

# Mechanism of Phosphatidylglycerol Activation Catalyzed by Prolipoprotein Diacylglyceryl Transferase

Warispreet Singh,<sup>†,‡,⊥</sup> Munir Bilal,<sup>§,⊥</sup> James McClory,<sup>†</sup> Daniel Dourado,<sup>†,‡</sup> Derek Quinn,<sup>‡</sup> Thomas S. Moody,<sup>‡,||</sup> Iain Sutcliffe,<sup>§</sup> and Meilan Huang<sup>\*,†,||</sup>

<sup>†</sup>School of Chemistry & Chemical Engineering, Queen's University Belfast, Belfast BT9 5AG, Northern Ireland, United Kingdom

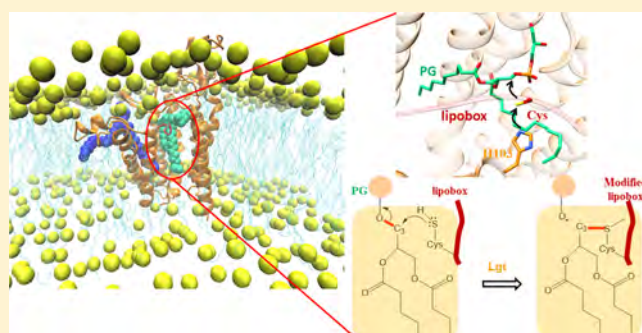
<sup>‡</sup>Department of Biocatalysis and Isotope Chemistry, Almac Sciences, Almac House, 20 Seagoe Industrial Estate, Craigavon BT63 5QD, Northern Ireland, United Kingdom

<sup>§</sup>Department of Applied Sciences, Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, United Kingdom

<sup>||</sup>Arran Chemical Company Limited, Unit 1 Monksland Industrial Estate, Athlone, Co., Roscommon, N37 DN24, Ireland

## Supporting Information

**ABSTRACT:** Lipoproteins are essential for bacterial survival. Bacterial lipoprotein biosynthesis is accomplished by sequential modification by three enzymes in the inner membrane, all of which are emerging antimicrobial targets. The X-ray crystal structure of prolipoprotein diacylglyceryl transferase (Lgt) and apolipoprotein *N*-acyl transferase (Lnt) has been reported. However, the mechanisms of the post-translational modification catalyzed by these enzymes have not been understood. Here, we studied the mechanism of the transacylation reaction catalyzed by Lgt, the first enzyme for lipoprotein modification using molecular docking, molecular dynamics, and quantum mechanics/molecular mechanics (QM/MM) calculations. Our results suggest that Arg143, Arg239, and Glu202 play a critical role in stabilizing the glycerol-1-phosphate head group and activating the glycerol C3–O ester bond of the phosphatidylglycerol (PG) substrate. With PG binding, the opening of the L6–7 loop mediated by the highly conserved Arg236 residue as a gatekeeper is observed, which facilitates the release of the modified lipoprotein product, as well as the entry of another PG substrate. Further QM/MM studies revealed that His103 acts as a catalytic base to abstract a proton from the cysteine residue of the prelipoprotein, initiating the diacylglyceryl transfer from PG to prelipoprotein. This is the first study on the mechanism of lipoprotein modification catalyzed by a post-translocational processing enzyme. The transacylation mechanism of Lgt would shed light on the development of novel antimicrobial therapies targeting the challenging enzymes involved in the post-translocational modification pathway of lipoproteins.



## INTRODUCTION

Bacterial lipoproteins are a functionally diverse class of peripheral membrane proteins that are well represented in prokaryotic proteomes.<sup>1–3</sup> During post-translocation, bacterial lipoprotein precursors are processed by the Sec or TAT systems and then modified by three enzymes sequentially. The first enzyme is prolipoprotein diacylglyceryl transferase (Lgt), which transfers a diglyceride from phospholipids onto an invariant cysteine of the signal peptide in precursor lipoproteins (prelipoproteins) anchored on cell membranes<sup>2,4,5</sup> (Figure 1A). Lipoprotein signal peptidase (LspA) then cleaves the signal peptide from diacylglyceryl-modified prelipoproteins<sup>6,7</sup> and finally, in many but not all taxa, apolipoprotein *N*-acyl transferase (Lnt) catalyzes the *N*-acylation of the apolipoprotein.<sup>8,9</sup> The post-translational modification pathways for processing functionally diverse lipoproteins encoded in bacterial genomes are conserved

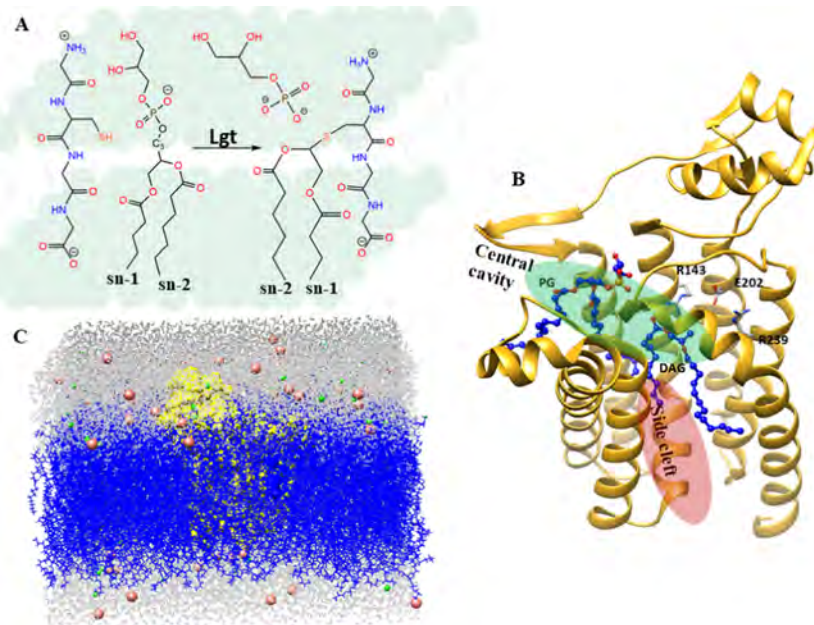
across the bacterial domain and are essential for bacterial physiology; they are unique to bacteria and thus represent potentially novel antimicrobial targets.<sup>10</sup>

Lgt, the gate and defining integral membrane protein of the pathway, recognizes a consensus sequence “lipobox” ([LVI]-[ASTVI]-[GAS]-C) located at the N-terminal signal peptide of prelipoproteins. The recently elucidated X-ray crystallographic structure of an *Escherichia coli* Lgt enzyme in complex with phosphatidylglycerol (PG) and diacylglyceryl (DAG, a hydrolyzed product of the PG in the active site) molecules (PDB: 5AZC<sup>2</sup>) revealed that Lgt has seven transmembrane (TM) helices, with a significant periplasmic domain (between TM4 and TMs) and two other loops forming “arms” (the first

Received: May 4, 2019

Revised: July 24, 2019

Published: July 25, 2019



**Figure 1.** (A) Schematic representation of the transfer of the diacylglycerol group from the phosphatidylglycerol (PG) to the conserved cysteine of the prelipoprotein in the lipid bilayer catalyzed by Lgt enzyme. (B) X-ray crystal structure of Lgt enzyme (pdb: 5AZC). The protein is shown in the cartoon representation using Chimera. The phosphatidylglycerol (PG) molecule in the first binding site and the diacylglyceryl (DAG) in the active site of Lgt are shown in blue. The active site residue of the Lgt enzymes Arg143, Arg239, and Glu202 are shown in the stick representation. The central cavity is shown in the green shade and the side cleft is illustrated by the red shade. (C) Lgt in the 1-palmitoyl-2-oleoyl-phosphatidylethanolamine (POPE) lipid bilayer. Lgt is shown by the surface representation in yellow and lipid bilayers shown in blue. The water molecules, sodium ion, and chloride ion are shown in gray, green, and pink, respectively.

preceding TM1 and the other between TM2 and TM3), a central cavity with two binding sites for the PG substrate (Figure 1B).<sup>2</sup> It was proposed that PG moves from the first site to the second site. In the second binding site, the diacylglycerol moiety of PG transfers to the lipobox of the prelipoprotein and the modified lipobox exits from the side cleft of Lgt. However, in the crystal structure of Lgt, the second PG binding site (the active site) is occupied by a DAG molecule (the hydrolyzed product of PG) instead of a PG substrate. In addition, a lipobox characteristic of the prelipoprotein signal peptide is missing in the crystal structure, indicating that this structure may be catalytically unproductive.

