>, Stefan Gerhardt<sup>1</sup>, Klaus Aktories<sup>2</sup> and Oliver Einsle<sup>1\*</sup>

<sup>1</sup>Institut für organische Chemie und Biochemie, Albert-Ludwigs-Universität Freiburg, D-79104 Freiburg, Germany

<sup>2</sup>Institut für Experimentelle und Klinische Pharmakologie und Toxikologie, Albert-Ludwigs-Universität Freiburg, D-79104 Freiburg, Germany

<sup>3</sup>Gamaleya Research Institute, Moscow 123098, Russia

Received 11 September 2009;  
received in revised form  
16 November 2009;  
accepted 17 November 2009  
Available online  
23 November 2009

Edited by M. Guss

The glucosyltransferase Lgt1 is one of three glucosylating toxins of *Legionella pneumophila*, the causative agent of Legionnaires disease. It acts through specific glucosylation of a serine residue (S53) in the eukaryotic elongation factor 1A and belongs to type A glycosyltransferases. High-resolution crystal structures of Lgt1 show an elongated shape of the protein, with the binding site for uridine diphosphate glucose at the bottom of a deep cleft. Lgt1 shows only a low sequence identity with other type A glycosyltransferases, and structural conservation is limited to a central folding core that is usually observed within this family of proteins. Domains and protrusions added to the core motif represent determinants for the specific recognition and binding of the target. Manual docking experiments based on the crystal structures of toxin and target protein suggest an obvious mode of binding to the target that allows for efficient transfer of a glucose moiety.

© 2009 Elsevier Ltd. All rights reserved.

**Keywords:** glucosyltransferases; Legionnaires disease; *Legionella pneumophila*; protein crystallography

## Introduction

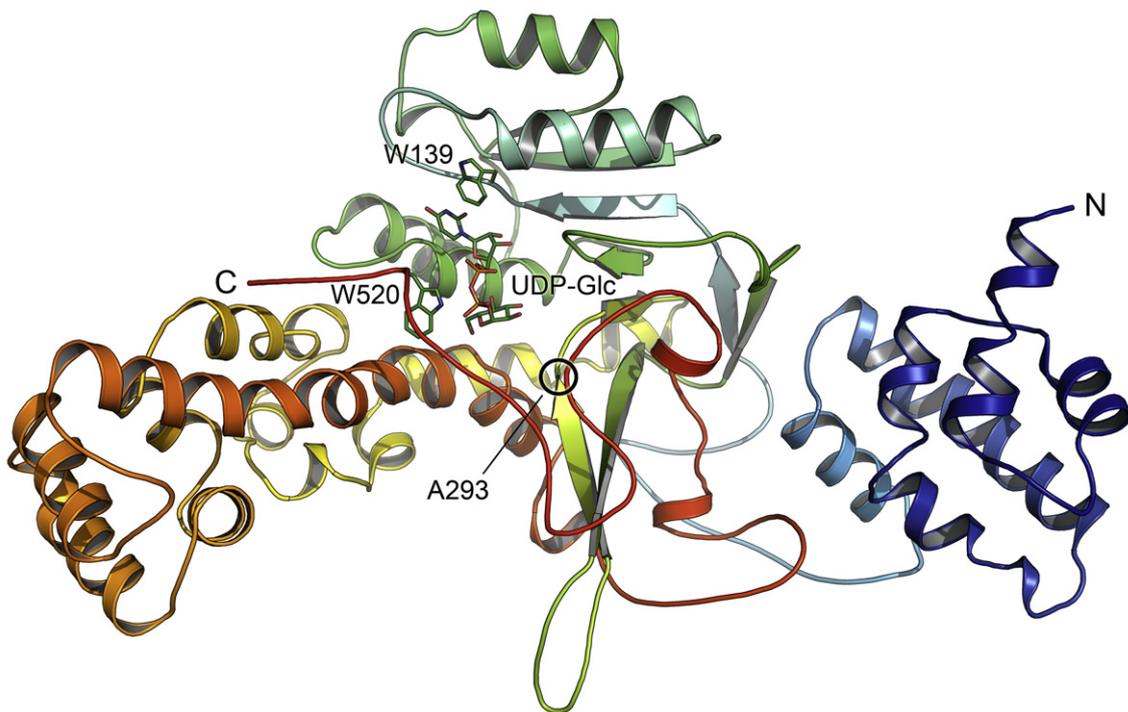
*Legionella pneumophila* is the causative agent of Legionnaires disease, which is characterized by severe pneumonia with high fatality rates.<sup>1,2</sup> The pathogens are intracellular parasites infecting protozoa (e.g., amoebae), as well as animals and humans. In humans, *L. pneumophila* is able to invade and proliferate in phagocytes. The bacteria severely affect the normal course of phagocytosis, inhibit acidification of phagosomes, and prevent phagosome-lysosome fusion.<sup>3–5</sup> Moreover, the pathogens alter phagocytic membrane biogenesis and trigger the formation of “replicative vacuoles,” which are characterized by ribosome-studded membranes.<sup>6,7</sup> After

replication of *L. pneumophila*, host cells die due to induction of apoptosis or necrosis,<sup>8–10</sup> and the pathogens are released for a new cycle of cell invasion and replication.<sup>10,11</sup>

Recent studies showed that the complex pathogen–host interaction of *L. pneumophila* depends on the release of a large array of bacterial factors into the cytosol of host cells.<sup>12</sup> Recently, several related glucosyltransferases from *L. pneumophila* strain Philadelphia-1, which severely affect the protein synthesis of host cells *in vitro* and *in vivo*, were identified.<sup>13–15</sup> The bacterial effectors termed Lgt1–Lgt3 mono-*O*-glucosylate eukaryotic elongation factor 1A (eEF1A) at residue S53. The modification of eEF1A blocks protein biosynthesis and causes the death of target cells. The prototype of the *L. pneumophila* glucosyltransferase Lgt1 is an ~60-kDa protein that exhibits a low sequence similarity with clostridial glucosylating toxins (21.5%/31.5% sequence identity/similarity with *Clostridium difficile* toxin B and 18.4%/31.0% sequence identity/similarity with *Clostridium sordellii* cytotoxin L), which modify small GTPases of the Rho/Ras GTPase family. Lgt2 and Lgt3 are about 24%

\*Corresponding author. E-mail address: einsle@biochemie.uni-freiburg.de.

Abbreviations used: eEF1A, eukaryotic elongation factor 1A; GT-A, type A glycosyltransferases; SeMet, selenomethionine; UDP-Glc, uridine diphosphate glucose; PDB, Protein Data Bank.



