A Computational and Modeling Study of the Reaction Mechanism of *Staphylococcus aureus* Monoglycosyltransferase Reveals New Insights on the GT51 Family of Enzymes

Kenneth Goossens, Rui PP Neves, Pedro A Fernandes, and Hans De Winter*

Cite This: https://dx.doi.org/10.1021/acs.jcim.0c00377

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: Bacterial glycosyltransferases of the GT51 family are key enzymes in bacterial cell wall synthesis. Inhibiting cell wall synthesis is a very effective approach for development of antibiotics, as this can lead to either bacteriostatic or bactericidal effects. Even though the existence of this family has been known for over 50 years, only one potent inhibitor exists, which is an analog of the lipid IV product and derived from a natural product. Drug development focused on bacterial transglycosylase has been hampered due to little being know about its structure and reaction mechanism. In this study, *Staphylococcus aureus* monoglycosyltransferase was investigated at an atomistic level using computational methods. Classical molecular dynamics simulations were used to reveal information about the large-scale dynamics of the enzyme–substrate complex and the importance of magnesium in structure and function of the protein, while mixed mode quantum mechanics/molecular mechanics calculations unveiled a novel hypothesis for the reaction mechanism. From these results, we present a new model for the binding mode of lipid II and the reaction mechanism of the GT51 glycosyltransferases. A metal-bound hydroxide catalyzed reaction mechanism yields an estimated free energy barrier of 16.1 ± 1.0 kcal/mol, which is in line with experimental values. The importance of divalent cations is also further discussed. These findings could significantly aid targeted drug design, particularly the efficient development of transition state analogues as potential inhibitors for the GT51 glycosyltransferases.
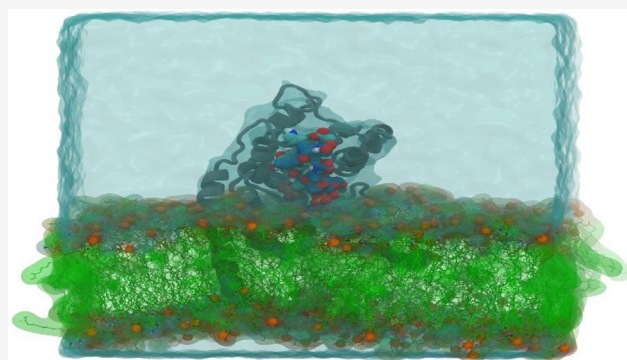
INTRODUCTION

Staphylococcus aureus is a Gram-positive bacterium and has been considered an important cause of human pathologies for a long time, being the most commonly isolated human bacterial pathogen. The rapid rise in resistance of *S. aureus* to antibiotics has led it to become an increasingly hot topic and a problematic pathogen to treat. Efforts are being made toward developing new antibiotics to treat methicillin-resistant *S. aureus* infections, but the bacterium has proven to be not only resistant to β-lactam antibiotics but also capable of rapidly developing resistance mechanisms against most antibiotics that are used to treat them.[1−4] Over the years, various attempts have been made toward identifying new antibiotics, but finding potent and druggable compounds has proven to be difficult.[4−6]

A relatively novel approach to finding new antibiotics is the inhibition of the bacterial glycosyltransferases (GTases) of family 51 (GT51). These membrane proteins are inverting GTases that link together lipid II substrates to create polysaccharide chains, which are the key component of the bacterial cell wall. While the GT51 family is classified under GTases, it is structurally more similar to the lysozyme family of enzymes, although the similarity is still limited.[7] The structural ambiguity of this family, combined with the lack of other

enzymes that are structurally similar, make it a difficult target for drug development. Most enzymes in the GT51 family are bifunctional membrane proteins containing both a domain that catalyzes acyltransfer and a domain that catalyzes glycosyltransfer, which is anchored to the lipid bilayer by a transmembrane helix. One exception is *S. aureus* monoglycosyltransferase (SAMGT), which only catalyzes glycosyltransfer. This 269 amino acid long membrane protein can be defined by five motifs that are conserved among other GT51 enzymes and a flexible flap domain, which is not fully resolved in all but one of the available crystal structures (Figure 1).[8]

The catalytic activity of SAMGT takes place when a lipid II (see Figure 2A for its structure) molecule or a polysaccharide strand binds to the donor site and a lipid II molecule binds to the acceptor binding site. Apart from its function, very little is known about the mechanistic details of GT51 activity. The
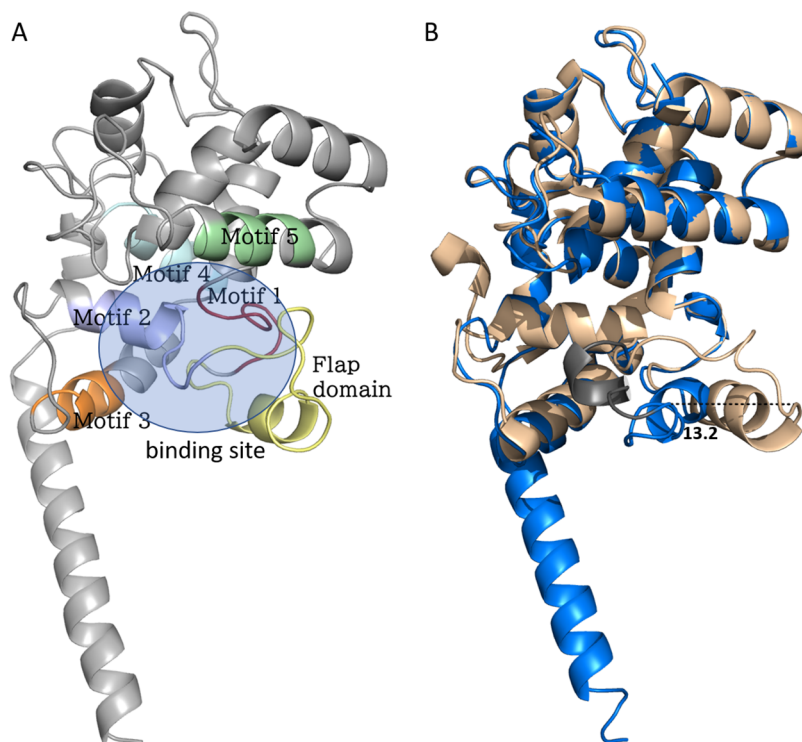
Figure 1. Structure of *S. aureus* MGT. (A) Conserved motifs and the flap domain are displayed in different colors. (B) Superposition of two different crystal structures (PDB ID 3VMT, shown in blue cartoon; PDB ID 6FTB, shown in wheat cartoon) shows significant conformational flexibility of the flap domain. The modeled part of the flap domain is colored gray. The Thr_{122} Cα distance between the structures is 13.2 Å (dotted line).
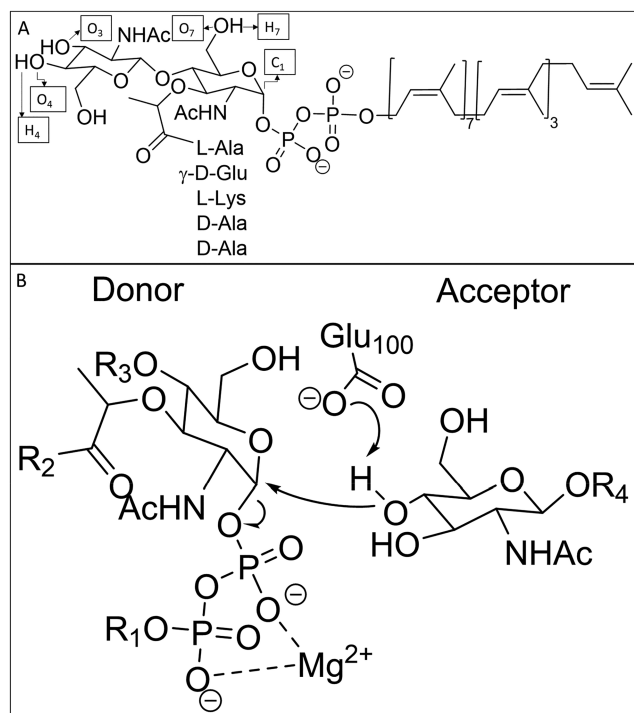

Figure 2. (A) Structure of lipid II. Reference names for key atoms are represented in the boxes. (B) Currently proposed reaction mechanism for enzymatic catalysis of lipid II linkage. A magnesium ion in the active site is stabilizing the diphosphate of the lipid II donor molecule.

currently proposed reaction mechanism is a glutamate-assisted S_N2 nucleophilic attack of the lipid II acceptor O_4 oxygen

(O_4a) onto the lipid II or polysaccharide donor C_1 carbon (C_1d) (Figure 2B).[9−12] A $\beta(1−4)$-linked polysaccharide is formed this way, and the newly elongated chain is subsequently shuffled into the donor site. The glutamate residue is thought to play a catalytic role in analogy to the mechanism proposed for many GTases of the GT-A and GT-B families[13] but also due to its conservation within the GT51 family,[14] the loss of activity upon mutation to glutamine,[14−18] and titration experiments suggesting a basic catalytic residue.[14,19] However, its direct implication in the reaction mechanism has never been shown.