Here, we investigated the binding mode of PG and the prelipoprotein in *E. coli* Lgt using molecular docking and molecular dynamics (MD) simulations. By using hybrid quantum mechanics/molecular mechanics (QM/MM) calculations, we elucidated the mechanism of lipid modification by Lgt and found that His103 functions as a catalytic base in the diglyceride transfer from PG to prelipoprotein, abstracting a proton from the conserved cysteine residue of the prelipoprotein. The present study provides important insight for lipid modification by Lgt and would shed light on the design of novel antimicrobial therapies targeting the key post-translocational modification enzymes.

## METHODS

**Structure Preparation and Molecular Docking.** The crystal structure of the *E. coli* Lgt (PDB: 5AZC<sup>2</sup>) in complex with the substrate PG and a DAG inhibitor was used as the initial structure. The DAG molecule at the second binding site was modified into a PG molecule by adding the head group (R) glycerol-1-phosphate using GaussView 5.0.<sup>11</sup> The protonation states of titratable residues were predicted using

H++ server<sup>12</sup> at pH 7 and was used in the subsequent modeling study. The His7 was protonated, and His24, His103, and His196 were in the neutral form with the epsilon nitrogen protonated (Table S1). The side chains of arginine and lysine residues were protonated, and all of the glutamic acid and aspartic acid side chains were deprotonated.

The parameters for the PG and DAG molecules were taken from the Lipid14 force field<sup>13</sup> implemented in the Amber 16 package.<sup>14</sup> The partial sequence of the substrate signal peptide of prelipoprotein, spanning the characteristic lipobox (GSTLLAGCSSN), was obtained from Mao et al. 2016<sup>2</sup> and docked into the active site of Lgt using the Flexpepdock.<sup>15,16</sup> Ten models were generated in total (Table S2), and the default parameters were adopted during the flexible docking. The prelipoprotein peptide was docked into the representative structure obtained from MD simulations of the Lgt–PG complex (i.e., Lgt bound with two PG molecules). Two different sites located at the (1) front cleft and (2) side cleft were considered, respectively.

**MD Simulations.** The docked poses were assessed for the liability for the diacylglycerol reaction to happen, based on the distance between the C3 atom of PG and cysteine sulfur of the lipobox and their orientation. The most favorable docked pose was selected for MD simulations. The productive atomistic MD simulations were performed at 310 K using the GPU version of the PMEMD<sup>17</sup> engine integrated with the Amber 16 package.<sup>14</sup> The MD simulations were conducted on three complex systems: (I) Lgt in complex with PG and DAG (Lgt–PG–DAG), (II) Lgt in complex with two PG molecules (Lgt–PG), and (III) Lgt in complex with two PG molecules and a docked prelipoprotein (Lgt–PG–lipobox). All of the MD simulations were performed in a lipid bilayer consisting of 1-

palmitoyl-2-oleoyl-phosphatidylethanolamine (POPE) (Figure 1C), which was adopted in the MD simulations of Lnt.<sup>8</sup>

The orientation of the Lgt protein in the lipid bilayer was assessed using the orientation of the protein in the membrane database.<sup>18</sup> The CHARMM-GUI lipid builder<sup>19</sup> was used to insert Lgt into the POPE lipid bilayer consisting of 250 POPE lipid molecules in each of the upper and lower leaflets of the bilayer. The TIP3P<sup>20</sup> water molecules were used to solvate the bilayer, and the counterions were added using Monte-Carlo simulations with a concentration of 0.15 M NaCl. The enzyme system consisted of a total size of  $\sim 128\,000$  atoms with a box dimension of  $125.81 \times 125.81 \times 85.36 \text{ \AA}^3$ .

The FF14SB<sup>21</sup> and Lipid14<sup>13</sup> force fields were used in all simulations. The entire system was placed in a rectangular box, and the periodic boundary conditions were employed. Long-range electrostatic interactions were calculated using the particle mesh Ewald method<sup>22</sup> with a cutoff of 10  $\text{\AA}$ , and van der Waals interactions were computed with the same real-space cutoff. The various systems were subjected to energy minimization using steepest descent (2000 steps) followed by conjugate gradient algorithms (1000 steps). The protein and the bilayer were first restrained using a potential of  $50 \text{ kcal mol}^{-1} \text{ \AA}^2$ , and only solvent and ions were allowed to relax. Then, a full minimization of the entire system was conducted using steepest descent (5000 steps) and conjugate gradient (2500 steps) methods.

The systems were then subjected to controlled heating from 0 to 100 K for 400 ps using the Langevin thermostat with a collision frequency of  $1 \text{ ps}^{-1}$  and a canonical ensemble (NVT). In the second round of the heating, the system was gently heated from 100 to 310 K using a Langevin thermostat for another 400 ps. Anisotropic Berendsen weak-coupling barostat was used to control the pressure during the heating. The bilayer was restrained using a harmonic potential of  $50 \text{ kcal mol}^{-1} \text{ \AA}^2$  during both heating cycles. The SHAKE algorithm was used to constrain all of the bonds involving hydrogen atoms.<sup>23</sup> To equilibrate the density and dimensions of the system, equilibration was performed at 310 K in an NPT ensemble for 50 ns without restraints on the solute molecules. The Berendsen barostat was used to maintain the pressure at 1 bar during the equilibration phase.<sup>24</sup> A 700 ns production MD run was performed for Lgt–PG–DAG, 2.5  $\mu\text{s}$  for Lgt–PG, and 150 ns for the Lgt–PG–lipobox in an NPT ensemble with a target pressure of 1 bar and a pressure coupling constant of 2 ps. The frames from the production run were saved every 20 ps. The trajectories were analyzed using CPPTRAJ,<sup>25</sup> VMD,<sup>26</sup> and UCSF Chimera.<sup>27</sup> The root-mean-square deviation (RMSD) of the backbone  $\text{C}\alpha$  atoms of the protein and the heavy atoms of the lipid bilayer were analyzed with reference to the minimized crystal structure. The root-mean-square fluctuations (RMSFs), electrostatic interactions, hydrogen bonding, principal component analysis, and cluster analysis were conducted using CPPTRAJ.<sup>25</sup>

**QM/MM Calculations.** The snapshots for the QM/MM calculations were obtained from the cluster analysis for the last 100 ns equilibrated MD trajectory of the Lgt enzyme in complex with the prelipoprotein. The criteria for the selections of the snapshot were based upon the distance between the side-chain  $\text{N}\epsilon$  atom of His103, the cysteine sulfur atom of the prelipoprotein, and the carbonyl atom C3 of the C3–O ester bond of the PG molecule (Figure 1A). To obtain a catalytically competent reactant complex, the MD snapshots were selected, where the thiol group of the cysteine

formed hydrogen bonds with the  $\text{N}\epsilon$  atom of His103, and the distance between the sulfur atom and the carbonyl C3 atom of C3–O ester bond was less than 4.0  $\text{\AA}$ . The lipid bilayer was excluded in the QM/MM setup, but the water molecules that were present near the PG binding site from the MD simulations were included in the setup.

These snapshots were first subjected to 2000 steps of steepest descent energy minimization followed by 2000 steps of conjugate gradient minimization using Amber 16.<sup>14</sup> The minimized snapshots were then prepared for QM/MM setup using TAO package<sup>28</sup> and optimized using ONIOM<sup>29,30</sup> in the Gaussian09.<sup>31</sup> The system was divided into two layers in the QM/MM calculations. The QM region was calculated with density functional theory (DFT) using B3LYP<sup>32–34</sup> functional and 6-31+G(d,p) basis set. A mechanical embedding method was applied to integrate the point charge from the MM layer into the QM layer. Single-point corrections were performed at the QM/MM level on the stationary points using the electrostatic embedding scheme and the B3LYP functional theory with 6-311G(3df,3pd) basis sets. Single-point calculations on the optimized structures of the stationary points were also performed using  $\omega\text{B97XD}$  functional<sup>35</sup> to include the empirical dispersion corrections with a high exchange for long-range electron–electron interactions.