**Fig. 1.** Three-dimensional structure of the glucosyltransferase Lgt1 from *L. pneumophila* shown in cartoon representation. The protein chain is shown from blue at the N-terminus to red at the C-terminus. The typical GT-A fold of glycosyltransferases forms the central part of the protein, while the N-terminal part constitutes a distinct domain. The UDP-Glc ligand and the side chains of residues W139 and W520 are shown in stick representation (see Fig. 2).

and 18% identical with Lgt1, but share the same eukaryotic substrate (eEF1A) and target site (S53). Recently, the minimal structural requirements for the recognition of eEF1A by Lgts have been determined. It turned out that the decapeptide 50-GKGSFKYAWV-59 of eEF1A is sufficient for modification by *L. pneumophila* glucosyltransferases. Moreover, with this small substrate peptide, it was possible to identify Lgt1 by NMR analysis as a retaining glucosyltransferase.<sup>16</sup> Sequence comparison of the genomes of various *L. pneumophila* strains revealed that variants of the three Lgt subfamilies of glucosyltransferases are present in different strains of *L. pneumophila*<sup>15</sup> (e.g., *L. pneumophila* strains Philadelphia-1, Paris, Lens, and Corby). *Legionella* glucosyltransferases are listed as members of glycosyltransferase family 88 in the carbohydrate-active enzymes database CAZy†.

Lgts were initially recognized as glycosyltransferases through their slight sequence similarity with regions of the active site of clostridial glycosylating toxins, including *C. difficile* toxins A and B and *C. sordellii* lethal toxin.<sup>13,17</sup> Recent crystal structure analyses of the catalytic domain of the clostridial toxins showed that they are members of the type A glycosyltransferases (GT-A) family.<sup>18,19</sup> They have a typical Rossmann-like folding core, arranged in two tightly associated adjoining  $\beta$ - $\alpha$ - $\beta$  domains that form a central  $\beta$ -sheet that is involved in the binding

of the nucleotide sugar. Members of the GT-A family frequently contain an amino acid signature motif DXD in which two aspartic acid residues participate in metal ion binding and catalysis.<sup>20,21</sup> This motif is found in both *Legionella* glucosyltransferases and clostridial glycosylating toxins. The second family of glycosyltransferases is the GT-B family, which also consists of two Rossmann-like folding units. However, in the GT-B family, the DXD motif is absent, and the two  $\beta$ - $\alpha$ - $\beta$  domains are not tightly associated, leaving a cleft in between.<sup>20</sup>

Here we report the 1.7-Å crystal structure of Lgt1 from *L. pneumophila* strain Lens, which is 89% identical in sequence with Lgt1 from *L. pneumophila* strain Philadelphia-1 (Fig. S1). In spite of a conserved glycosyltransferase core fold, the overall structure of Lgt1 differs markedly from that of the clostridial glycosylating toxins, and the shape of the molecule itself provides a clear indication as to how interaction with the target eEF1A might take place.

## Results

### Structure of Lgt1

Work on the glucosyltransferase Lgt1 started originally with the enzymes from *L. pneumophila* strains Philadelphia-1 and Lens. As we intended to study the Lgt1 protein-substrate complex, we decided to start with inactivated enzymes. Earlier

† <http://www.cazy.org>

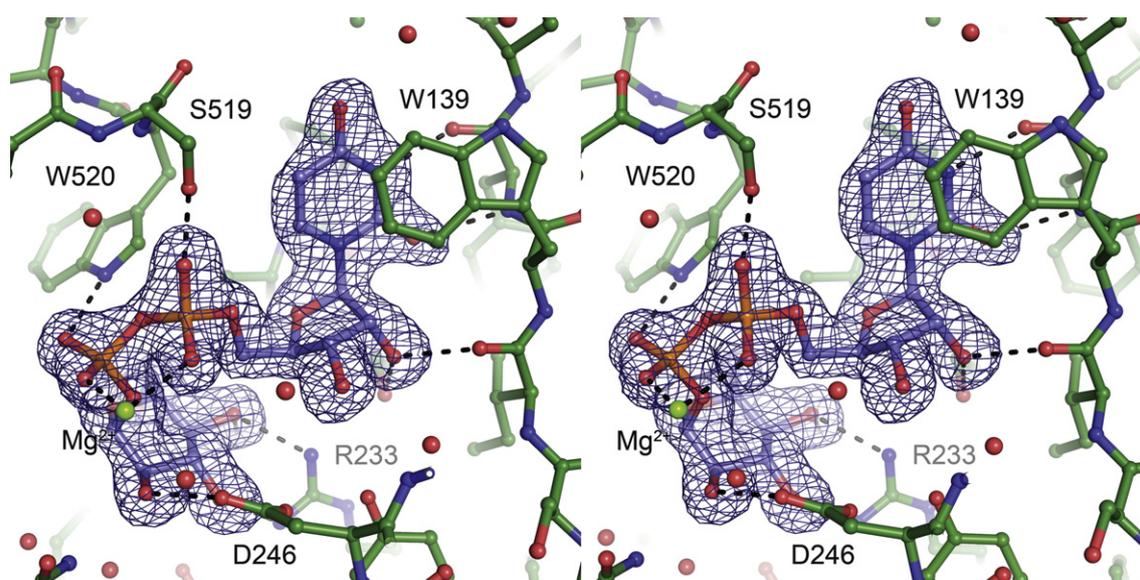
mutational studies of the related *C. difficile* toxin B revealed that residue N384 is important for enzymatic activity.<sup>20</sup> This asparagine is conserved in all clostridial enzymes and also in Lgt1 from *L. pneumophila* strains Philadelphia-1 and Lens (Fig. S2). Exchange of the equivalent N293 with alanine in Lgt1 caused inhibition of the glucosylation of eEF1A (Fig. S3). As the yield of expression of the Lgt1 N293A variant in *L. pneumophila* strain Lens was very high and was expected not to glucosylate its substrate, it was subsequently used for crystallization trials.

The structure of *L. pneumoniae* Lgt1 was solved by SeMAD, and models were obtained for the selenomethionine (SeMet)-labeled protein in complex with the substrate uridine diphosphate glucose (UDP-Glc) at a resolution of 1.7 Å and for the unlabeled protein in complex with the nucleotide UDP at a resolution of 2.3 Å. Both forms were crystallized under identical conditions and show only minor deviations, as discussed below. The structure of Lgt1 groups into three distinct domains that yield an elongated overall shape of the molecule (Fig. 1). The helical N-terminal domain encompasses amino acids 5–104. A linker region connects this domain to the central part of the protein (residues 128–322), where a mixed  $\alpha/\beta$  fold of complex topology forms the actual glucosyltransferase domain that also contains the binding site for UDP-Glc. The third domain of Lgt1 is less well defined and forms an extended protrusion between residues 323 and 444, after which the course of the peptide chain returns to the UDP-Glc binding site only to subsequently terminate in an elongated C-terminal loop. The final part of the peptide chain, beyond residue 513, was only found to be structured in the UDP-Glc

complex, where it covers the binding cleft for the nucleotide-sugar, possibly locking it in place (Fig. 1).