Existing crystal structures support two possible binding modes. The protein either has two binding pockets separated by a (partially unresolved) flexible flap domain (Figure 1B, blue cartoon)[20] or has the flap domain fully resolved and shifted away from motif 3, exposing only a single, large binding pocket (Figure 1B, wheat cartoon).[8] The first model is interesting as crystal structures exist with a lipid IV analog in the donor pocket or with a lipid II analog in the acceptor pocket. Furthermore, the "second" pocket is also suited for substrate binding, as it has a patch of positively charged residues that could bind and stabilize the diphosphate moiety of lipid II. However, the model has one mechanistic shortcoming that cannot be ignored. Lipid II binding in two separated pockets would significantly hinder shuffling of the formed product into the donor site, as the hydrophobic tail of the formed lipid IV product is obstructed by the flap domain, which separates the two pockets. Since the flap domain lies at the bilayer interface, its flexibility along the z-axis is limited. The shuffling mechanism has been confirmed in one study, where the length of the polysaccharide products was measured

using thin-layer chromatography,[16] thus crediting a model with one large pocket for SAMGT.

Although various drug discovery efforts have been made targeting this enzyme, the only well-defined GT inhibitors currently available belong to the moenomycin family.[21] While these lipid IV analogues are potent inhibitors, their poor pharmacokinetic properties hinder their potential as pharmaceutical compounds.[9,21] *In silico* drug discovery is hindered by the lack of knowledge about which of the two proposed enzyme conformations is the one that is suitable for substrate binding and the lack of a description of the reaction mechanism and kinetics of the protein. This is exacerbated by the relatively low resolution and quality of the existing crystal structures, which is recurring in membrane proteins, a problem that will likely not be solved in the near future. In the face of all these uncertainties, we have tried to draw a clearer image of substrate binding and catalysis for this enzyme. In this work, we have combined classical molecular dynamics (cMD), free energy calculations, and hybrid quantum mechanics/ molecular mechanics (QM/MM) calculations to create a new model for the binding mode and reaction mechanism of SAMGT that is consistent with experimental data. Even though this enzyme has not given up all its mysteries, we believe that this study helps to clarify many aspects of its substrate binding and catalysis.

■ RESULTS

In order to investigate the most likely binding conformation of the SAMGT enzyme and the lipid II substrates and the role of magnesium ions, multiple microsecond-time scale cMD simulations were performed. These are summarized in Table 1. A more detailed description of the simulated systems is given in the Methods section. Reaction coordinates will be referred to in the form "d(X−Y)", with "d" signifying a distance-based reaction coordinate and X and Y identifying the involved atoms. Names of lipid II atoms can be found in Figure 2A. Atom suffixes "d" and "a" stand for donor lipid II and acceptor lipid II, respectively.

Open versus Closed Binding Conformation of SAMGT. Large-time scale cMD simulations were used to investigate the most likely conformation of the flap domain when SAMGT gets bound to its substrates. In particular, both the closed and open conformation were simulated and the critical interatomic distances required for catalysis were monitored (Figure 2B). A view of the starting conformations can be found in Figure 3.

To evaluate the first hypothesis, in which the donor and acceptor lipid II are separated by the flexible flap domain (the closed conformation), eight cMD simulations with an aggregated simulation time of up to 4.2 μs were set up (system 1 in Table 1; Figure 1B, blue cartoon). Analysis of the distance between the C_1d and O_4a atoms of the respective lipid II molecules as a function of time showed that this distance never got shorter than 8 Å, with an average value of 11.7 ± 1.1 Å over the course of the entire 4.2 μs long simulation (Figure S1). The fact that the closed conformation is shown to be very unsuitable for catalysis, even when explored though 4.2 μs MD simulations, led us to propose that the closed SAMGT conformation prevents a favorable orientation between the two lipid II molecules and the subsequent formation of the covalent linkage. Preliminary steered MD simulations also suggested that the flap domain prevents closer contacts between donor and acceptor lipid II molecules (data not shown). Therefore, it seems that the flap domain prevents close contacts between the donor and acceptor lipid II molecules in the closed conformation. As a consequence, we dropped the first hypothesis and focused on the second putative binding mode, in which both substrates occupy one large pocket with the flap domain in an open conformation (system 2 in Table 1; Figure 1B, wheat cartoon). A 20.8 μs long simulation on this open conformation system resulted in a mean C_1d−O_4a distance of 8.8 ± 3.3 Å, which is much shorter than the mean distance in the closed conformation. Close contacts of less than 5 Å between C_1d and O_4a occur in 9.4% of the simulation frames and in 21.7% of the simulation frames when a cutoff of 6 Å is taken (Figure S2).

Table 1. Summary of the Classical MD Simulations That Were Performed

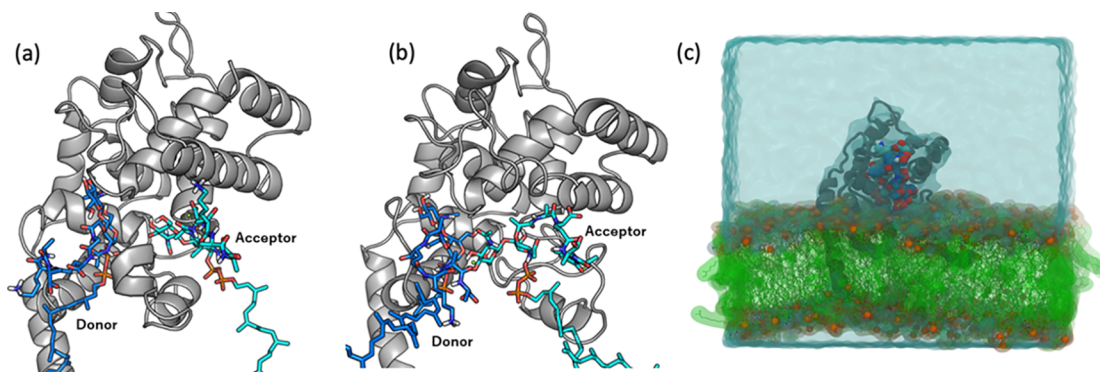| system	number of simulations	aggregated time (μs)	force field	description
system 1	8	4.2	Gromos54A7	closed flap domain, full system
system 2	20	20.8	Gromos54A7	open flap domain, full system

Figure 3. Representation of the SAMGT system. SAMGT is shown in gray cartoon. Donor and acceptor lipid II are shown in dark blue and cyan, respectively. Magnesium is shown as green spheres. (A) Closed conformation. (B) Open conformation. (C) Full system in the open conformation. Water is illustrated by the transparent blue surface. Bilayer phospholipids are shown as green sticks with a transparent green surface, with the phosphorus atoms of the phospholipid head groups shown as orange spheres.
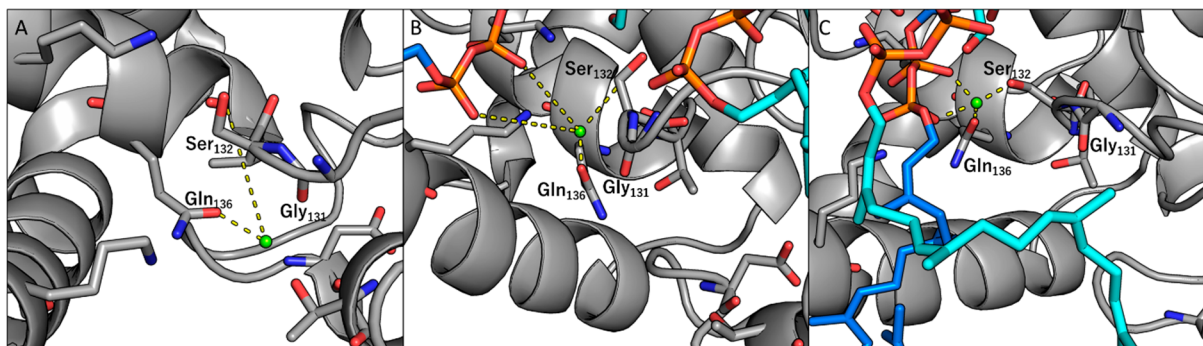
Figure 4. Initial position of the magnesium ion in the active site. The yellow dashed lines mark the contacts present in the final coordination sphere. (A) Magnesium position in the crystal structure (PDB ID 3VMT) (closed conformation). (B) Magnesium position after initial modeling (open conformation). (C) Stable magnesium coordination sphere after simulation (open conformation).

Even though very large time scale rearrangements of the flap domain to bring the closed conformation to a productive position cannot be absolutely excluded, they are very unlikely, as during the very long MD simulations the flap domain already moves so much that the C_1d-O_4a distance fluctuates more than 10 Å. This means that large-scale motions of the flap domain are explored and observed during the MD simulations, and these never brought the enzyme to a productive conformation in the closed conformation, but it did, and quite frequently, in the open conformation.