The QM region is composed of His103, Arg143, Asn146, Glu202, Arg239, Tyr235, the cysteine residue of the prelipoprotein, and the PG molecule in the second binding site. The residues (including water molecules) that are within 10  $\text{\AA}$  of the QM region were allowed to move freely and the rest of the system was frozen during geometry optimization. The link atom approach was adopted to saturate the dangling bond in the QM/MM calculation. The reaction coordinates used in the potential energy scan were the distance between the sulfur atom of a cysteine residue of the prelipoprotein and the C3 carbon atom of the C3–O ester bond of the PG molecule. The optimized reactant structure was used as the starting points for potential energy scan, which was conducted along the designated reaction coordinates with a decrement of 0.1  $\text{\AA}$ . Minima and the transition states (TS) obtained from the potential energy surface (PES) were fully optimized and validated by frequency calculations. The TS was characterized by the presence of one imaginary frequency. Intrinsic reaction coordinate (IRC) calculations were performed to validate that the TS connects the reactant and product.

## RESULTS AND DISCUSSION

**Conformational Dynamics of Lgt in the Presence of PG.** Microbial lipoproteins are essential for the growth of Gram-negative bacteria, and lipid modification is an important post-translocational process that can be targeted for the development of antimicrobial therapies.<sup>10</sup> So far, there is little information on the reaction mechanism of lipid modification by lipoprotein modification enzymes during post-translocation process. Lgt is the first and gate enzyme in the lipid modification of prolipoproteins. To study the reaction mechanism of diacylglycerol transfer, it is necessary to obtain an active conformation membrane-bound Lgt in complex with the PG substrate and lipobox. However, in the resolved crystal structure of Lgt,<sup>2</sup> the active site is occupied by a DAG molecule (a PG analogue in the absence of the head group) and the prolipoprotein is also absent.

Here, we built the complex structure of Lgt bound with PG and studied the conformational dynamics by molecular

dynamics simulations. Multiple replicas of MD runs can give reliable conformational information.<sup>35</sup> Five replicas of 500 ns MD simulations were run. The structural fluctuation of the Lgt enzyme in complex with PG was examined by the RMSD analysis on the C $\alpha$  atoms of the protein with reference to the minimized crystal structure, which showed that the simulations of the Lgt enzyme in complex with a PG substrate or a DAG molecule at the second binding site reached equilibration around 200 ns with an average RMSD value of 2.0 Å (Figure S1).

The cluster analysis of the equilibrated trajectory of the Lgt–PG MD simulations showed that the two PG molecules interact with each other through flexible aliphatic chains (Figure S2). The flexibility of the residues of the Lgt enzyme was analyzed by the RMSF analysis based on the last 300 ns of the equilibrated MD simulation trajectory (Figure 2).

In Lgt, a large central cavity with a periplasmic exit is composed of two clefts, “front cleft” and “side cleft”.<sup>2</sup> While the enzyme is bound with a second PG (Lgt–PG), the  $\eta$ 2 loop region (residues 95–101) of arm-2 and the L6–7 region near the front cleft adopted a more open conformation in the Lgt–PG complex than that in the Lgt–PG–DAG complex, making

the front cleft more open, liable for the PG molecule to enter (Figures 3 and S3B). In addition, the loop region L6–7 (residues 245–254) displayed an obvious structural fluctuation throughout the 2.5  $\mu$ s MD simulations (Figure S3A). The comparison of the X-ray structure and MD-simulated structures of Lgt–PG–DAG complex and Lgt–PG complex showed that L6–7 loop in Lgt bound with a second PG molecule changed into a more open conformation than that of Lgt–PG–DAG complex (Figure S3B). The TM3 region at the side cleft of the Lgt–PG complex (residues 101–120), which harbors the catalytic residue His103,<sup>2,36</sup> showed increased flexibility in comparison with Lgt–PG–DAG complex (Figure 2). The high flexibility of the front cleft and the side cleft may allow the sliding of the PG molecule in the large central cavity from the first binding site to the catalytic site of Lgt.

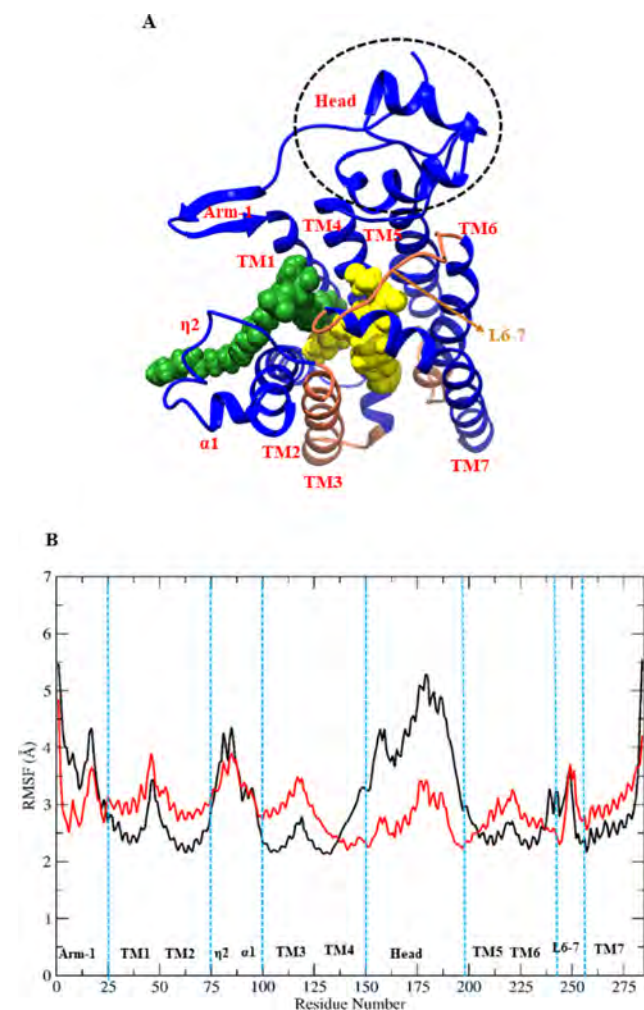
**First Binding Site of Lgt.** The binding mode of PG in the first binding site of Lgt was studied. Tyr26 located on the TM1 was suggested to be important for the Lgt activity.<sup>4</sup> In the crystal structure of Lgt, the hydroxyl group of the Tyr26 was 3.7 Å away from the O2 oxygen (ester oxygen of the sn2 lipid chain) of the PG. The side chain of Asn149 on TM4 and Glu151 formed a hydrogen bond with the glycerol head group of the PG molecule in the X-ray structure. With time evolution, PG in the MD-simulated structure of Lgt–PG moved away from these residues due to the extra room in the central cavity formed by the opening of loop L6–7 (Figure S4).

This indicates that the PG molecule is rather mobile after it is initially bound in the first binding site, liable to slide in the large binding cavity of the Lgt enzyme. In the X-ray structure, the side chain of the Arg73 was 6.2 Å away from the O11 of the PG sn-1 chain and was consistent with the average distance between Arg73 and O11 in the representative structures from the five replica runs of MD simulations (5.7 Å) (Figure S4). Superposition of the X-ray crystal structure and MD cluster structure of Lgt–PG showed a displacement of PG molecule in the first binding site toward the arm-2 of Lgt (Figure S5). These results indicate the mobile nature of the PG in the first binding site, which is attributed to the high flexibility of arm-2 and L6–7 at the front cleft of the enzyme.