UDP-Glc is bound by Lgt1 at the bottom of a deep binding cleft (Figs. 1 and 2). The nucleotide base stacks against the indole moiety of W139, and the protein assures base specificity through two hydrogen bonds from uridine-N3 to the backbone carbonyl of F140 and from the backbone amide nitrogen of the same residue to uridine-O2. The ribose moiety forms only a single hydrogen bond between its 2'-OH group and the side chain of S229, while the 3'-OH group only contacts two water molecules. The  $\alpha$ -phosphate of UDP-Glc is hydrogen-bonded to S519, and the  $\beta$ -phosphate of UDP-Glc is hydrogen-bonded to the indole nitrogen of W520. Both coordinating residues are located in the very C-terminal part of the protein, and the correct binding of UDP-Glc may be a necessary prerequisite for an ordered conformation of this region. The glucopyranose moiety of UDP-Glc is firmly locked in the binding pocket, with the 2''-OH, 3''-OH, 4''-OH, and 6''-OH groups of the sugar in direct hydrogen-bonding contact with the side chains of a characteristic triad consisting of D230, R233, and D246. Lgt1 is thus able to probe the conformations of all hydroxy groups on the sugar and confirm its identity. Galactose, as an example of another sugar that is activated by uridylylation, is the C4'' epimer of glucose. To distinguish glucose from galactose, the enzyme ascertains the equatorial positioning of the 4''-OH group of glucose by short hydrogen bonds to both D230 (2.6 Å) and R233 (2.8 Å).

The C1'' atom of glucose faces the solvent and will be the site of an attack by S53 of the target eEF1A. However, Lgt1 has most recently been shown to retain the  $\alpha$ -conformation of the anomeric C1'' atom,



**Fig. 2.** Binding of UDP-Glc in *L. pneumophila* Lgt1. The stereo representation shows an experimental  $F_o - F_c$  omit electron density map contoured at  $3.0\sigma$  around the UDP-Glc molecule. Hydrogen bonds to the protein are depicted as dotted black lines. The relevant residues for UDP-Glc binding are highlighted, and  $Mg^{2+}$  coordinates the two phosphate moieties.

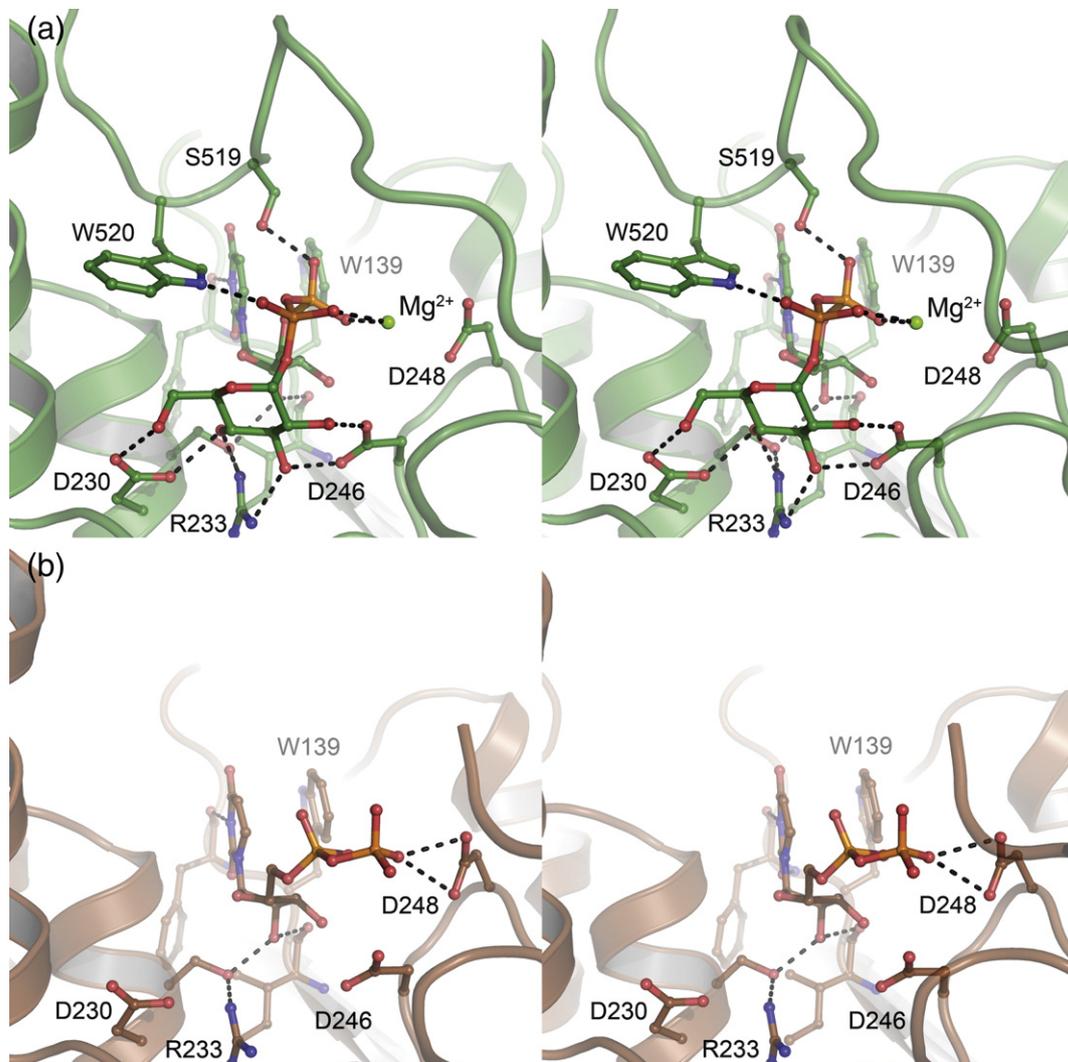
as seen in UDP-Glc also in the glucosylated target.<sup>16</sup> The same has been found for clostridial glycosylating toxins (e.g., *C. sordellii* lethal toxin),<sup>22</sup> and it has been suggested that an internal nucleophilic substitution ( $S_Ni$ ) proceeding through a short-lived oxocarbenium intermediate allows for the retention of the  $\alpha$ -anomeric conformation.<sup>18,23</sup>

In the UDP-bound form of Lgt1, only minimal structural changes are observed in the protein chain. The triad of residues that coordinates the glucosyl moiety in the UDP-Glc-bound form of Lgt1—D230, R233, and D246—retains its exact position in the absence of the carbohydrate, indicating that the side chains are fixed within the protein in order to fulfill their role and to ascertain the exclusive binding of a glucosyl nucleotide (Fig. 3). The stabilization of the glucose moiety by Lgt1 is further underlined by the observed displacement of the two phosphates in the

UDP-bound form: The  $\beta$ -phosphate moves significantly within the binding cleft and forms two weak hydrogen bonds of 2.9 Å and 3.0 Å to the  $\beta$ -carboxy group of D248. Accordingly, the phosphates are also not in a position to stabilize the C-terminus of Lgt1 by hydrogen-bonding with S519 and W520, leaving the entire region of the protein beyond residue L513 in a disordered state (Fig. 3b).

### Structural relation to other glycosyltransferases

A DALI<sup>24</sup> search using *L. pneumophila* Lgt1 as search model yields the highest homologies with GT-A toxin B from *C. difficile* [Protein Data Bank (PDB) ID 2BVL]<sup>19</sup> and cytotoxin L ('lethal toxin') from *C. sordellii* (PDB ID 2VKH).<sup>18</sup> The Z-scores from DALI were 16.0 for cytotoxin L and 15.3 for toxin B. The structural similarities of these two



**Fig. 3.** Ligand binding in Lgt1. (a) Binding of UDP-Glc. In particular, the equatorial hydroxy groups of glucose are specifically coordinated by a triad consisting of D230, R233, and D246. Residues S519 and W520 interact with the phosphate groups of the ligand, and this stabilized the conformation of the C-terminus of Lgt1. (b) In the absence of glucosyl moiety in the UDP complex, the phosphate groups are not held in place and relocate to form weak hydrogen bonds to D248. The C-terminus of the protein is disordered. The stereo figures are shown with identical orientations of Lgt1.