This distance criterion indicated that the second binding mode was by far the most likely one and was taken as the model system to investigate subsequently the position of the active site Mg^{2+} ion, the different lipid II binding conformations, and the catalytic mechanism of SAMGT.

Location of the Active Site Mg^{2+} Ion. Magnesium ions are not resolved in all of the available SAMGT crystal structures. Only two crystal structures exhibit a Mg^{2+} ion near the donor lipid II binding site (PDB ID 3VMT and 3VMQ; Figure 2B), while in only one crystal structure (PDB ID 3VMT) a second Mg^{2+} ion, located in a peripheral domain at approximately 13 Å from the attacking O_4 atom of the lipid II acceptor molecule, has been reported. Even though the other available crystal structures do not report any resolved Mg^{2+} ions in their electron densities, it has been shown that divalent metals are imperative for GTase function and more specifically for the function of SAMGT.[14] While the importance of the divalent metals appears to be dependent on the bacterial strain and the specific isozyme, they have a moderate to strong influence on activity, with the activity of SAMGT being almost completely abolished in their absence. Since the inclusion of the Mg^{2+} ions at the active site is imperative for the QM/MM study in this work, classical MD simulations were performed to investigate the most likely binding motif of the active site Mg^{2+} ion. The initial position of the active site Mg^{2+} ion was based on the position reported in literature[20] but was slightly adjusted in order to prevent clashes with the donor lipid II molecule and residues in close proximity (Figure 4).

These classical MD simulations of system 2 (SAMGT in the open conformation) revealed a very stable binding motif for the active site Mg^{2+} ion, but one that is different compared to the binding motif as observed in the known crystal structures. Early on in the cMD simulations, contacts between the Mg^{2+} ion near the donor lipid II binding site and the diphosphate moiety were formed spontaneously, resulting in a stabilized coordination sphere involving two water molecules, residues

Gln_{136} and Ser_{132}, and two salt bridges involving the diphosphate moiety of the donor lipid II molecule. This coordination sphere was already formed after a few nanoseconds following the start of the dynamics and remained extremely stable during the remainder of the 20.8 μs long simulation. We stress that the two Mg-containing crystal structures of SAMGT do not contain lipid II as substrate or any analog diphosphate moiety. In this circumstance, it is not expected for the Mg^{2+} ion to be present. The observed coordination sphere is in line with our expectations, as numerous cases are described in which diphosphate and triphosphate binding and release in enzymes are stabilized by Mg^{2+}, suggesting that Mg^{2+} plays a key role in stabilizing the diphosphate moiety.[22−24] In fact, it is widely common in enzymes that bind di- and triphosphates have divalent cations (mostly Mg^{2+} ions) coordinated to the phosphates. Broad examples include the families of polymerases, reverse transcriptases, and RNases, among many others. Based on large precedent, it is expected to find a Mg^{2+} ion coordinating the phosphate group. Therefore, we opted to model the position of the Mg^{2+} ion as predicted by the cMD simulations in the subsequent QM/MM studies on the reaction mechanism. Despite the modeling of the Mg^{2+} ion at this position being solid and precedented, we cannot exclude that other positions for the ion might exist. However, it is improbable that placing the Mg^{2+} ion in other coordination shells not involving the phosphates can have the strong effect in catalysis that is experimentally observed.[14]

Binding Conformations of the lipid II Substrates. Lipid II is a large substrate and has much conformational freedom due to the large number of rotational degrees of freedom. The lipid II binding pocket is very solvent exposed, meaning that the relative orientation of the lipid II substrates can also vary greatly. Markov state modeling was used to investigate the different macrostates that could be associated with the different lipid II binding modes, including an approach toward a feasible conformation for the catalytic elongation of the polysaccharide. The choice of metrics from which to generate a Markov state model that makes sense with regard to the that we wanted to observe was not trivial. We finally chose to discretize our large simulation using the coordinates of the lipid II sugar and diphosphate moieties as metric, because we were particularly interested in finding plausible orientations of these groups for lipid IV formation by SAMGT. The final Markov model, generated from the microsecond scale cMD simulations of system 2, resulted in
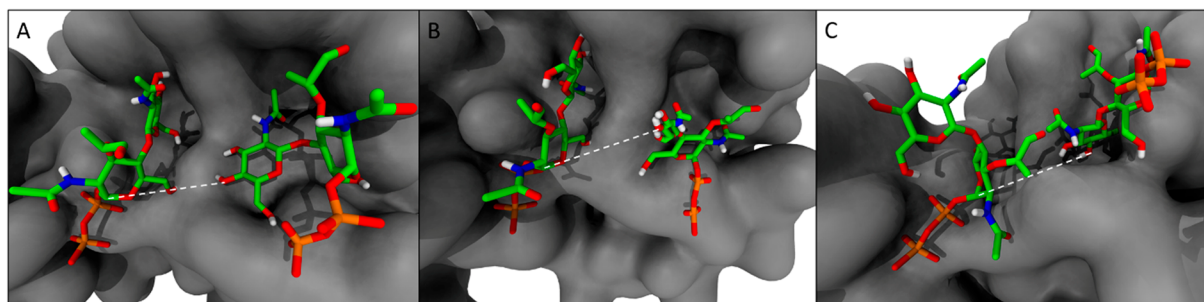
Figure 5. Averaged conformation of every macrostate for the three-state Markov model of SAMGT. Only the lipid II core (sugar moiety and diphosphate) and the protein are shown for clarity. States A−C are shown in panels A−C, respectively. The white dotted line in each panel denotes the distance between O_4a and C_1d. Donor molecules are located on the left and acceptor molecules are located on the right side in each panel.

Table 2. Equilibrium Probability, Relative Gibbs Free Energy and Key Distances between the Atoms Involved in the Proposed Reaction Coordinates for the Different Macrostates and over the Full Trajectory That Was Used to Generate the Markov Model

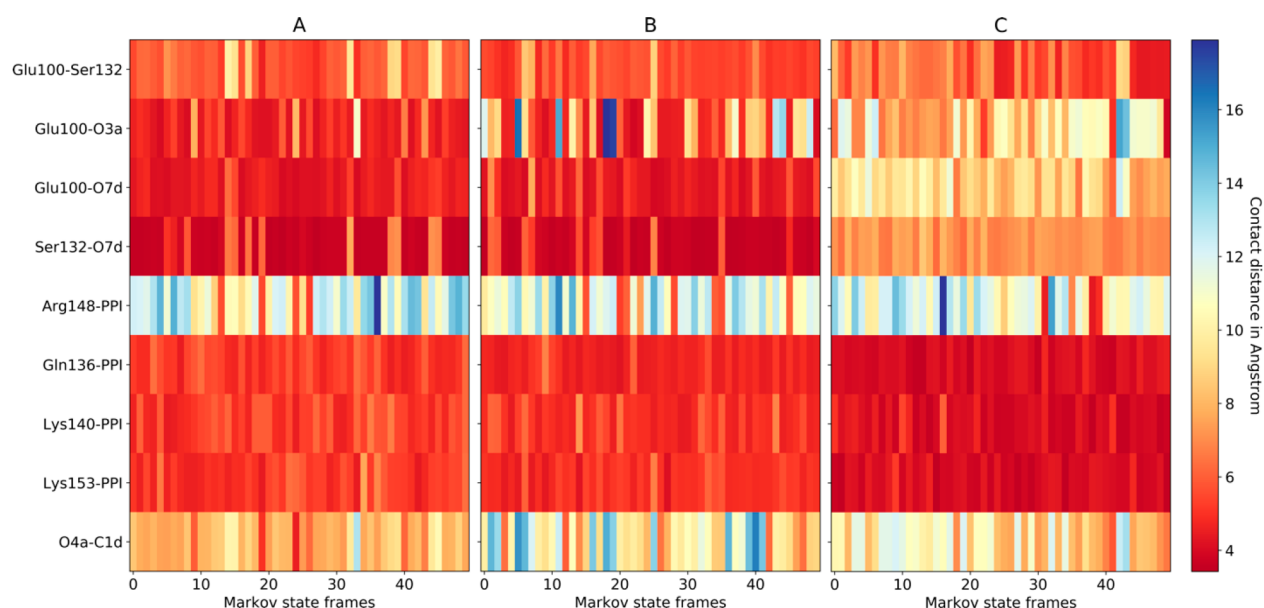|	state A	state B	state C	full trajectory
$d(O_4a−C_1d)$ (Å)	7.0 ± 1.5	11.2 ± 3.2	8.7 ± 2.0	8.9 ± 3.3
$d(O_4a−C_1d) < 5$ Å (%)	14	0	2	9
$d(O_4a−C_\delta \text{ Glu}_{100}) < 5$ Å (%)	88	32	6	36
equilibrium probability (%)	1.6	2.3	96.1	
relative Gibbs free energy (kcal/mol)	$+1.7 \pm 0.9$	$+1.1 \pm 1.0$	0	

Figure 6. Color maps of key distances within every macrostate. Each panel A−C describes the corresponding states A−C. PPI, diphosphate moiety of donor lipid II. All distances reported are between heavy atoms. In the case of glutamate and arginine, distances were taken from the $C\delta$ and the $C\zeta$ carbons, respectively. Distances from the diphosphate were taken from the P1 and P2 phosphorus atoms.

three distinct macrostates in which both lipid II substrates were bound to SAMGT (Figure 5). This number of macrostates was chosen as the implied time scales suggested that there are three processes that take place on a statistically relevant time scale (Figure S3). Trial and error with different numbers of states confirmed that three macrostates resulted in a Markov model with the highest quality.