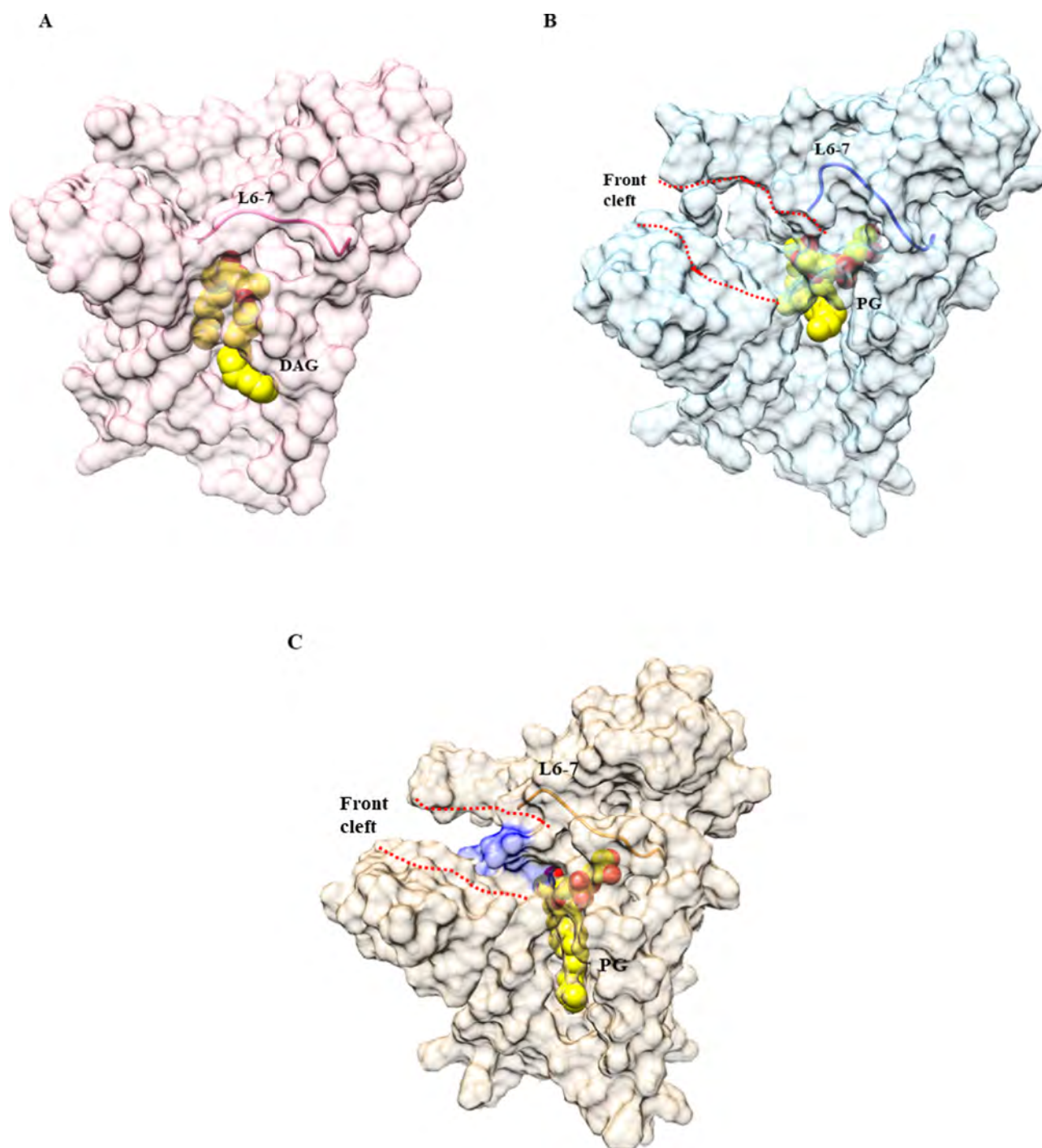
**Catalytic Site of Lgt.** The second binding site of PG in the Lgt enzyme is where catalysis takes place.<sup>2</sup> In the crystal structure of Lgt–PG–DAG complex, Arg239 formed a salt bridge interaction with Glu243 (Figure 4A), although it is far away from the DAG molecule. With time evolution, the DAG molecule in the Lgt–PG–DAG complex flipped far away from the active site (Figure 4B).

In the crystal structures of Lgt bound with a DAG and an inhibitor palmitic acid (PDB: 5AZC and 5AZB) at the second site, a highly conserved residue Arg239 on TM6 exhibited multiple conformations.<sup>2</sup> Interestingly, during the MD simulations of Lgt in complex with a PG molecule (with the glycerol head group added to the DAG molecule) in the second binding site, Arg239 turned toward PG, forming an electrostatic interaction with the phosphate of the glycerol head group of PG with an average distance of 3.8 Å (Figures 4C and S6). Thus, the binding of the PG molecule in the second site induced a conformational change of Arg239, which stabilizes the negatively charged head group of the PG by electrostatic interactions.

In the X-ray crystal structure, Ser198 pointed away from the catalytic site (Figure 4A). In contrast, in the MD-simulated structure of Lgt in complex with PG in the catalytic site, Ser198 turned toward the catalytic site and interacts with



**Figure 2.** MD simulation of the Lgt–PG complex. (A) Secondary structures represented in coral color exhibit the highest flexibility for the Lgt–PG complex. The two PG molecules are shown in green and yellow spheres. (B) RMSF analyses of the Lgt–PG and Lgt–PG–DAG are shown in red and black colors, respectively.