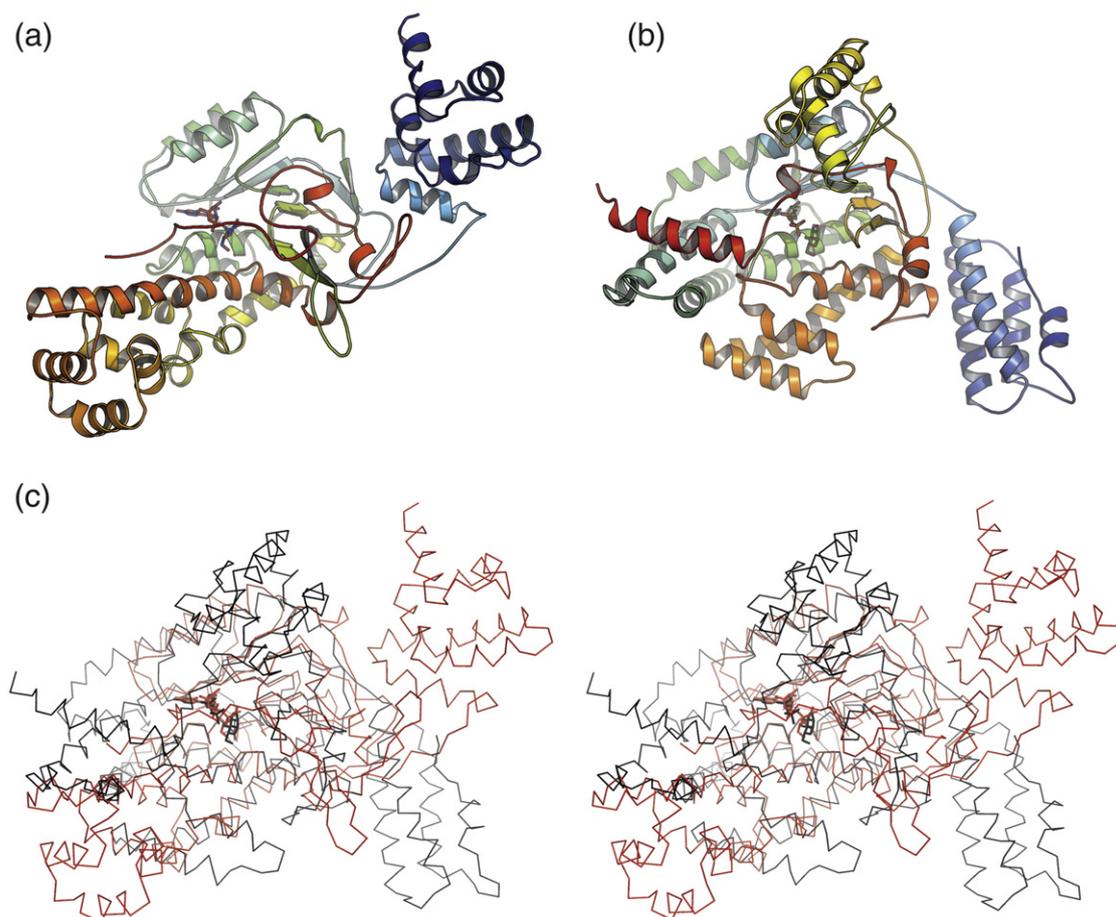
proteins have been discussed in detail previously,<sup>18</sup> and their folds are virtually identical, with a root-mean-squared deviation (rmsd) of 1.04 Å for 483 C $\alpha$  atoms. However, upon first inspection, the overall shape of *Lgt1* differs drastically from that of the other two enzymes (Fig. 4a and b). The similarities found by DALI are limited exclusively to the central domain of *Lgt1*, which shows homologies to the common fold of GT-A, although sequence homologies to the members of this family are very low. A structure-based alignment of *Lgt1* and toxin B reveals their kinship for approximately 230 amino acid residues in the central part of both proteins, albeit with a relatively high rmsd of 3.42 Å for the C $\alpha$  atoms (Fig. 4c). Like most members of the GT-A family, *Lgt1* shares the DXD motif (D246-X-D248 of *Lgt1*) with clostridial glucosylating cytotoxins. In clostridial glucosylating cytotoxins, this motif is involved in Mn<sup>2+</sup>, UDP, and glucose binding. The same holds true for *Lgt1*, with the exception that *Lgt1* was crystallized with a magnesium ion bound. However, D248 does not have the same functional importance as D246, as the D248A or D248N variants exhibited only 70% and 80% of wild-type glucosyltransferase activity, respectively, whereas

**Table 1.** Glucosylation by wild-type *L. pneumophila* glucosyltransferase *Lgt1* and mutants

	UDP-Glc (%)
Wild type	100.0
D246A	0.2±0.3
D246N	0.4±0.3
D248A	72.8±6.5
D248N	84.4±3.2
D246A/D248A	0.7±0.5
D246N/D248N	0.0
W520A	0.4±0.6
W520F	11.4±1.9
W520H	6.5±3.3

The glutathione *S*-transferase fusion peptide (eEF1A fragment 29 Y-73I; 3  $\mu$ M), which is a preferred substrate, was incubated with wild-type *Lgt1* from *L. pneumophila* strain Lens, and the indicated mutant enzyme protein (1  $\mu$ M) was incubated with UDP-[<sup>14</sup>C]Glc. Incubation was performed for 2 min (at 37 °C) to meet the linear phase of the glucosylation reaction. The labeled proteins were then analyzed by SDS-PAGE and PhosphorImaging. Data are given as percentage of wild-type enzyme activity (100%)  $\pm$ SD ( $n=3$ ).

D246A and D246N were almost completely inactive (Table 1). Another example of structural similarity is that between residues D230 and R233 within the



**Fig. 4.** Structural comparison of *Lgt1* with toxin B from *C. difficile*. (a) Cartoon representation of *Lgt1*. (b) Cartoon representation of *C. difficile* toxin B. (c) Stereo representation of the superposition of the C $\alpha$  positions of both toxins, with *Lgt1* in red and with toxin B in black. Orientations are identical with (a) and (b). The structural homology is fully restricted to the central glucosyltransferase module of both proteins.