In state A, the attacking O_4a oxygen is found to be oriented toward the donor lipid II, and the C_1d position of the donor lipid II molecule is oriented in such a way that a nucleophilic attack would be possible from a geometrical point of view. In state B, the sugar core of the attacking lipid II molecule is more

flexible, making it less suitable as a catalytic starting conformation. This is further confirmed by the larger mean distance $d(O_4a−C_1d)$ that was observed in state B as compared to state A (Table 2). In state C, both substrates are oriented significantly differently compared to state A and state B, which seems to be the result of a conformational change in the flap domain. While the mean distance between O_4a and C_1d is smaller in state C than in state B, only in 2% of the representative state C frames the distance is smaller than 5 Å, while in state A this occurs in 14% of the frames (Table 2). State C is thermodynamically the most favored state for SAMGT bound by two substrates, while state A, with a relative

Gibbs free energy of $+1.7 \pm 0.9$ kcal/mol when compared to state C, is the least thermodynamically favored. Figures S3–S5 show additional statistics on the kinetics between the states.

The key distances in Table 2 indicate that state A, which is the highest energy metastable state, would be the only state possible to take into consideration for catalytic activity. The reaction coordinate distance between O_4a and C_1d of less than 5 Å is found only in states A and C. However, the relative orientation of the lipid II molecules makes state C not suitable for initiation of the nucleophilic attack. The thermodynamic penalty to access state A from state C is however quite small.

Figure 6 shows that the key contacts in state A and state B are similar, with exception of the distance that is important for catalysis, $O_4a–C_1d$. While state C is clearly distinct from A and B, it can be noted that the only persistent contacts over all three states are those that assist in stabilizing the diphosphate moiety, signaling the importance of this moiety to lipid II binding.

Investigation of the Catalytic Mechanism. State A of the Markov model provided insight on a putative transition state (TS)-like orientation of lipid II in the active site and was used as starting point for the investigation of the catalytic mechanism. After visual analysis of the state, several evaluation functions were used to score every frame along the MD trajectories based on interactions within the active site (see Methods for more details). In total, 17 conformations were extracted from the full trajectory as starting points for the evaluation of the reaction mechanism of SAMGT using QM/MM calculations. In order to elucidate the most likely reaction mechanism, a number of potential mechanisms were evaluated by calculating potential energy surface scans using the appropriate guess reaction coordinate(s) as variable parameter(s). In each case, the corresponding reaction coordinate was scanned using a hybrid QM/MM approach that was performed using Gaussian09 with the ONIOM scheme.[25−28] All energy scans were performed at the B3LYP/6-31G(d):AMBER level.[29−32]

S_N2-type Mechanism with Glu_{100}. Deprotonation of the O_4aH-hydroxyl did not occur when scanning the $d(O_4a–C_1d)$ reaction coordinate alone (data not shown). In order to force proton transfer nevertheless, a two-dimensional potential energy scan across the reaction coordinates $d(O_4a–C_1d)$ and $d(H_4d–Glu_{100})$, which are the respective reaction coordinates for the nucleophilic attack and the deprotonation of the lipid II hydroxyl, was performed. The potential energy surface scan confirmed the 1D-results and showed this to be an endothermic process with an estimated activation energy of about 76 kcal/mol (Figure 7).

S_N1-type Mechanism with a Carbocation Intermediate and Glu_{100} as the Catalytic Base. According to this mechanism, the first step would be the spontaneous breakage of the bond linking C_1d with the O of the connected phosphate moiety. It would be expected that a local minimum would appear along the reaction coordinate corresponding to the carbocation intermediate formed after cleaving the donor lipid II diphosphate moiety from the sugar core. In reality, this local minimum could not be reproduced, as the free energy only increased along this reaction coordinate to a plateau of 25 kcal/mol (data not shown), hence suggesting that the carbocation intermediate is too unstable and therefore not a suitable intermediate.

In both potential energy surface scans, Glu_{100} was considered to be the catalytic base. However, neither of
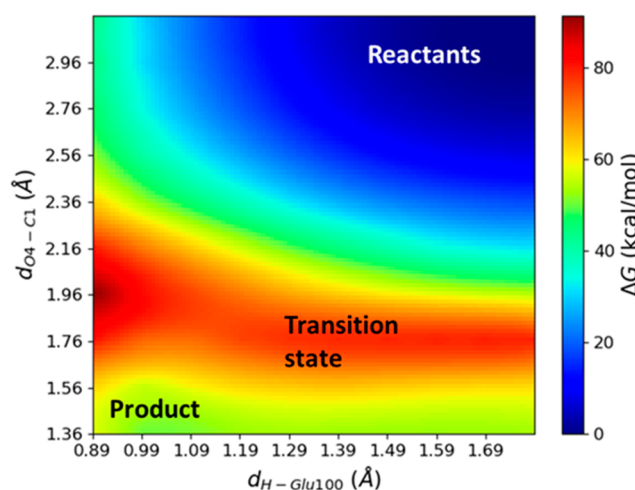

Figure 7. Two-dimensional potential energy surface for the S_N2-type mechanism with Glu_{100} as the catalytic base and SAGMT modeled in an open loop conformation, calculated at the B3LYP/6-31G-(d):AMBER level. The reaction coordinate for the deprotonation step is plotted on the x-axis, and the reaction coordinate for the nucleophilic attack by O_4a on C_1d is plotted on the y-axis.

these mechanisms was able to generate a favorable reaction energy profile. These calculations suggest that it seems unlikely that Glu_{100} is the base that is deprotonating the acceptor lipid II-hydroxyl, which might be explained by the fact that the negative carboxylate of the side chain of this residue is found to be heavily stabilized by hydrogen bonds and salt bridges with Arg_{241}, Asn_{224}, Ser_{132}, donor and acceptor lipid II sugars, and water. Overstabilization of the Glu_{100} residue decreases its pK_a, making it unsuitable to act as a base. Proton transfer to Glu_{100} does not result in a stable intermediate, since no local minimum could be observed along the reaction coordinate $d(H_4d–Glu_{100})$ describing the path between the transferred proton and the Glu_{100} oxygen (data not shown).

The failure to explain the catalytic mechanism with Glu_{100} as catalytic base led us to explore an alternative mechanism with a different catalytic base.

A Metal Hydroxide-Catalyzed Nucleophilic Attack. The alternative mechanism that was explored was a concerted S_N2 mechanism in which the proton transfer was occurring between the acceptor lipid II hydroxyl and a magnesium-bound hydroxide ion, as similar mechanisms have been found to be the main mechanism of catalysis for several enzymes.[23,33,34] These mechanisms take into consideration the existence of an exchange equilibrium between water molecules or hydroxide ions in bulk solvent and at the active site (eq 1):

$$H_2O_{(prot)} + HO^-_{(aq)} \rightleftharpoons H_2O_{(aq)} + HO^-_{(prot)} \qquad (1)$$
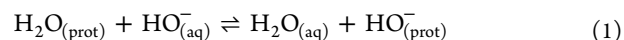
The exchange free energy for the transfer is usually small and has been calculated to be close to zero in other enzymes as will be discussed below. The coordination sphere of the active site magnesium ion contains two water molecules in a geometry that was extremely consistent during MD simulations. The water molecule positioned in a suitable orientation for assisting in the nucleophilic attack was transformed into a hydroxide ion in order to study this reaction mechanism. The configuration of the transformed active site can be seen in Figure S6.