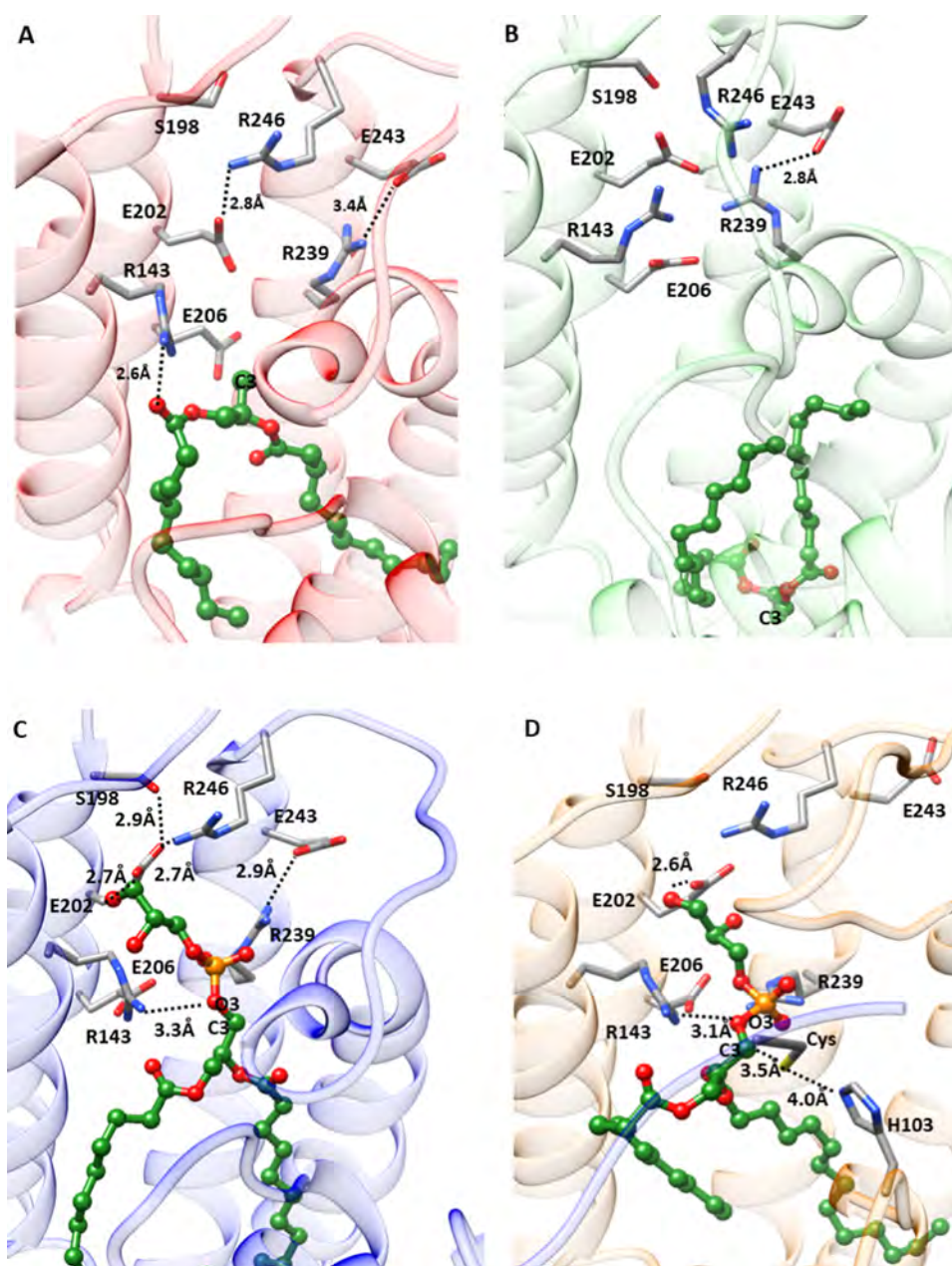


**Figure 3.** Opening of the front cleft from the MD-simulated structures. (A) Lgt–DAG (Lgt in complex with DAG and PG), (B) Lgt–PG (Lgt in complex with two PG), (C) Lgt–PG–lipobox (Lgt in complex with two PG and lipobox). The loop L6–7 at the front cleft is shown in ribbon representation; the lipobox is shown in the blue surface; DAG/PG in the MD-simulated structures are shown in CPK mode. Only the PG/DAG in the second site is shown.

Glu202, which, in turn, interacts with Arg246, stabilizing the glycerol head group of PG by a hydrogen bond (Figure 4C).

The C3–O ester bond needs to be activated to facilitate the subsequent nucleophilic attack by the deprotonated cysteine residue of the prelipoprotein. In the X-ray structure of Lgt bound with a DAG molecule, Arg143 was hydrogen-bonded with the O1 atom (ester oxygen of the *sn*-1 lipid chain) of DAG (Figure 4A). In contrast, in the MD-simulated structure of Lgt in complex with a PG molecule (a phosphate head was

added to DAG), Arg143 formed stable electrostatic interactions with the phosphate O3 of PG molecule in the second binding site (Figure 4C). In addition, Arg143 made stable ionic interactions with Glu206. Thus, the presence of the glycerol head group in the PG molecule helps to orientate the catalytically important residues Arg143 and Arg239 and therefore organize the active site for the diacylglycerol transfer. This is in accordance with the previous mutagenesis study



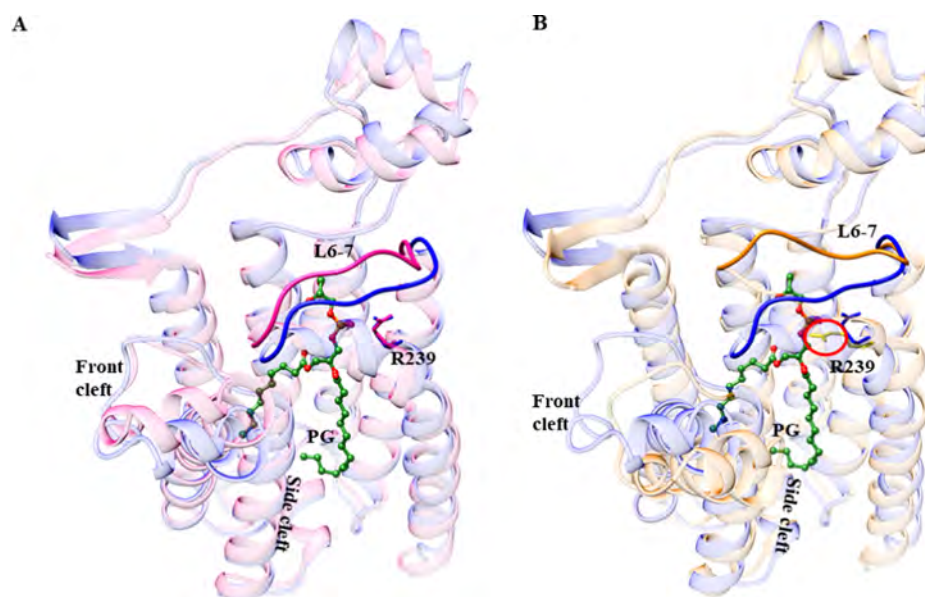
**Figure 4.** (A) X-ray structure of Lgt enzyme in complex with DAG molecule at the second site. The representative structures of Lgt complexes from the cluster analysis of MD simulation are shown in (B)–(D): (B) the Lgt–DAG complex (Lgt in complex with DAG and PG). (C) Lgt–PG (Lgt in complex with two PG molecules). (D) Lgt–PG–lipobox (Lgt in complex with two PG molecules and prelipoprotein signal peptide). The DAG/PG molecule is shown in stick and ball, and active site residues are shown in stick representations. The lipobox is shown in blue ribbon cartoon.

where the mutation of these residues had a dramatic effect on the enzyme activity of the Lgt enzyme.<sup>2</sup>

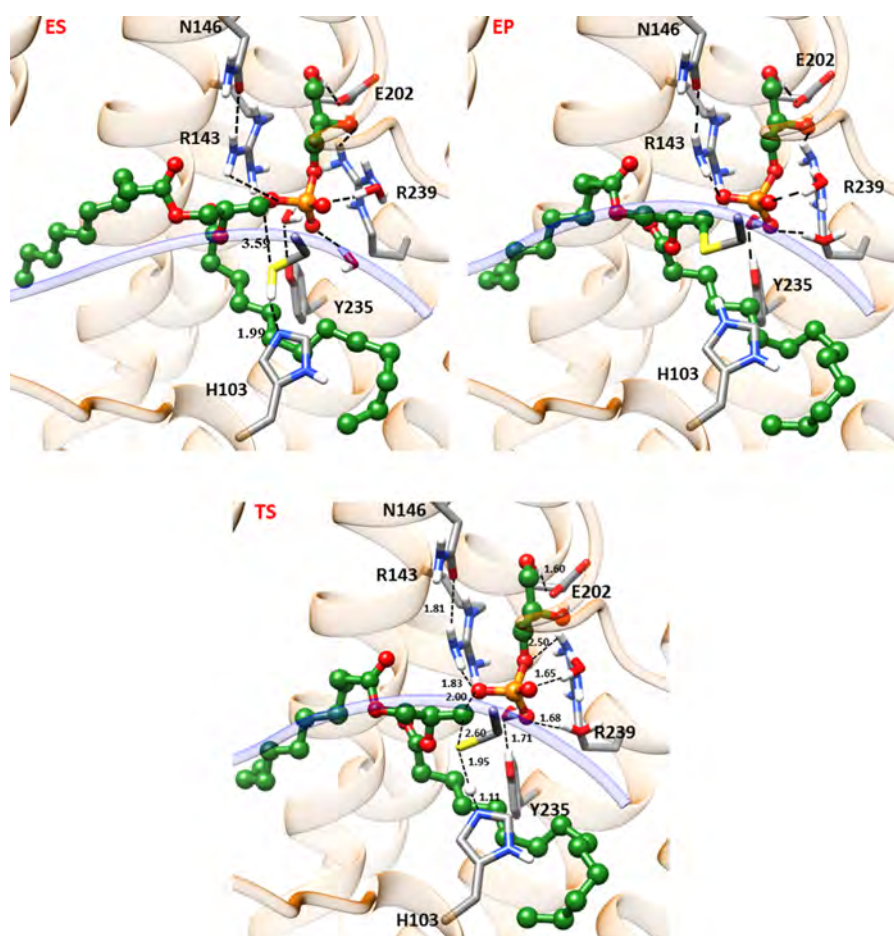
**Lgt–PG–Prelipoprotein (Enzyme–Substrate) Complex.** Aside from the head group of the PG molecule, the prelipoprotein signal peptide is also missing in the X-ray structure of the Lgt enzyme.<sup>2</sup> To understand the transacylation mechanism, the structure of Lgt in complex with both PG molecules and a lipobox signature peptide was constructed by docking the partial prelipoprotein signal peptide spanning the lipobox motif (GSTLLAGCSSN) to the representative structures of Lgt–PG complex obtained from 2.5  $\mu$ s MD simulations. The potential binding site of the prelipoprotein signal peptide (lipobox) in *E. coli* Lgt was proposed to be the

H<sup>103</sup>GGL motif near the side cleft of the enzyme.<sup>2</sup> The prelipoprotein peptide was docked in the front cleft and side cleft (Figure 1B), respectively, and the docked poses were further studied by MD simulations (Figure S7A,B).