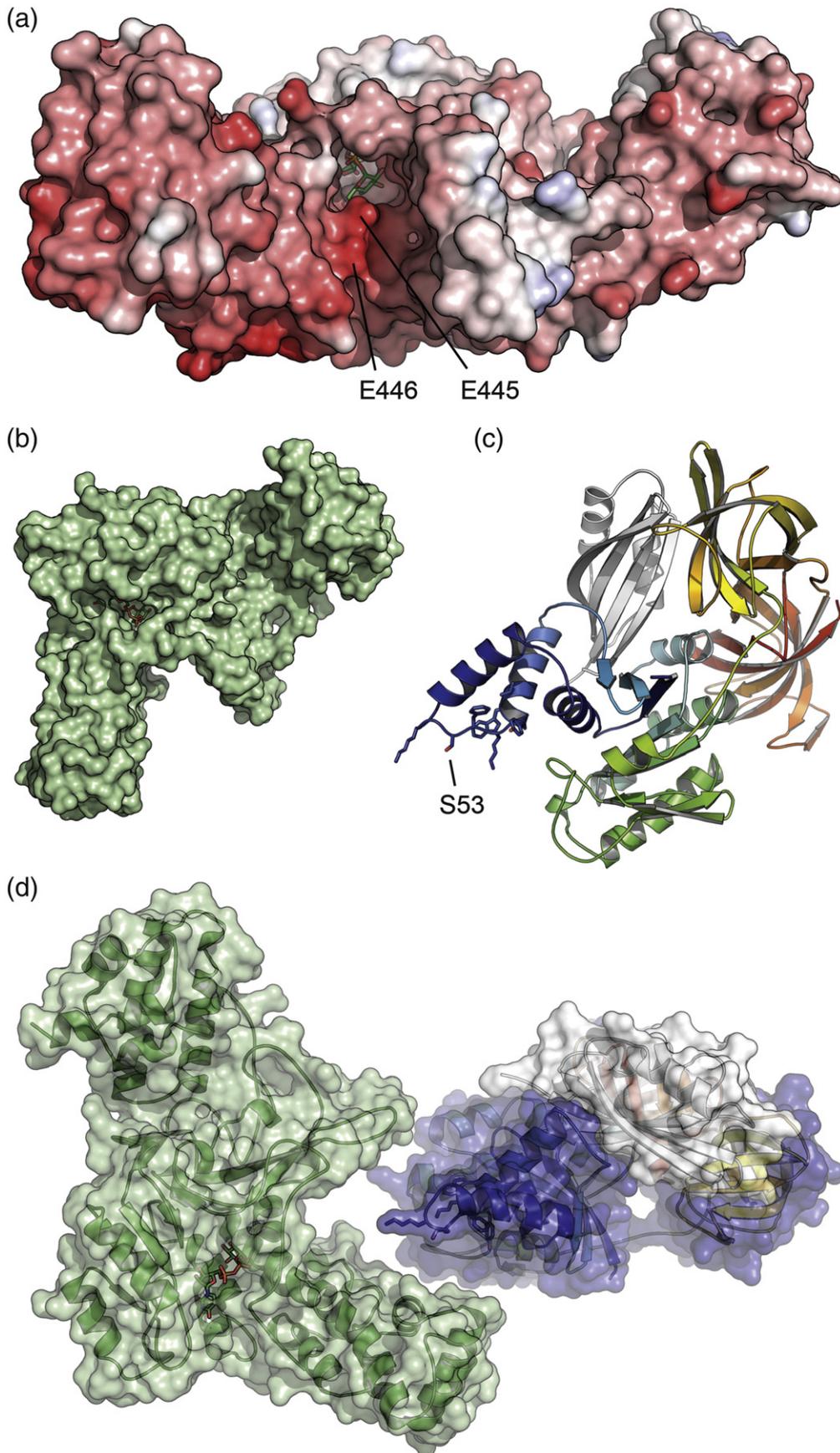


Fig. 5 (legend on next page)

glucose recognition triad of Lgt1 (Fig. 3), which are structurally and functionally equivalent to residues D270 and R273 of *C. difficile* toxin B, a clostridial glucosylating cytotoxin. Moreover, in toxin B, W520 was shown to belong to a “flexible loop,” which undergoes major conformational changes dependent on the cosubstrate bound, indicating an open and closed conformation of the glucosyltransferase. The equivalent W520 of Lgt1 that binds to the  $\alpha$ -phosphate in the UDP-Glc-bound form (Fig. 3a) is in a disordered state in the UDP form of Lgt1 (Fig. 3b) and appears to swing out to form an open conformation. It has been suggested that W520 is involved in the mechanism of the glucosylation reaction catalyzed by clostridial glucosylating toxins,<sup>18</sup> but our data on Lgt1 do not support a similar mechanism in this case. Exchange of W520 for alanine resulted in a reduction in enzyme activity by >95%, but residual enzymatic activity was retained in the W520F and W520H mutants. W520 may therefore be important for substrate binding and possibly for the formation of specific interactions with the substrate, but it is not directly involved in catalysis (Table 1). The binding mode observed for UDP and glucose in *C. difficile* toxin B was found to be analogous to the one for UDP-Glc in Lgt1 (Fig. S4).

In the glycosylating toxins cytotoxin L and toxin B, the inserts and protrusions in which the proteins differ from the basic glycosyltransferase fold have been suggested to be the structural determinants for the recognition of target proteins.<sup>18</sup> While both clostridial toxins have been shown to glucosylate human Rho/Ras GTPases,<sup>25,26</sup> the target of Lgt1 differs substantially, and the structural changes in the protein fold likely constitute an adaptation to eEF1A.

Both toxin B and Lgt1 contain a distinct  $\alpha$ -helical N-terminal domain that does not form part of the general fold of GT-A. It constitutes residues 1–90 in toxin B and residues 1–104 in Lgt1; however, while the domain of toxin B folds into an anti-parallel four-helix bundle, the N-terminal domain of Lgt1 consists of a total of six helical fragments and is topologically unrelated. In both cases, the N-terminal domains do not face the side of the enzyme where interaction with the target protein takes place; on direct comparison, they locate at very different positions with respect to the glycosyltransferase domain (Fig. 4). A role for this domain has so far not been assigned, and further studies will be required in order to clarify its significance.

The active-site environment of Lgt1 shows strong similarities with that of toxin B, in particular with respect to the orientation of the UDP-Glc ligand. A notable difference lies in the presence of two

negatively charged residues, E445 and E446, located at the funnel-like entrance to the active site (Fig. 5a). This distinct negative patch might play a role in the recognition and binding of the target protein; indeed, 50-GKGSFKYAWV-59, the minimal peptide fragment of eEF1A required for binding to Lgt1, contains two lysine residues. Note that exchange of any of the lysine residues results in only a modest reduction in the ability of the peptide to serve as a substrate for Lgt1. A much stronger effect was observed when F or W was exchanged for A. This might be due to rearrangements of the S53 loop of eEF1A upon removal of these bulky side chains, but it could also indicate an interaction of the aromatic residues with Lgt1. This could take place in the form of  $\pi$ -stacking interactions (e.g., with the indole moiety of W520), but direct experimental or structural evidence for this interaction remains to be provided.

In order to elucidate the mechanism of peptide binding, we attempted the cocrystallization of Lgt1 with the decapeptide and obtained crystals, but the peptide fragment was not visible in the structure. To prevent the reaction from taking place in the crystal, we used an inactive variant of Lgt1, N293A, for these experiments. The failure to form a complex with the target peptide may, in part, be due to the extended conformation of the C-terminus of Lgt1 that may have to be rearranged for complex formation with eEF1A.