Potential energy scans similar to the previous ones with Glu_{100} as base were performed starting from nine different

Table 3. Energy Barriers and Key Interaction Distances for the Nine Optimized TS Conformations That Were Tested in the Metal Hydroxide-Catalyzed Reaction Mechanism[a]

energy barrier (kcal/mol)	distance (Å)								
	diphosphate–Arg$_{148}$	diphosphate–Lys$_{140}$	diphosphate–Lys$_{153}$	diphosphate–Gln$_{136}$	O$_7$d–Glu$_{100}$	O$_3$a–Glu$_{100}$	Arg$_{241}$–Glu$_{100}$	O$_4$a–C$_1$d	C$_1$d–diphosphate
16.1	1.9	1.0	1.7	2.1	1.6	2.0	3.6	2.4	2.2
18.6	1.9	1.8	1.8	1.9	4.4	1.9	3.6	2.4	2.3
19.7	1.9	1.5	1.7	2.2	1.7	1.7	3.2	2.1	2.7
23.6	2.7	2.6	1.9	2.0	1.6	1.7	3.6	2.2	2.4
26.7	1.7	3.7	1.5	1.9	1.7	1.7	1.8	2.2	2.2
27.2	2.1	3.6	1.6	1.6	1.8	2.1	3.7	2.1	2.5
29.6	2.7	3.2	1.6	2.0	1.7	1.8	3.0	2.2	2.3
33.1	2.5	1.6	1.6	1.7	1.6	1.8	2.1	2.2	2.2
35.5	5.1	2.8	1.8	3.3	1.6	1.8	5.4	2.3	2.5

[a]Distances are given in angstrom. Distances to diphosphate, Glu$_{100}$, and Arg$_{148}$ were measured as the distance to the nearest oxygen or nitrogen.
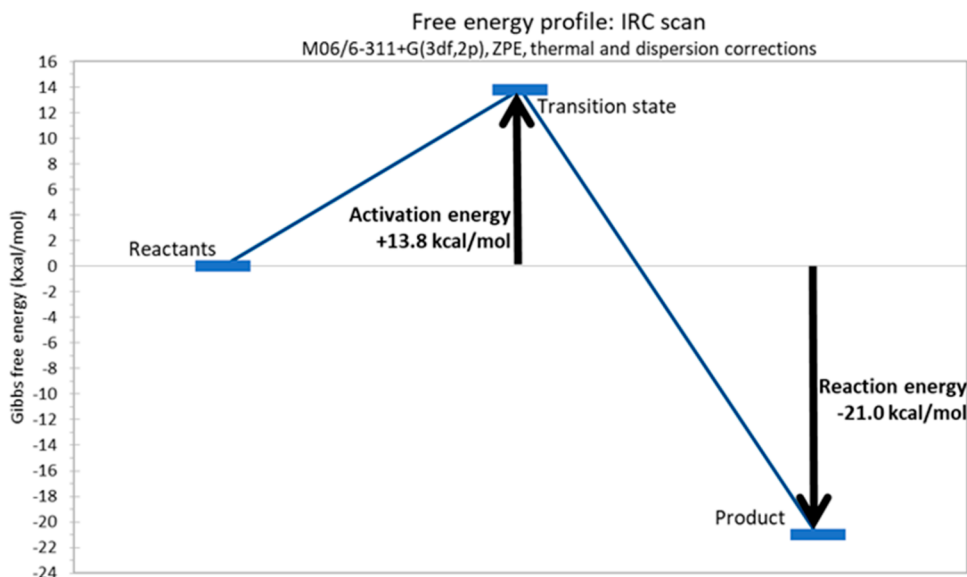
Figure 8. Accurate free energy profile of the metal hydroxide-catalyzed reaction mechanism, calculated at the M06/6-311+G(3df,2p):AMBER level and including zero-point energy, thermal, and dispersion corrections.
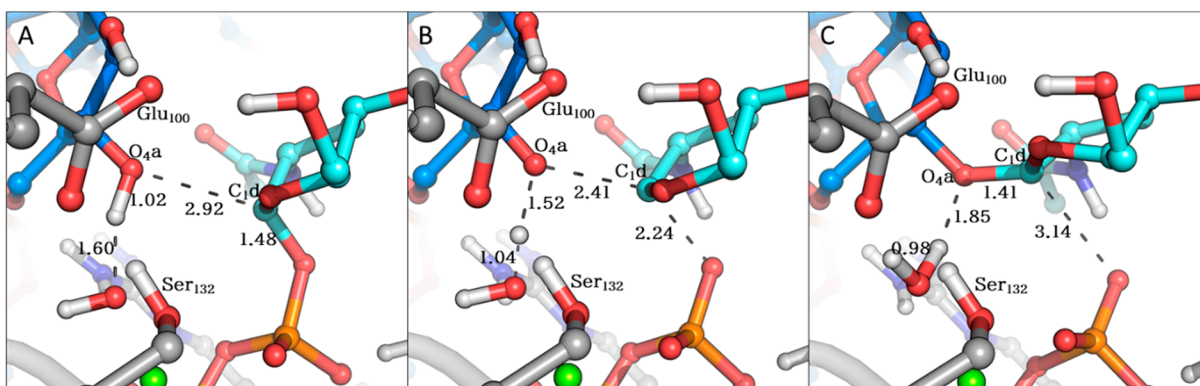
Figure 9. Stationary states along the intrinsic reaction coordinate, going from reactants (A) over the transition state (B) and to the final product (C). The protein is shown in gray cartoon with key residues in gray balls and sticks. Key residues are labeled. Donor lipid II is shown in light blue balls and sticks; acceptor lipid II is shown in dark blue balls and sticks. Mg^{2+} is represented as a green sphere.

starting conformations, taken from state A of the Markov model. The $d(O_4a–C_1d)$ reaction coordinate was scanned using a hybrid QM/MM approach that was performed using Gaussian09 with the ONIOM scheme. All energy scans were performed at the B3LYP/6-31G(d):AMBER level (Table 3).

These scans showed this alternative mechanism to be far more favorable than any of the other tested mechanisms with Glu$_{100}$ as base. The reaction was found to be exothermic. The lowest TS resulted in an activation barrier of 16.1 kcal/mol at the B3LYP/6-31G(d):AMBER level, which is well within the
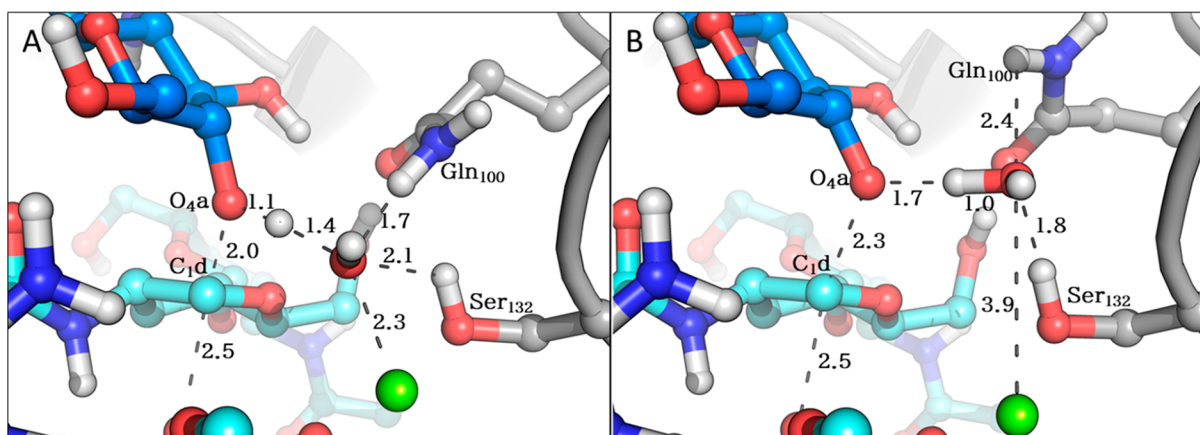
Figure 10. Active site distances before (A) and after (B) geometry optimization of the transition state-like structure in the $E_{100}Q$ mutated structure. The protein is shown in gray cartoon with key residues in gray balls and sticks. Key residues are also labeled. Donor lipid II is shown in cyan-colored balls and sticks; acceptor lipid II is shown in dark blue balls and sticks. Mg^{2+} is represented as a green sphere.

expected range of enzyme-catalyzed reactions. Moreover, this conformation of the enzyme does not differ significantly from the crystal structures considered: the backbone RMSD of residues 65 to 269 is 1.263 Å to the closed conformation (PDB ID 3VMT) and 1.070 Å to the open conformation (PDB ID 3HZS). Table 3 gives a full overview of activation energies and possible key interaction distances for the nine conformations in which this mechanism was studied.

After full TS optimization and confirming that the TS relaxes back to reactants and product by following the corresponding intrinsic reaction coordinate (IRC), a more accurate estimate of the activation energy was calculated for the conformation with the lowest energy barrier in Table 3. The three stationary states of this system (reactants, TS, and product), were fully optimized using the M06 density functional with an extended 6-311+G(3df,2p) basis set. After applying the necessary zero-point energy, thermal, and dispersion corrections, the activation energy barrier (ΔG_A) was found to be 13.8 kcal/mol (Figure 8). A detailed figure of the resulting stationary states is given in Figure 9.