The prelipoprotein peptide was first docked near the side cleft and subjected to MD simulations; however, with time evolution, the Cys residue moved on average 12.0 Å away from the PG in the second binding site and His103 made no interaction with the side chain of Cys (Figure S7C); hence, the docked prelipoprotein peptide near the side cleft of Lgt represented an unfavorable pose for the catalysis and was not studied further.



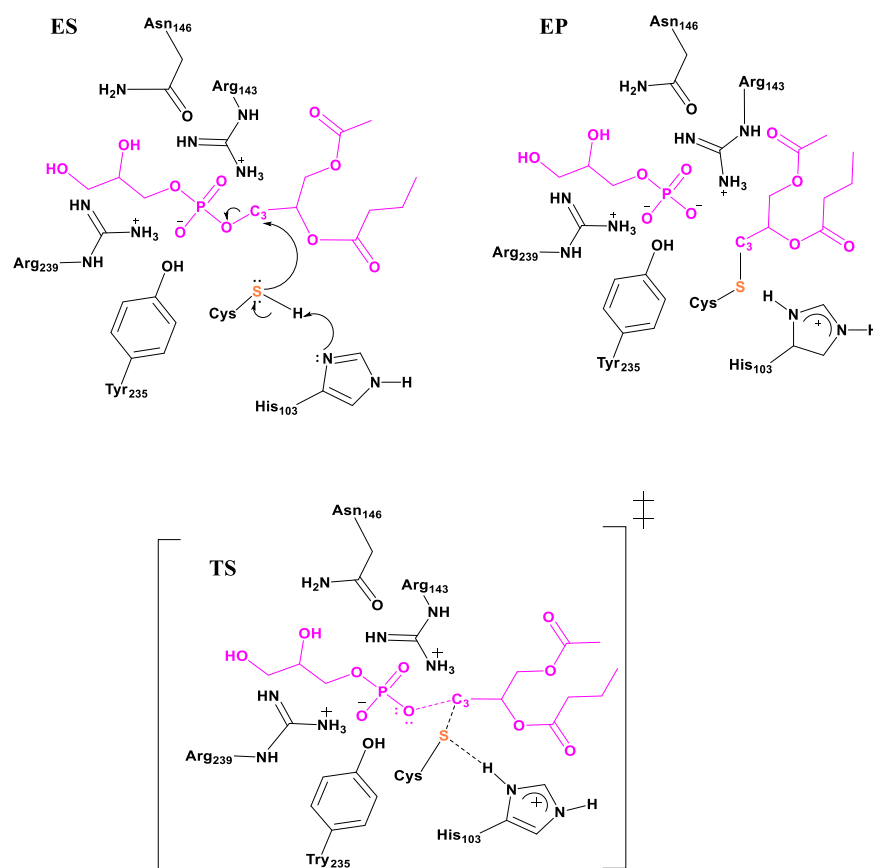
**Figure 5.** Gate opening mediated by Arg239 resulted in the opening of loop L6–7, which facilitates the substrate entry and product release. (A) Superimposing of the crystal structure with the MD structure of Lgt–PG, (B) superimposing of the MD structures of Lgt–PG and Lgt–PG–lipobox. The X-ray structure of Lgt–PG–DAG (PDB: 5AZC) is in pink, the MD structure of Lgt–PG is in blue, and the MD structure of Lgt–PG–lipobox is shown in orange. Arg239 of the MD structure of Lgt–PG–lipobox is shown in yellow and highlighted by a red circle.



**Figure 6.** QM/MM optimized structures of the Lgt enzyme. (ES) Lgt–reactant complex, (TS) Lgt–transition state complex, (EP) Lgt–product complex.

Then, the prelipoprotein peptide was positioned in the proximity of the front cleft (Table S2, Figure S7A), and a 150

ns MD simulation was run for the resulted Lgt–PG–lipobox complex (Figure S8). Similar to the simulated structure of the



**Figure 7.** Schematic representation of the reaction mechanism of diacylglycerol transfer from PG to a preprolipoprotein signal peptide in Lgt.

Lgt–PG complex (Figure 4C), the catalytic residue Arg143 formed a hydrogen bond with Glu206 and the O3 oxygen atom of the PG molecule (Figure 4D). In addition, Glu202 stabilized the head group of the PG molecule by H-bond interaction with the glycerol hydroxyl. Notably, in the presence of the preprolipoprotein peptide, Glu243 turned away from Arg239 so that the ionic interaction between Glu243 and Arg239 observed in the absence of the lipobox was lost; in the meantime, Arg239 turned toward the phosphate group of PG molecule, forming an electrostatic interaction with it (Figures 4D and S9).

Hence, we propose that Arg239 located on HM6 serves as a gate (Figure 5). In the absence of the lipobox, the gate is open, whereas the gate is closed with lipobox binding, and the L6–7 loop exhibits an open conformation, which may facilitate the release of the product sideways, as well as the entry of another PG molecule in the first binding site. This is in accordance with the previous hypothesis that PG sliding may induce the opening of the gate formed around the loop region L6–7.<sup>2</sup> The protein–lipid interaction still represents high complexity<sup>37</sup> although the entry and exit of membrane-embedded ligands to the protein have been extensively studied.<sup>38</sup> Actually, the opening of the binding channel triggered by TM helix tilting has been reported in other lipid-processing enzymes<sup>39–41</sup> and was also observed during substrate binding in some GPCR proteins such as sphingosine-1-phosphate receptor 1<sup>42</sup> and rhodopsin.<sup>43,44</sup>

Sankaran et al., 1997 performed site-directed mutagenesis on all of the histidine present in the Lgt enzyme and found that His103 was essential for the enzyme activity.<sup>45</sup> Mao et al., 2016 also identified His103 as a critical residue for the Lgt enzyme

activity.<sup>2</sup> The MD-simulated structure of Lgt–PG–lipobox showed the average distances between the cysteine thiol group of the preprolipoprotein signal peptide in relation to the crucial C3 carbon of the ester bond of the PG molecule and that between His103 of the Lgt enzyme were 3.5 Å ( $\pm 0.3$ ) and 4.0 Å ( $\pm 0.5$ ), respectively (Figures 4D and S9). This indicates that His103 may function as a catalytic base to abstract a proton from the lipobox cysteine, which subsequently undergoes a nucleophilic attack on the C3 atom of the PG substrate.

It should be noted that although His103 is well conserved in *Proteobacteria* Lgt sequences, it is not strictly conserved in all Lgt enzymes. Alternative residues such as tyrosine, phenylalanine, and tryptophan replace the histidine residue based on the multiple sequence alignment of Lgt proteins from a range of bacterial phyla (Figure S10). Further modeling and structural studies are therefore needed to determine the deprotonation mechanism of cysteine in other Lgt enzymes where the histidine residue is absent at position 103.

**Reaction Pathway of Lgt Revealed by QM/MM.** Hybrid QM/MM method can be used to study the reaction in enzyme systems.<sup>46</sup> To elucidate the reaction mechanism of the nucleophilic attack of the C3–O ester bond in the PG molecule by cysteine of the docked preprolipoprotein, QM/MM studies were performed on the Lgt–PG–lipobox complex. The important residues identified from our MD simulations and previous mutagenesis studies of Lgt<sup>2,45</sup> were included in the QM region. These include His103, Arg143, Asn146, Tyr235, and Arg239, the cysteine residue of the lipobox, and three water molecules identified in the vicinity of the phosphate head group in the active site of Lgt.