## Discussion

### The N293A variant of Lgt1

The conserved residue N384 was discussed to be essential for catalytic activity in *C. difficile* toxin B,<sup>20</sup> and the same is true for its counterpart N293 in *L. pneumophila* Lgt1 (Fig. S3). We therefore created a variant protein N293A with the goal of obtaining a complex of the glucosyltransferase with a substrate peptide. However, the structural analysis revealed N293 not be located in immediate proximity to the UDP-Glc ligand, but rather in the obvious access pathway for the substrate eEF1A (see the text below). It is therefore assumed that the role of N293 lies in the guidance and/or binding of the substrate to the active site of Lgt1, and that it does not participate in the mechanism of glucosyl transfer itself. Inactivity of the above variants is thus likely due to reduced binding of substrate, such that a new strategy for obtaining a Lgt1–eEF1A complex will have to be devised.

**Fig. 5.** Glucosylation of eEF1A by *L. pneumophila* Lgt1. (a) Surface representation of Lgt1, colored according to the electrostatic surface potential [contoured from  $-10k_B T$  (red) to  $+10k_B T$  (blue)]. Two negatively charged residues, E445 and E446, line the entrance to the binding site for UDP-Glc. (b) Top view of Lgt1 highlighting the binding cleft for the target protein. UDP-Glc is bound at the bottom of this cleft. (c) Cartoon representation of eEF1A (PDB ID 2B7C)<sup>27</sup> shown from blue at the N-terminus to red at the C-terminus. The structure shows a complex with eukaryotic elongation factor 1B $\alpha$ . Side chains are depicted for the decapeptide 50-GKGSFKYAWV-59 of eEF1A, which is the minimally recognized peptide for Lgt1. S53 is the residue that is glucosylated by Lgt1. (d) Hypothetical model for the interaction of Lgt1 with eEF1A. The relative orientation of the proteins was chosen based on surface complementarity considerations.

In toxin B, the role of N384 may well be different, as this residue, although corresponding in sequence to N293 of Lgt1, is located closer to the UDP-Glc moiety and is placed directly below the sugar ring, albeit without forming any direct hydrogen bonds. The finding that this asparagine, although conserved, might function very differently in both glucosyltransferases underlines the importance of studying other members of this diverse class of proteins.

### Glycosylation of eEF1A by Lgt1

In order to perform its physiological function, Lgt1 has to recognize and specifically bind to eEF1A. Surface complementarity is expected to play a major role in this specific interaction, and a second important factor may be the complementarity of protein surface charges. The electrostatic surface potential of Lgt1, as calculated with DELPHI,<sup>28</sup> is distinctly negative, surrounding the obvious entrance to the binding site for UDP-Glc (Fig. 5a). The aforementioned residues E445 and E446 are in a prominent position to interact with the positively charged lysines in the recognition peptide. In a surface view of Lgt1, the clasp-like access funnel is seen as a dominant structural feature of the protein (Fig. 5b); in the target eEF1A, the fragment surrounding the glucosylation site at S53 is located at the periphery of the very N-terminal domain of the elongation factor (Fig. 5c). The crystal structure of eEF1A in complex with eukaryotic elongation factor 1B $\alpha$  (PDB ID 2B7C)<sup>27</sup> shows this site to be clearly exposed, making it the obvious target for the binding of Lgt1. The role of the negative electrostatic surface potential of Lgt1 is then to promote complex formation, as nucleotide-binding proteins such as elongation factors typically show a positive electrostatic surface potential (Fig. S5).

Based on the requirements of a close approach of Lgt1 and eEF1A and surface charge complementarity considerations, manual docking has been performed in order to assess the possible architecture of the glycosylation complex (Fig. 5d). Without any further modifications, the protein surfaces of the two binding partners are a clear match, with the binding cleft of Lgt1 being able to accommodate the exact surroundings of S53 of eEF1A without a need for major conformational rearrangements. We have refrained from presenting a more detailed docking model mainly due to the fact that the C-terminus of Lgt1 is not in a position to allow for an optimal interaction with eEF1A, and it can be assumed that in the actual complex of both proteins, the interaction with eEF1A induces a different conformation of this region of the glucosyltransferase. This presumed mode of interaction of the proteins explains the obvious protrusions observed in the structure of Lgt1 as an optimization of affinity. In a toxin/target complex such as the one of Lgt1 with eEF1A, the evolutionary pressure working on both binding partners is divergent: While the toxin will be selected for increased affinity, the target protein will try to evade by increasing the variability of the interface.<sup>19</sup> It is then not surprising to find a tendency in the toxin

to increase the interaction surface and, in the long run, to create structures such as the protrusions observed around the glycosyltransferase domain. The target protein, on the other hand, is very limited in its freedom to vary its surface, as it needs to retain its physiological function.

## Experimental Procedures

### Cloning of genes

The gene Lpl1319 coding for *L. pneumophila* Lgt1 was PCR-amplified with PfuII Turbo (Stratagene) from the genomic DNA of *L. pneumophila* strain Lens (accession no. NC\_006369) using the primers Lpl1319\_fw (gccGGATCC-atgaaagcaagaaggagtaacg) and Lpl1319\_rev (gccCTCG-AGctaccctactgaaggcaacc). The PCR product was cleaved with EcoRI and XhoI restriction endonucleases and ligated to the corresponding sites of the vector pET28a-TEV, which is a modified version of the pET28a vector (Novagen) containing a TEV protease cleavage site between the N-terminal His tag and the authentic protein sequence. In order to mutate Lpl1319 into Lpl1319-N293A, we performed site-directed mutagenesis, using Quikchange (Stratagene), with the primers listed in Table S1. The other glucosyltransferase variants listed in Table S1 were constructed accordingly. The sequence of the corresponding plasmids was confirmed at GATC (Konstanz, Germany).

### Expression and purification of recombinant proteins

*Escherichia coli* BL21\* CodonPlus cells (Stratagene) transformed with the respective plasmids were grown in LB medium to an OD<sub>600</sub> of approximately 0.6. Protein expression was induced by adding IPTG up to a final concentration of 1 mM, and the cells were grown for a further 4 h at 23 °C. The cells were resuspended in lysis buffer [20 mM Tris-HCl (pH 7.4), 150 mM NaCl, 25 mM imidazole, 30  $\mu$ g/ml DNase I, 10 mM  $\beta$ -mercaptoethanol, and 1 mg/ml lysozyme] supplemented with Proteinase Inhibitor Cocktail (Roche) and lysed by sonication. Subsequently, the cell lysate was applied onto a 1-ml HisTrap HP column (GE Healthcare) equilibrated in buffer A [20 mM Tris-HCl (pH 7.4), 150 mM NaCl, 10 mM  $\beta$ -mercaptoethanol, and 25 mM imidazole] and eluted with a linear gradient from 90 mM to 500 mM imidazole. The protein-containing fractions were pooled and dialysed against a dialysis buffer [20 mM Tris-HCl, 150 mM NaCl, 0.5 mM dithiothreitol, and 10% glycerol (pH 7.4)]. The proteins were >95% pure, as estimated by SDS-PAGE and Coomassie staining, eluting in a single symmetric peak from a Superdex 200 column.