To obtain a complete estimate of the barrier associated with this catalytic mechanism, two additional terms had to be included. The free energy for the exchange of a water molecule in the coordination sphere of the active site Mg^{2+} with a bulk solvent hydroxide at a pH of 7 has to be included (eq 1). This correction term has two components. The first component is the standard state free energy associated with exchange of a water molecule in the Mg^{2+} coordination sphere with a hydroxide ion (ΔG_{ex}), which was calculated with alchemical transformations using Langevin dynamics, paired with the Multistate Bennett Acceptance Ratio.[35,36] The value was found to be -9.7 ± 1.0 kcal/mol (see Methods). Expectably, it is favorable to exchange water by hydroxide in the Mg^{2+} coordination shell because the hydroxide binds much more strongly to the Mg^{2+} ion. The second term is a correction for the non-standard-state concentrations (10^{-7} mol dm^{-3} for hydroxide and 55.6 mol dm^{-3} for water), which can be calculated as an entropic penalty, $\Delta G_{exchange}$, associated with the restriction of free volume accessible to a hydroxide when moving from the bulk solvent to the Mg^{2+} coordination sphere, which has only two sites available for hydroxide binding. This entropic penalty $\Delta G_{exchange}$ has been calculated in many works before and is derived from the particle-in-a-box model:

$$\Delta G_{exchange} = -k_B T \ln\left(\frac{V_f}{V_i}\right) \tag{2}$$

with V_f being the volume of two water molecules for Mg^{2+} coordination with a reference water density of 0.997 g/cm^3 and V_i being the accessible volume of hydroxide in bulk solvent, which was calculated by assuming a pH of 7. At a temperature of 310 K, eq 2 yields a free energy penalty $\Delta G_{exchange}$ of +12.0 kcal/mol.[37−39] As such, the calculated total apparent free energy barrier $\Delta G^{\ddagger} = \Delta G_A + \Delta G_{ex} + \Delta G_{exchange}$ is found to be 16.1 ± 1.0 kcal/mol. Experimental studies on SAMGT have shown it to have a turnover number, k_{cat}, of 0.40 ± 0.063 s^{-1}.[20] Using transition state theory on this turnover number, the experimental free energy barrier $\Delta G^{\ddagger}$ for the reaction is calculated to be around 18.7 kcal/mol,[40] in excellent agreement with our theoretical estimates based on the model with a magnesium-bound hydroxide ion as catalytic base. It is well-known that the activation free energy is very sensitive to mechanistic and structural inconsistencies, and such inconsistencies are reflected in a higher value for the activation free energy. Despite the limitations that the modeling of the Mg^{2+} ion might bring, the excellent reproduction of the experimental activation free energy is a strong indication that the modeling was essentially correct. This validates both the predicted structure and the predicted mechanism.

Validation of the Proposed Catalytic Mechanism. Multiple papers have shown the importance of the glutamate residue in the active site for glycosyltransfer activity. Mutational studies have proven that enzyme activity is reduced 500-fold when the glutamate is mutated to a glutamine, suggesting an important role in the reaction mechanism. This formed the basis of the hereto proposed mechanism with Glu_{100} as catalytic base.[14,16−19] However, as the here proposed reaction mechanism does not predict Glu_{100} to be directly involved in the mechanism, a critical view on the role of this Glu_{100} was needed. In addition, the importance of divalent metals is heavily implied in several papers.[14,19] To confirm this, the impact of Mg^{2+} in the active site should also be investigated.

In the first instance, the impact on the activation energy of an *in silico* mutation of Glu_{100} to Gln was investigated. Glu_{100} is involved in a number of key hydrogen bonds with nearby residues as well as both substrates. The residue seems to play an important role in organization of the binding site.

Consistent interactions of Glu_{100} could be seen with Ser_{132}, H_7d, and H_4a in all starting conformations that were used in this study. The reaction mechanism was tested with the $E_{100}Q$ mutation for the three initial conformations with the lowest activation energy in the wild type system. In two cases, the reaction was not able to proceed because Gln_{100} disrupted the active site configuration by forming a hydrogen bond with the catalytic hydroxide, whereas the original Glu_{100} forms a hydrogen bond with Ser_{132} instead. In the third conformation, optimization of the TS-like structure was successful. However, while the negative vibrational frequency of the saddle point structure corresponded to the reaction coordinate, indicating that the TS is correct for this reaction, the conformation differs significantly from the initial guess TS structure. More specifically, the catalytic hydroxide diffused away from the Mg^{2+} coordination sphere, indicating that the metal hydroxide becomes unstable upon forming a hydrogen bond with Gln_{100} (Figure 10). It therefore seems that one of the functions of the Glu_{100} residue could be to keep the hydroxide ion in its optimal location near the Mg^{2+} ion by means of electrostatic repulsion. Upon an IRC scan of the mutated enzyme toward reactants and product, the diffused hydroxide did not return to the Mg^{2+} coordination sphere. After corrections, the activation energy for this $E_{100}Q$ mutated system was 26.2 kcal/mol, well above the 13.8 kcal/mol that was calculated for the unmutated system. The reaction energy for this system was −5.6 kcal/mol, compared to the reaction energy of −21.0 kcal/mol that was calculated for the wild type system.

Second, we also evaluated the effect of replacing the active site Mg^{2+} with Na^+. While the energy scan along the $d(O_4a−C_1d)$ reaction coordinate had an estimated activation energy barrier of around 23 kcal/mol, it was not possible to calculate accurate energies for this mechanism. Upon optimization of the transition state, the protonated hydroxide always drifted away from the active site, making it impossible to localize a saddle point. Furthermore, the lipid II hydroxyl group got fully deprotonated immediately at the start of the reaction coordinate scan, suggesting that the hydroxide would be too unstable to catalyze this reaction in the absence of a divalent metal ion.

Third, an S_N1-type mechanism, with a carbocation intermediate and hydroxide as the catalytic base, was evaluated as well to assess if the presence of a stronger base to catalyze the nucleophilic attack has an influence on this mechanism. However, the same observation was made as with the Glu_{100} mechanism, in which the energy kept increasing along the reaction coordinate, indicating that the carbocation is not stabilized and therefore unlikely to be formed.

■ DISCUSSION

Available crystal structures of SAMGT reveal two distinct binding pockets for a conformation in which the flap domain is in the closed position and one binding pocket for a conformation in which the flap domain is in the open position. It is well-known that it is difficult to obtain crystal structures for membrane proteins and that they usually have a low resolution. The electron densities and B-factors of the SAMGT crystal structures reveal that there is a significant structural uncertainty regarding the active site. The crystal structure with a lipid II analog in the acceptor binding site (PDB ID 3VMT), for example, has B-factors of 91 Å^2 and 89 Å^2 for the two Mg^{2+} ions. The lipid II analog itself has a mean B-factor of 103 ± 27 Å^2. The flexible region including residues 100−116 has a mean

B-factor of 106 ± 35 Å^2. Comparatively, the full protein has a mean B-factor of 74 ± 35 Å^2. A recent paper states that for crystal structures at low resolution, B-factors with an upper limit of 80 Å^2 are acceptable, with anything over 100 Å^2 resulting in unsubstantiated coordinates.[41] Important parts of the active site have consistently higher B-factors than this cutoff, and the mean B-factor of the entire structure is around 80 Å^2 as well. Therefore, the X-ray structures provided a good starting ground for our structural and mechanistic studies, and we added a layer of modeling over the initial structures to further define the missing details. Given the great flexibility of large parts of the protein and substrates, the cMD simulations performed in this study have given insight into why the B-factors might be of this magnitude. A novel theory of the binding mode of lipid II to SAMGT is presented, which conforms to the existing experimental observations.

Function of Divalent Metals in SAMGT and the GT51 Family. Previous studies have indicated the importance of divalent metal ions for activity of enzymes within the GT51 family, but their involvement in the reaction mechanism has never been studied in detail.[14,19,42] This study strongly suggests that, in the case of SAMGT, divalent metal ions have a function in catalysis. In the active site, the metal ion is necessary for correct orientation of the catalytic hydroxide. Presumably, the divalent metal also aids in stabilization of the donor lipid II diphosphate group after it is cleaved. However, this was not observed when reconstructing the activation free energy with Mg^{2+} replaced by Na^+, where the main impact appeared to be the stabilization of the catalytic hydroxide ion. On the other hand, it is possible that the divalent metal enters the active site in complex with lipid II and improves binding affinity toward SAMGT.

Reaction Mechanism of Glycosyltransfer in the GT51 Family. The activation free energies calculated in this study can be compared to experimental findings through transition state theory.[40] We can see large fluctuations in the predicted activation energy barriers, even if conformational states of reactants and transition states are very similar between different systems. These discrepancies likely stem from differences in nonbonded interactions due to thermal conformational fluctuations.[37] More specifically, the data suggests that the degree of stabilization of the donor lipid II diphosphate moiety is strongly correlated to the predicted activation energies (Figure S7). As we do not seek to accurately replicate experimental measurements of the activation barriers but rather are looking for feasible hypotheses for the reaction mechanism, we accumulated data in search for the lowest observable barrier instead of a statistically significant measurement of the energy barriers in the investigated reaction mechanism. Such a study would require many more starting conformations to be tested before a statistically significant result can be shown.[37,43] To make this computationally feasible, the Hamiltonian would have to be downgraded, limiting even further the accuracy of the converged free energy. However, our findings do suggest that a metal hydroxide-catalyzed S_N2 reaction appears to be more likely than the alternative reaction mechanisms that were tested, including the S_N1-type mechanism that had been proposed over two decades ago[17] and the S_N2-type mechanism with a catalytic glutamate that has been generally accepted since the early 2000s.[19] The predicted energy barrier of 16.1 ± 1.0 kcal/mol is in line with the experimentally observed value of 18.7 kcal/mol, suggesting that the metal hydroxide catalyzed

nucleophilic attack is potentially the main pathway for polysaccharide synthesis in the GT51 family.