In the QM/MM optimized reactant complex, the side chain of His103 formed a hydrogen bond with the thiol group of cysteine with a distance of 1.99 Å (Figure 6). The distance between the sulfur atom of the cysteine and the carbon of the C3–O ester bond is 3.59 Å. In the QM/MM optimized structure, the C3–O ester bond of PG formed an H-bond with Arg143, which was further involved in the H-bond interaction with Asn146. The phosphate group of the PG is stabilized by the hydrogen bonds with three water molecules. Tyr235 makes water-assisted hydrogen bonds with the oxygen of the phosphate group. Glu202 and Arg239 stabilize the glycerol oxygen atoms of the PG by hydrogen-bonding interactions.

A potential energy surface (PES) scan was conducted by scanning the distance between the sulfur atom of the cysteine and the C3 carbon of the C3–O ester bond. In the QM/MM optimized TS complex, the distance of the epsilon nitrogen to the proton of cysteine was 1.11 Å and the distance between the proton and the sulfur atom was 1.95 Å (Figure 6). The distance between the C3 carbon atom of the C3–O ester bond and the phosphate oxygen became 2.00 Å. The deprotonated sulfur atom of cysteine was 2.60 Å away from the C3. This distance is in good agreement with the transition state associated with the acylation reaction in bile salt hydrolase enzyme, where the nucleophile cysteine anion was 2.60 Å away from the carbonyl carbon of the amide bond of tauroconjugated bile acid.<sup>47</sup> The developed negative charge due to the cleavage of the C3–O on the phosphate group was stabilized by the electrostatic interaction with Arg143. The side chain of Asn146 approached Arg143, maintaining a stable hydrogen bond with it.

The QM/MM calculations show that the diacylglyceryl transfer from PG to the prolipoprotein peptide substrate proceeds via an SN<sub>2</sub>-like concerted mechanism (Figure 7). His103 acts as a catalytic base to abstract the proton from the cysteine residue of the prolipoprotein peptide. The deprotonated cysteine then undergoes nucleophilic attack on the C3 carbon of the C3–O ester in PG to form a diacylglyceryl–prolipoprotein. The estimated activation energy of the reaction was 18.6 kcal mol<sup>-1</sup> (Figure 8). It is in good agreement with



**Figure 8.** QM/MM reaction profile of the Lgt enzyme computed using the B3LYP functional with 6-31+G(d,p) basis set. The values in parentheses and square bracket are the single-point QM/MM calculations performed with the 6-311++G(3df,3pd) basis set using B3LYP and  $\omega$ B97XD functional, respectively.

the experimental activation energy barrier of Lgt enzyme ( $V_{\max} = 25.5 \text{ nmol mg}^{-1} \text{ min}^{-1}$ ,<sup>45</sup>  $\sim 21.0 \text{ kcal mol}^{-1}$ ), since the QM/MM DFT calculations using the B3LYP functional are known to underestimate the organic reaction barrier with a mean absolute error of  $\sim 3 \text{ kcal mol}^{-1}$ .<sup>48</sup> The energy of the product formed was  $-13.0 \text{ kcal mol}^{-1}$  lower than that of the ES,

indicating that the overall reaction is exothermic in nature. The transition state was located and validated by a unique imaginary frequency of  $-442.3 \text{ cm}^{-1}$ , and intrinsic reaction coordinate (IRC) scan calculations were performed on the transition state, which showed that it connects the reactant and product (Figure S11). Potential surface scan starting from an alternative starting structure selected from MD-simulated trajectory and gave similar reaction profiles (Figures S12 and S13).

We also examined the possibility of the stepwise reaction mechanism by studying the proton transfer from the cysteine thiol group of the lipobox to His103 of Lgt. The QM/MM potential energy scan in the quest for a possible ionic pair showed that the energy kept increasing (Figure S14). Optimization of the structure closest to an ionic pair on the reaction profile leads to the neutral form, i.e., the proton moved back to cysteine. In addition, no transition state could be located corresponding to the proton transfer from cysteine to His103, indicating that the reaction does not go through an ionic form and therefore the stepwise mechanism is not viable.

In summary, here, we provide a catalytically active model of the *E. coli* Lgt enzyme in complex with prelipoprotein and PG substrates; our MD simulations of Lgt–PG–lipobox show that the product is indeed released sideways relative to the lipid bilayer, in agreement with the previous hypothesis proposed by Mao et al. that diacylglyceryl–prolipoprotein is released sideways relative to the lipid-bilayer-based;<sup>2</sup> in addition, our QM/MM studies demonstrate that His103 plays a crucial role in the catalytic process, acting as a catalytic base to trigger the reaction.

## CONCLUSIONS

Lgt is an important bacterial enzyme in post-translocational modification of bacterial lipoproteins, and understanding of the plasticity of the active site of Lgt should provide valuable insight into the design of novel antibiotic drugs. Here, the Lgt structure in complex with a PG substrate and prelipoprotein signature peptide was modeled using MD simulations. We found a catalytically active conformation of the enzyme attributed by notable conformational changes of loop L6–7. In addition, we validated the importance of Arg143, Arg239, and Glu202 in the stabilization of the glycerol-1-phosphate head group of the PG substrate. The conserved Arg239 functions as a gate to regulate the opening of the loop L6–7 so as to facilitate the entry of the PG substrate and the release of the modified lipoprotein product.

The QM/MM study of the reaction mechanism of *E. coli* Lgt enzyme showed that His103 acts as a catalytic base, facilitating the thiolate anion (conserved cysteine of the prelipoprotein signal peptide) to attack the C3 carbon of PG with an activation barrier of  $18.6 \text{ kcal mol}^{-1}$ . After lipid modification, diacylglyceryl–prolipoprotein then exits the Lgt enzyme sideways relative to the lipid bilayer, allowing it to undergo further modification catalyzed by the LspA and Lnt enzymes.

In summary, here, we present a catalytically active model of Lgt enzyme in complex with the prolipoprotein and PG substrate and elucidated the mechanism of diacylglyceryl transfer from the PG substrate to the prolipoprotein catalyzed by Lgt. The information obtained on the binding mode of the prolipoprotein in the Lgt enzyme along with the reaction mechanism of the diacylglyceryl transfer would provide us novel insight on the structure-based design of broad-spectrum

antimicrobial therapies targeting the enzymes involved in the post-translocational modification of lipoproteins.

## ■ ASSOCIATED CONTENT

### ■ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpccb.9b04227.

$pK_a$  values of the titratable residues of Lgt, the docking scores of the lipobox in Lgt, the interactions of the PG molecule with first binding site residues, the MD-simulated poses of Lgt–PG–lipobox, the sequence alignment of Lgt proteins, the QM/MM reaction profile from an alternative starting structure, and the QM/MM reaction profile in the quest for stepwise mechanism (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: m.huang@qub.ac.uk.

### ORCID

Thomas S. Moody: 0000-0002-8266-0269

Meilan Huang: 0000-0001-9063-8056

### Author Contributions

<sup>†</sup>W.S. and M.B. contributed equally to this paper.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors acknowledge the financial support from INVEST NI Research and Development Programme, part-financed by the European Regional Development Fund under the Investment for Growth and Jobs programme 2014–2020. The authors are grateful for the computing resources from QUB High Performance Computing Centre.