SeMet labeling of Lpl1319-N293A was performed using the autoinduction medium PASM5052 by Studier.<sup>29</sup> *E. coli* BL21\* CodonPlus cells transformed with the respective plasmids were grown for 36 h at 23 °C in PASM5052 and subsequently harvested. Purification of the protein was carried out as described above.

### Glucosyltransferase assay

Assays were carried out as published previously.<sup>14</sup> Briefly, 1  $\mu$ M recombinant enzyme was incubated with

3  $\mu$ M glutathione S-transferase eEF1A deletion peptide (29-YKCGIDKRTIEKFEKEAAEMGKGSFKYAWVLDKLIK-AERERGITI-73) as substrate in 20 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1 mM MnCl<sub>2</sub>, and 10  $\mu$ M [<sup>14</sup>C]UDP-glucose (American Radiolabeled Chemicals, St. Louis, MO) for 2 min at 37 °C. The reaction was stopped by adding Laemmli buffer and boiling the samples at 100 °C for 5 min. The samples were then subjected to SDS-PAGE and analyzed by PhosphorImaging (PhosphorImager Storm 820; Molecular Dynamics, Vienna, Austria).

### Crystallization and data collection

Both wild-type Lgt1 and the N293A variant were crystallized using the sitting-drop vapor-diffusion method. One microliter of protein solution (4 mg/ml) was mixed with 1  $\mu$ l of reservoir solution containing 23% (wt/vol) of polyethylene glycol 3350, 0.06 M ammonium acetate, and 0.1 M sodium acetate buffer at pH 5.0. As an additive, 0.2  $\mu$ l of 2 M nondetergent sulfobetaine 211 (NDSB-211; Hampton Research, Laguna Niguel, CA) was added to the drop, and the mixture was equilibrated against the reservoir solution. Single crystals appeared within 2 days and reached their maximum size after approximately 4 days. Crystals of both wild-type and SeMet-labeled proteins were obtained under identical conditions. Crystals were harvested into a reservoir buffer, and the open drop was left on air for 5 min such that some water evaporated and the remaining buffer proved suitable as a cryoprotectant for low-temperature diffraction experiments. Data were collected at beamline X06SA at the Swiss Light Source (Villigen, Switzerland), with SeMet-labeled crystals diffracting to a maximum resolution of 1.7 Å and with native crystals diffracting to a maximum resolution of 2.3 Å. For phase determination, a four-wavelength multiple-wavelength anomalous dispersion experiment was carried out at the K-absorption edge of selenium to a maximum resolution of 1.7 Å (Table 2). All data sets were processed using MOSFLM<sup>30</sup> and scaled with SCALA.<sup>30</sup>

### Structure solution and refinement

The structure of Lgt1 N293A, in complex with UDP-Glc, was solved by SeMAD. The positions of anomalous scatterers were determined using SHELXD,<sup>31</sup> and SHARP<sup>32</sup> was used for phase-angle calculations. SHELXD correctly identified 10 of 15 Se sites in the asymmetric unit, and site refinement yielded a phasing power of 1.373 at a figure of merit of 0.501 with SHARP. Phase improvement was carried out with SOLOMON,<sup>30</sup> and the electron density maps obtained were readily interpretable. A molecular model for Lgt1 was built using Coot<sup>33</sup> and refined with REFMAC.<sup>34</sup> The model consisted of residues 6–525 of the protein sequence; the first five residues, as well as the N-terminal TEV protease cleavage site and the hexahistidine affinity tag, were disordered. The final model of Lgt1-N293A was refined to an *R*-factor of 0.166 (*R*<sub>free</sub>=0.193) at a resolution of 1.7 Å, with all residues falling within the allowed or additionally allowed regions of the Ramachandran plot (data not shown). The wild-type protein, in complex with UDP, was refined to an *R*-factor of 0.185 (*R*<sub>free</sub>=0.262) at 2.3 Å resolution. All figures were prepared using PyMOL.<sup>35</sup> During purification and crystallization, neither Mg<sup>2+</sup> nor Mn<sup>2+</sup> was added. Nevertheless, a Mg<sup>2+</sup> bound to the phosphate groups of UDP-glucose was modeled in the N293A variant protein structure, based on a near-octahedral coordination geometry and a total electron density that was too low for Mn<sup>2+</sup>. We acknowledge that this is not an unambiguous identification of Mg<sup>2+</sup>; we refrain from drawing conclusions regarding the function of the protein. The nucleotide-sugar moieties were fitted manually into the clear difference electron density features and refined using REFMAC.

### Accession numbers

Coordinates and structure factors have been deposited in the PDB with accession numbers 3JSZ (Lgt1\_N293A with UDP-Glc) and 3JT1 (Lgt1 with UDP).

**Table 2.** Data collection and refinement statistics for Lgt1

Construct data set	Lgt1-N293A SeMet				Lgt1 wild-type native
	Se $\lambda_{lowE}$	Se $\lambda_{infection}$	Se $\lambda_{peak}$	Se $\lambda_{highE}$	
Wavelength (Å)	1.00000	0.97973	0.97957	0.97205	1.00000
Resolution (Å)	22.0–1.7 (1.8–1.7)	23.6–1.9 (2.0–1.9)	23.6–1.9 (2.0–1.9)	23.6–1.9 (2.0–1.9)	31.8–2.3 (2.4–2.3)
Number of reflections	433,102 (60,773)	402,084 (51,514)	410,579 (54,369)	409,984 (53,308)	135,112 (19,778)
Completeness (%)	99.9 (100.0)	99.9 (99.9)	99.9 (99.9)	99.9 (99.9)	99.8 (99.8)
Multiplicity	6.9 (6.6)	8.9 (7.9)	9.1 (8.3)	9.1 (8.1)	5.1 (5.1)
Space group	R3	R3	R3	R3	R3
Unit cell axes					
<i>a=b</i>	122.23	122.23	122.23	122.23	122.57
<i>c</i>	102.78	102.78	102.78	102.78	103.98
<i>R</i> <sub>merge</sub>	0.062 (0.391)	0.094 (0.482)	0.094 (0.468)	0.085 (0.434)	0.102 (0.657)
<i>R</i> <sub>pim</sub>	0.027 (0.176)	0.037 (0.194)	0.039 (0.185)	0.033 (0.172)	0.054 (0.357)
<i>I</i> / $\sigma$ ( <i>I</i> )	7.3(1.9)	6.5(1.5)	6.5(1.6)	6.9(1.7)	8.4 (2.1)
Ligand	UDP-glucose				UDP
<i>R</i> <sub>cryst</sub>	0.166 (0.196)				0.185 (0.277)
<i>R</i> <sub>free</sub>	0.193 (0.231)				0.262 (0.317)
rmsd bond lengths (Å)	0.008				0.023
rmsd bond angles (°)	1.182				1.949
Average <i>B</i> -factor (Å <sup>2</sup> )	17.9				38.6
Diffraction precision index (Å)	0.061				0.205

Values in parentheses represent the highest-resolution shell.

## Acknowledgements

The authors wish to thank Elke Gerbig-Smentek for excellent technical assistance. Diffraction data were collected at beamline X06SA at the Swiss Light Source. This work was supported by Deutsche Forschungsgemeinschaft (Ei-520/3 to O.E. and Ak-6/17 to K.A.).

## Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2009.11.044](https://doi.org/10.1016/j.jmb.2009.11.044)

## References

- Fields, B. S., Benson, R. F. & Besser, R. E. (2002). *Legionella* and Legionnaires' disease: 25 years of investigation. *Clin. Microbiol. Rev.* **15**, 506–526.
- Marston, B. J., Plouffe, J. F., File, T. M., Hackman, B. A., Salstrom, S. J., Lipman, H. B. *et al.* (1997). Incidence of community-acquired pneumonia requiring hospitalization—results of a population-based active surveillance study in Ohio. *Arch. Intern. Med.* **157**, 1709–1718.
- Horwitz, M. A. (1983). The Legionnaires' disease bacterium (*Legionella pneumophila*) inhibits phagosome-lysosome fusion in human monocytes. *J. Exp. Med.* **158**, 2108–2126.
- Horwitz, M. A. & Maxfield, F. R. (1984). *Legionella pneumophila* inhibits acidification of its phagosome in human monocytes. *J. Cell Biol.* **99**, 1936–1943.
- Summersgill, J. T., Raff, M. J. & Miller, R. D. (1988). Interactions of virulent and avirulent *Legionella pneumophila* with human polymorphonuclear leukocytes. *Microb. Pathog.* **5**, 41–47.
- Coers, J., Monahan, C. & Roy, C. R. (1999). Modulation of phagosome biogenesis by *Legionella pneumophila* creates an organelle permissive for intracellular growth. *Nat. Cell Biol.* **1**, 451–453.
- Horwitz, M. A. (1983). Formation of a novel phagosome by the Legionnaires' disease bacterium (*Legionella pneumophila*) in human monocytes. *J. Exp. Med.* **158**, 1319–1331.
- Byrne, B. & Swanson, M. S. (1998). Expression of *Legionella pneumophila* virulence traits in response to growth conditions. *Infect. Immun.* **66**, 3029–3034.
- Gao, L. Y. & Abu Kwaik, Y. (1999). Apoptosis in macrophages and alveolar epithelial cells during early stages of infection by *Legionella pneumophila* and its role in cytopathogenicity. *Infect. Immun.* **67**, 862–870.
- Gao, L. Y. & Abu Kwaik, Y. (2000). The mechanism of killing and exiting the protozoan host *Acanthamoeba polyphaga* by *Legionella pneumophila*. *Environ. Microbiol.* **2**, 79–90.
- Molmeret, M., Bitar, D. M., Han, L. H. & Abu Kwaik, Y. (2004). Disruption of the phagosomal membrane and egress of *Legionella pneumophila* into the cytoplasm during the last stages of intracellular infection of macrophages and *Acanthamoeba polyphaga*. *Infect. Immun.* **72**, 4040–4051.
- Andrews, H. L., Vogel, J. P. & Isberg, R. R. (1998). Identification of linked *Legionella pneumophila* genes essential for intracellular growth and evasion of the endocytic pathway. *Infect. Immun.* **66**, 950–958.
- Belyi, I., Popoff, M. R. & Cianciotto, N. P. (2003). Purification and characterization of a UDP-glucosyltransferase produced by *Legionella pneumophila*. *Infect. Immun.* **71**, 181–186.
- Belyi, Y., Niggeweg, R., Opitz, B., Vogelsang, M., Hippenstiel, S., Wilm, M. & Aktories, K. (2006). *Legionella pneumophila* glucosyltransferase inhibits host elongation factor 1A. *Proc. Natl Acad. Sci. USA*, **103**, 16953–16958.
- Belyi, Y., Tabakova, I., Stahl, M. & Aktories, K. (2008). Lgt: a family of cytotoxic glucosyltransferases produced by *Legionella pneumophila*. *J. Bacteriol.* **190**, 3026–3035.
- Belyi, Y., Stahl, M., Sovkova, I., Kaden, P., Luy, B. & 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.
- Belyi, Y., Niggeweg, R., Opitz, B., Vogelsang, M., Hippenstiel, S., Wilm, M. & Aktories, K. (2006). *Legionella pneumophila* glucosyltransferase inhibits host elongation factor 1A. *Proc. Natl Acad. Sci. USA*, **103**, 16953–16958.
- Ziegler, M. O. P., Jank, T., Aktories, K. & Schulz, G. E. (2008). Conformational changes and reaction of clostridial glycosylating toxins. *J. Mol. Biol.* **377**, 1346–1356.
- Reinert, D. J., Jank, T., Aktories, K. & Schulz, G. E. (2005). Structural basis for the function of *Clostridium difficile* toxin B. *J. Mol. Biol.* **351**, 973–981.
- Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. (2008). Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**, 521–555.
- Unligil, U. M. & Rini, J. M. (2000). Glycosyltransferase structure and mechanism. *Curr. Opin. Struct. Biol.* **10**, 510–517.
- Vetter, I. R., Hofmann, F., Wohlgenuth, S., Herrmann, C. & Just, I. (2000). Structural consequences of mono-glucosylation of Ha-Ras by *Clostridium sordellii* lethal toxin. *J. Mol. Biol.* **301**, 1091–1095.
- Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R. & Davies, G. J. (2002). Insights into trehalose synthesis provided by the structure of the retaining glucosyltransferase OtsA. *Chem. Biol.* **9**, 1337–1346.
- Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138.
- Just, I., Selzer, J., Hofmann, F., Green, G. A. & Aktories, K. (1996). Inactivation of Ras by *Clostridium sordellii* lethal toxin-catalyzed glucosylation. *J. Biol. Chem.* **271**, 10149–10153.
- Just, I., Selzer, J., Wilm, M., von Eichel-Streiber, C., Mann, M. & Aktories, K. (1995). Glucosylation of Rho-proteins by *Clostridium difficile* toxin B. *Nature*, **375**, 500–503.
- Pittman, Y. R., Valente, L., Jeppesen, M. G., Andersen, G. R., Patel, S. & Kinzy, T. G. (2006). Mg<sup>2+</sup> and a key lysine modulate exchange activity of eukaryotic translation elongation factor 1B alpha. *J. Biol. Chem.* **281**, 19457–19468.
- Honig, B. & Nicholls, A. (1995). Classical electrostatics in biology and chemistry. *Science*, **268**, 1144–1149.
- Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expression Purif.* **41**, 207–234.
- Collaborative Computational Project No. 4. (1994). The CCP4 Suite: programs for protein crystallography. *Acta Crystallogr. Sect. D*, **50**, 760–763.
- Schneider, T. R. & Sheldrick, G. M. (2002). Substructure solution with SHELXD. *Acta Crystallogr. Sect. D*, **58**, 1772–1779.
- La Fortelle, E. d., Irwin, J. J. & Bricogne, G. (1997).

- SHARP: a maximum-likelihood heavy-atom parameter refinement and phasing program for the MIR and MAD methods. *Crystallogr. Comput.* **7**, 1–9.
33. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D*, **60**, 2126–2132.
  34. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D*, **53**, 240–255.
  35. DeLano, W. L. (2002). *The PyMOL Molecular Graphic System*. DeLano Scientific, San Carlos, CA.