We note that slight differences from the actual energy barrier are expected, as each parcel from the total activation free energy has associated errors derived from methodological approximations. This can be particularly relevant in the determination of the free energy required to place a hydroxide molecule from the bulk solvent into the active site of SAMGT (ΔG_{ex}), as it is well-known that alchemical transformations leading to a change in the net charge of the system can imply artifacts during the PME calculation step. The coupled insertion of a counterion could be performed to keep the neutrality of the system;[44] however, given the size of our system, we do not expect such artifacts to have as much of an impact on the calculations as the alternative insertion of an additional counterion. Furthermore, the parameters for the hydroxide ion generated by the ATB server could also imply some level of error, because we did not perform extensive validation. Nevertheless, given that the absolute hydration free energy of the hydroxide calculated by this methodology was within 30% of the experimental values,[45] we believe that our results are qualitatively correct, aside from a small bias due to the force field parametrization.

Other tested mechanisms, like a glutamate-catalyzed mechanism, a mechanism without Mg^{2+} in the active site, and the metal hydroxide catalyzed mechanism in an $E_{100}Q$ mutant, resulted in either failure to find a physical path from reactants to product or unrealistic activation energy barriers.

The work by Terrak and Nguyen-Distèche on the pH dependence of the glycosyltransferase activity of SAMGT shows that the enzyme activity is increasing with increasing pH up to an optimal of around 7.5−8.[14] Similar conclusions were made by Schwartz and co-workers on the glycosyltransfer activity of PBP1b from *E. coli*[19] and by Nakano et al. in their work studying ribozyme-catalyzed self-cleavage of a phospho-diester bond.[34] These experimental observations are in line with the here proposed mechanism, as well as with the Glu_{100}-based mechanism. To allow a further distinction between the possible reaction mechanisms, the pH-dependent behavior of the enzyme would have to be further characterized at higher pH, as inconsistencies can be seen here between studies. An alternative approach would be the assessment of the kinetic isotope effects of ^{18}O-water in order to distinguish a metal-bound hydroxide mechanism from other possible reaction mechanisms.[46]

Different Binding Modes for lipid II and lipid IV. The proposed binding mode and reaction mechanism support the experimentally observed shuffling mechanism in which newly synthesized lipid IV molecules get shuffled into the donor site, and a new lipid II molecule can bind in the acceptor pocket to maintain a processive linking mechanism. A second, distributive mechanism has also been observed experimentally when the only available substrate is lipid IV.[16] Numerical modeling has previously shown that a model in which polysaccharide synthesis happens via both a lipid II (processive) and a lipid IV (distributive) linking mechanism is the most accurate, supporting that some proteins in the GT51 family can have both a processive and a distributive reaction mechanism.[47] However, not all GTase enzymes in the GT51 family appear to accept the same substrates.[48] It can be speculated that the distributive nature of the lipid IV linking mechanism could mean that the binding pocket that was previously observed in a crystal structure of SAMGT (PDB ID 3VMT) is the acceptor binding site for the lipid IV linking mechanism. The mechanism would then be distributive because the flap domain is blocking the shuffling of the formed product into the donor site, causing it to dissociate eventually. Additional studies would have to be performed to confirm this proposal.

CONCLUSION

The binding mode of lipid II to SAMGT was investigated at an atomistic level using cMD simulations. Simulations with a combined length of over 20 μs revealed that the binding mode of the donor lipid II is similar to that of moenomycin A in the crystal structures. However, while it was generally assumed that the donor and acceptor binding site are separated by the flexible flap domain, it is more likely that both lipid II substrates occupy one large pocket, with the flap domain in an open conformation. A Markov model gave more insight into the different macrostates associated with this hypothesis, with one macrostate being associated with an "activated" binding mode in which the acceptor lipid II is correctly oriented for attack and in close proximity to the donor lipid II. Classical MD simulations revealed the likely position of a divalent metal ion in the active site, which was shown to be essential for catalytic activity. The reaction mechanism of SAMGT was investigated using hybrid QM/MM calculations. While several mechanisms were tested, the only mechanism that was found to have a reasonable activation energy happens through an S_N2-type attack with a metal hydroxide as the nucleophilic catalyst. The observed mechanism has an estimated activation energy of 13.8 kcal/mol and an apparent free energy barrier of 16.1 ± 1.0 kcal/mol. This in line with the experimentally observed value of 18.7 kcal/mol, providing support both to the proposed mechanism and the modeled structures. We were unable to find a reasonable energy barrier for the generally proposed S_N2-type mechanism with a glutamate residue as the nucleophile, even though this mechanism was supported by both the 500-fold reduced activity of the $E_{100}Q$ mutant and the pH-profile of the enzymatic activity. However, this mechanism is endothermic and activation energy barriers of over 70 kcal/mol were observed. The metal hydroxide-catalyzed mechanism was also tested on the $E_{100}Q$ mutant, where a significantly higher estimated activation energy of 26.2 kcal/mol was observed. Our results suggest that the mutant glutamine could disrupt the active site configuration by forming a hydrogen bond with the catalytic hydroxide. Therefore, Glu_{100} is still an essential residue for SAMGT activity in the hydroxide-catalyzed mechanism. We thus propose that a lipid II linking mechanism utilizing a metal hydroxide as catalytic base is the main catalytic pathway of the GT51 family of GTases.

In summary, this work provides valuable atomic-level insights into SAMGT. The study helps us to understand many fine, atomic-level structural and mechanistic details SAMGT and helps to guide further rounds of experiments to further clarify the missing links in the understanding of this important antibacterial target.

METHODS

Structure Preparation. The initial model of system 1 was based on the crystallographic structures of lipid II analogue-bound and moenomycin-bound SAMGT (PDB ID 3VMR and 3VMT, respectively). Donor lipid II positioning was modeled after the diphosphate moiety and rings E and F of moenomycin (Figure S8). Acceptor lipid II positioning was

modeled after the binding pose of the lipid II analogue. Two Mg^{2+} ions were positioned according to interactions in crystal structure 3VMT. The starting conformation for the protein was taken from 3VMT, being the crystal structure with the higher resolution. Protein residues were kept in their predominant protonation states as predicted by the PDB2PQR server, and residue flips were predicted with MolProbity.[49,50] The first 40 N-terminal amino acids were missing and were initially modeled as an α-helix in Chimera.[51] Amino acids 123 to 129 were also missing and were modeled with MODELLER.[52]

A similar approach was taken for system 2, but the initial protein conformation was taken from a different crystal structure of moenomycin-bound SAMGT (PDB ID 6FTB) instead. Because of the shift in the flexible flap domain in the active site, it was necessary to manually infer new positions for the Mg^{2+} ions. In this crystal structure, the first 59 amino acids are missing. Amino acids 41−59 were modeled after the 3VMT crystal structure by superposition of the protein structures with the SPRUCE tool of the OpenEye software package.[53] The first 40 amino acids were modeled as above.

Molecular Dynamics. *System Setup.* An initial box was set up with the SAMGT complex in the center, aligned with the Z-axis. Phospholipids were added to a rectangular box in random orientations. The phospholipid composition within the system was chosen to reflect experimentally determined *S. aureus* bilayer composition, with a lipid composition of 152 phosphatidylglycerol molecules, 122 lysyl-phosphatidylglycerol molecules, and 68 cardiolipin molecules.[54−56] The Gromos54A7 united-atom force field was used to parametrize the system.[57] The Gromos suite of force fields provides an efficient approach for long MD simulations of membrane systems as they have been parametrized with partitioning behavior at the hydrophobic−hydrophilic surface in mind.[58] As no Gromos54A7 parameters were available for lipid II and lipid II was too complex for parametrization with the ATB server, parameters were derived from existing building blocks. Di-*anteiso*16:0-lysyl-phosphatidylglycerol, di-*anteiso*16:0-phosphatidylglycerol, and tetra-*anteiso*16:0-cardiolipin parameters were adapted from previous works.[8,59] The charges for phospholipid phosphate moieties described in the Gromos53A6 force field paper resulted in strong interaction of the lysyl phosphatidyl phospholipid head groups with the phosphate moieties, deforming the bilayer. Partial charges of the phosphate moieties of the phospholipids were thus calculated using Spartan (https://www.wavefun.com/products) at the CAM-B3LYP/cc-pVTZ level and RESP[60] fitted (Figure S9), which alleviated the issue. In order to prevent lipid II sugar ring flipping and configurational changes due to elevated temperatures, dihedral constraints were applied to the 12 ring dihedrals with a force constant of 10 kJ mol^{-1} rad^{-2}. The simulation box was then solvated in an explicit SPC water box.[61] Na^+ ions were added accordingly to neutralize the system. The resulting system had box dimensions of $11 \times 11 \times 18$ nm^3 and contained roughly 155 000 atoms.

Energy minimization was performed with the steepest descent method followed by the conjugate gradient method. The system was then equilibrated under constant volume and temperature (NVT) conditions for 500 ps at 310 K and later at constant pressure and temperature (NPT) for 1 ns. Conditions were set at 1 atm pressure and 310 K with a Berendsen barostat and thermostat.[62] All bonds were constrained using the LINCS algorithm to allow for a 2 fs time step.[63] Periodic

boundary conditions were applied and PME electrostatics and nonbonded cutoff values of 1.4 nm were used without dispersion correction. All protein heavy atoms and lipid II atoms were constrained with a force constant of 1000 kJ mol^{-1} nm^{-2} during equilibration.