## ■ REFERENCES

- (1) Sutcliffe, I. C.; Harrington, D. J.; Hutchings, M. I. A Phylum Level Analysis Reveals Lipoprotein Biosynthesis to Be a Fundamental Property of Bacteria. *Protein Cell* **2012**, *3*, 163–170.
- (2) Mao, G.; Zhao, Y.; Kang, X.; Li, Z.; Zhang, Y.; Wang, X.; Sun, F.; Sankaran, K.; Zhang, X. C. Crystal Structure of *E. coli* Lipoprotein Diacylglycerol Transferase. *Nat. Commun.* **2016**, *7*, No. 10198.
- (3) Babu, M. M.; Priya, M. L.; Selvan, A. T.; Madera, M.; Gough, J.; Aravind, L.; Sankaran, K. A Database of Bacterial Lipoproteins (DOLOP) with Functional Assignments to Predicted Lipoproteins. *J. Bacteriol.* **2006**, *188*, 2761–2773.
- (4) Pailler, J.; Aucher, W.; Pires, M.; Buddelmeijer, N. Phosphatidylglycerol: Prolipoprotein Diacylglycerol Transferase (Lgt) of *Escherichia coli* Has Seven Transmembrane Segments, and Its Essential Residues Are Embedded in the Membrane. *J. Bacteriol.* **2012**, *194*, 2142–2151.
- (5) Sankaran, K.; Wu, H. C. Lipid Modification of Bacterial Prolipoprotein. Transfer of Diacylglycerol Moiety from Phosphatidylglycerol. *J. Biol. Chem.* **1994**, *269*, 19701–19706.
- (6) Vogeley, L.; El Arnaout, T.; Bailey, J.; Stansfeld, P. J.; Boland, C.; Caffrey, M. Structural Basis of Lipoprotein Signal Peptidase II Action and Inhibition by the Antibiotic Globomycin. *Science* **2016**, *351*, 876–880.
- (7) Tjalsma, H.; Zanen, G.; Venema, G.; Bron, S.; van Dijk, J. M. The Potential Active Site of the Lipoprotein-Specific (Type II) Signal Peptidase of *Bacillus Subtilis*. *J. Biol. Chem.* **1999**, *274*, 28191–28197.
- (8) Noland, C. L.; Kattke, M. D.; Diao, J.; Gloor, S. L.; Pantua, H.; Reichelt, M.; Katakam, A. K.; Yan, D.; Kang, J.; Zilberley, I.; et al. Structural Insights into Lipoprotein N-Acylation by *Escherichia coli*

Apolipoprotein N-Acyltransferase. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, E6044–E6053.

(9) Wiktor, M.; Weichert, D.; Howe, N.; Huang, C.-Y.; Olieric, V.; Boland, C.; Bailey, J.; Vogeley, L.; Stansfeld, P. J.; Buddelmeijer, N.; et al. Structural Insights into the Mechanism of the Membrane Integral N-Acyltransferase Step in Bacterial Lipoprotein Synthesis. *Nat. Commun.* **2017**, *8*, No. 15952.

(10) Xia, J.; Feng, B.; Wen, G.; Xue, W.; Ma, G.; Zhang, H.; Wu, S. Bacterial Lipoprotein Biosynthetic Pathway as a Potential Target for Structure-Based Design of Antibacterial Agents. *Curr. Med. Chem.* **2018**, DOI: 10.2174/0929867325666181008143411.

(11) Dennington, R.; Keith, T.; Millam, J. *GaussView*, version 5; Shawnee Mission, 2015.

(12) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H<sup>++</sup>: A Server for Estimating PKas and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* **2005**, *33*, W368–W371.

(13) Dickson, C. J.; Madej, B. D.; Skjevik, ÅA.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.* **2014**, *10*, 865–879.

(14) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a Package of Computer Programs for Applying Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Comput. Phys. Commun.* **1995**, *91*, 1–41.

(15) London, N.; Raveh, B.; Cohen, E.; Fathi, G.; Schueler-Furman, O. Rosetta FlexPepDock Web Server–High Resolution Modeling of Peptide-Protein Interactions. *Nucleic Acids Res.* **2011**, *39*, W249–W253.

(16) Raveh, B.; London, N.; Schueler-Furman, O. Sub-Angstrom Modeling of Complexes between Flexible Peptides and Globular Proteins. *Proteins* **2010**, *78*, 2029–2040.

(17) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878–3888.

(18) Lomize, M. A.; Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. OPM: Orientations of Proteins in Membranes Database. *Bioinformatics* **2006**, *22*, 623–625.

(19) Jo, S.; Lim, J. B.; Klauda, J. B.; Im, W. CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophys. J.* **2009**, *97*, 50–58.

(20) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(21) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

(22) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N-log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(23) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(24) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(25) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.

(26) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 27–28.

(27) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—a Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.

- (28) Tao, P.; Schlegel, H. B. A Toolkit to Assist ONIOM Calculations. *J. Comput. Chem.* **2010**, *31*, 2363–2369.
- (29) Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. Combining Quantum Mechanics Methods with Molecular Mechanics Methods in ONIOM. *J. Chem. Theory Comput.* **2006**, *2*, 815–826.
- (30) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels–Alder Reactions and Pt(P(t-Bu)<sub>3</sub>)<sub>2</sub> + H<sub>2</sub> Oxidative Addition. *J. Phys. Chem. B* **1996**, *100*, 19357–19363.
- (31) Frisch, M.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A. et al. *Gaussian 09*, revision D. 01; Gaussian, Inc: Wallingford CT, 2009.
- (32) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (33) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (34) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785–789.
- (35) Knapp, B.; Ospina, L.; Deane, C. M. Avoiding False Positive Conclusions in Molecular Simulation: The Importance of Replicas. *J. Chem. Theory Comput.* **2018**, *14*, 6127–6138.
- (36) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
- (37) Corradi, V.; Sejdiu, B. I.; Mesa-Galloso, H.; Abdizadeh, H.; Noskov, S. Y.; Marrink, S. J.; Tieleman, D. P. Emerging Diversity in Lipid–Protein Interactions. *Chem. Rev.* **2019**, 5775–5848.
- (38) Marrink, S. J.; Corradi, V.; Souza, P. C. T.; Ingólfsson, H. I.; Tieleman, D. P.; Sansom, M. S. P. Computational Modeling of Realistic Cell Membranes. *Chem. Rev.* **2019**, 6184–6226.
- (39) Zhang, X. C.; Liu, Z.; Li, J. From Membrane Tension to Channel Gating: A Principal Energy Transfer Mechanism for Mechanosensitive Channels. *Protein Sci.* **2016**, *25*, 1954–1964.
- (40) Deplazes, E.; Louhivuori, M.; Jayatilaka, D.; Marrink, S. J.; Corry, B. Structural Investigation of MscL Gating Using Experimental Data and Coarse Grained MD Simulations. *PLoS Comput. Biol.* **2012**, *8*, No. e1002683.
- (41) Louhivuori, M.; Risselada, H. J.; van der Giessen, E.; Marrink, S. J. Release of Content through Mechano-Sensitive Gates in Pressurized Liposomes. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 19856–19860.
- (42) Stanley, N.; Pardo, L.; Fabritiis, G. D. The Pathway of Ligand Entry from the Membrane Bilayer to a Lipid G Protein-Coupled Receptor. *Sci. Rep.* **2016**, *6*, No. 22639.
- (43) Park, J. H.; Scheerer, P.; Hofmann, K. P.; Choe, H.-W.; Ernst, O. P. Crystal Structure of the Ligand-Free G-Protein-Coupled Receptor Opsin. *Nature* **2008**, *454*, 183–187.
- (44) Hildebrand, P. W.; Scheerer, P.; Park, J. H.; Choe, H.-W.; Piechnick, R.; Ernst, O. P.; Hofmann, K. P.; Heck, M. A Ligand Channel through the G Protein Coupled Receptor Opsin. *PLoS One* **2009**, *4*, No. e4382.
- (45) Sankaran, K.; Gan, K.; Rash, B.; Qi, H. Y.; Wu, H. C.; Rick, P. D. Roles of Histidine-103 and Tyrosine-235 in the Function of the Prolipoprotein Diacylglycerol Transferase of *Escherichia coli*. *J. Bacteriol.* **1997**, *179*, 2944–2948.
- (46) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.
- (47) Lodola, A.; Branduardi, D.; Vivo, M. D.; Capoferri, L.; Mor, M.; Piomelli, D.; Cavalli, A. A Catalytic Mechanism for Cysteine N-Terminal Nucleophile Hydrolases, as Revealed by Free Energy Simulations. *PLoS One* **2012**, *7*, No. e32397.
- (48) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.