In order to simulate spontaneous membrane self-assembly, a flat-bottomed potential was applied on the head and tail groups of the phospholipids after an initial equilibration, while keeping the protein−substrate complex in place using force constraints. All protein heavy atoms and lipid II atoms were constrained with a force constant of 1000 kJ mol^{-1} nm^{-2}. For the flat-bottomed potential, force constants of 10 kJ mol^{-1} nm^{-2} and 50 kJ mol^{-1} nm^{-2} were used on the terminal carbon of every phospholipid tail and on every phosphorus atom in the phospholipid head groups, respectively. Assembly of a full bilayer was seen within 20 ns of simulation time.

After membrane self-assembly, N-terminal amino acids 1−66 were allowed to relax and equilibrate for 100 ns, while the rest of the complex was still restrained. After membrane assembly and relaxation of the N-terminal residues, a new round of equilibration was performed as described above. After equilibration, the force constraints were removed and the thermostat and barostat were switched to velocity-rescale with a time constant of 0.5 ps, and Parrinello−Rahman with a relaxation time of 5 ps, respectively.[64,65] Pressure was regulated through semi-isotropic scaling. Long-range forces were calculated every step, while the neighbor list was updated every five steps. Recent work has shown that when using newer versions of GROMACS in conjunction with the Gromos force fields, bilayer properties deviate from those obtained upon parametrization of the force field when using larger intervals for calculation of long-range forces with the original twin-cutoff scheme. However, using a single cutoff of 1.4 nm allows for correct reproduction of force field behavior.[66]

Simulations. All classical MD simulations were performed using GROMACS 2018 or GROMACS 2019.[67] For system 1, an initial simulation of 200 ns was performed after the system was equilibrated. After equilibration of the N-terminal domain, excessive water molecules were removed from the system to reduce its size. Afterward, five independent runs of 500 ns were started with different starting velocities. Three more simulations of 500 ns were performed at elevated temperature, for an aggregated simulation time of 4.2 μs. For system 2, an initial simulation of 200 ns was performed after the system was equilibrated and excessive waters were removed after equilibration of the N-terminal domain. Afterward, five independent simulations of 100 ns were performed with different starting velocities. From these trajectories, 20 individual frames were picked, and a new simulation of 1 μs was started from each frame, for an aggregated simulation time of 20 μs. All simulations of system 2 were performed at 330 K to accelerate dynamics.

Markov State Model. Snapshots where unbinding of lipid II was observed were discarded before generating the Markov model, as this process was nonergodic in the limit of the simulation length. Several different metrics were tested, such as RMSD and RMSF of the full protein−substrate complex, RMSD of active-site residues, RMSD of full lipid II and truncated lipid II (core sugar and diphosphate groups), and contacts between key interacting residues and substrate functional groups, as well as combinations of these metrics. Ultimately, using the coordinates of the lipid II core groups resulted in the most consistent results. The systems were

aligned based on the protein Cα-atoms, and the Cartesian coordinates of the simulations were projected onto a feature space defined by the coordinates of the lipid II core groups (sugar moiety and diphosphate). The high-dimensional data from the simulation trajectories was then projected onto a two-dimensional space using time-lagged independent component analysis, with a lag time of 2 ns, and conformations were clustered into 200 microstates using the K-means clustering algorithm.[68] The coordinates of the lipid II core groups were used as the metric, by which a kinetic variance of 95% was retained.[69] Convergence was confirmed by plotting the implied time scales. Finally, the Markov model was generated with a lag time of 17 ns by further clustering the conformations into three macrostates. In order to test the self-consistency of the model, a Chapman–Kolmogorov test was performed.[70] Error estimates on the free energy differences between states were calculated by 100 bootstrapping iterations, in each of which 10% of the trajectories were randomly excluded. Generation and analysis were done within the HTMD environment.[71]

QM/MM Calculations. Hybrid QM/MM methods have been widely used to study enzymatic reaction mechanisms.[72] Here, we have used an adiabatic mapping method, which we believe is best suited for this kind of system.

Collecting Starting Conformations and System Setup. The different terms A, B, C, D, and E of the evaluation functions accounted, respectively, for $d(O_4a-C_1d)$, $d(H_4a-Glu_{100})$, the sum of distances between the carboxylate oxygens of Glu_{100} and potential hydrogen bond donors (Ser_{132}, Asn_{224}, Gln_{137}, H_7d), the sum of distances between diphosphate oxygens, and any nearby positively charged residues (Mg^{2+}, Lys_{140}, Arg_{148}, and Lys_{153}), and the distance between O_4a and the water molecule coordinating the nearby Mg^{2+}, which is replaced by a hydroxide ion in several mechanisms. These functions facilitated a cluster-like manual filtering method and allowed for the selection of a plausible but diverse set of starting conformations.

$$\frac{1}{A} + \frac{1}{B}$$

$$\frac{1}{A} + \frac{1}{B} + \frac{1}{C} - \frac{1}{D}$$

$$4\left(\frac{1}{A} + \frac{1}{B}\right) + \frac{1}{C} - \frac{1}{D}$$

$$\frac{1}{A} + \frac{1}{B} + 4\left(\frac{1}{C} - \frac{1}{D}\right)$$

$$\frac{1}{A} + \frac{1}{E}$$

The 10 highest scoring frames for each function were chosen (regardless of overlap in chosen conformations between different scoring functions), which resulted in 29 unique frames. These were then subjected to a visual inspection, with emphasis on relative orientation and key contacts. Finally, 17 conformations were selected for further investigation. Depending on the hypothesis, post-MD modifications were made to several residues. Water could be transformed into a hydroxide ion, Mg^{2+} could be transformed into Na^+, or Glu_{100} could be transformed into Gln_{100}. AMBER parameters for the QM/MM models were obtained using Antechamber 17.3, as available in AMBER18, using the AMBER14SB force field for amino acids

and the GAFF force field for phospholipids and lipid II.[73−75] Partial charges for phospholipids and lipid II were obtained using the RESP method, with charges fitted to the electrostatic surface potentials calculated with Gaussian09 at the HF/6-31G(d) level of theory.[28]

All systems were prepared using the MolUP plugin for VMD and GaussView5.[76−78] The system was truncated by removing all residues and molecules that were not in a 20 Å radius of O_4a, which was defined as the center of the active site, and at least two solvation layers were included for protein side chains. All atoms that were not within a 15 Å radius of O_4a and all water molecules in the MM region were constrained. The QM region contained 122 atoms in the final model (Figure S6). It consisted of one sugar ring from each lipid II as well as the diphosphate of the donor lipid II, the Mg^{2+} ion stabilizing the negative charge of the diphosphate group of the donor lipid II and its coordination sphere, all residues stabilizing the donor lipid II diphosphate group, Glu_{100}, a phosphate group of a phospholipid in close proximity to the catalytic hydroxide, and a glutamate forming a salt bridge with one of the diphosphate stabilizing residues. Several other hypotheses were also tested for the reaction mechanism, for which the residues in the QM layer could vary.

Calculations. QM/MM calculations were performed using Gaussian09 with the ONIOM scheme.[25−28] All geometry optimizations were performed at the B3LYP/6-31G-(d):AMBER level,[29−32] which has been shown to be sufficiently reliable for reproducing geometries of several enzyme reaction mechanisms.[72] Hydrogen atoms were used as link atoms upon truncation of the QM region.[79] The used reaction coordinate was the distance, $d(O_4a-C_1d)$. The structure that was highest in energy in the reaction path scans was used as the initial guess for the TS geometry optimization. Stationary points for reactants and product were then confirmed by following the IRC from the TS.[80] Vibrational frequencies of the stationary points were determined to confirm the nature of the stationary points (one imaginary frequency for the TS and no imaginary frequencies for the minima) at the B3LYP/6-31G(d):AMBER level, and zero-point energy and thermal contributions to the Gibbs energy were calculated using the harmonic oscillator/rigid-rotor formalism. Single point energy calculations were performed using the M06 density functional and an extended basis set (6-311+G(3df,2p)), as this hybrid meta density functional has been shown to give accurate estimates for the energy barrier in similar reactions.[81,82] Several other functionals were also tested to ensure that the results were qualitatively similar. Dispersion corrections were calculated using the DFT-D3 approach.[83]

Free Energy Calculations. The Multistate Bennett Acceptance Ratio method was used to calculate the difference in free energy of the system when one of the Mg^{2+}-coordinating waters was deprotonated.[35] Calculations were performed with the GROMACS 2019 software, using the Gromos54A7 force field. Energies were calculated using Pymbar.[84] The relative free energy was assessed using two different systems: one in which a hydroxide in bulk water solvent was alchemically transformed into a water molecule and one in which the Mg^{2+}-bound water was alchemically transformed into a hydroxide ion in the active site of SAMGT. Figure 11 shows the thermodynamic cycle used for the free energy calculations.
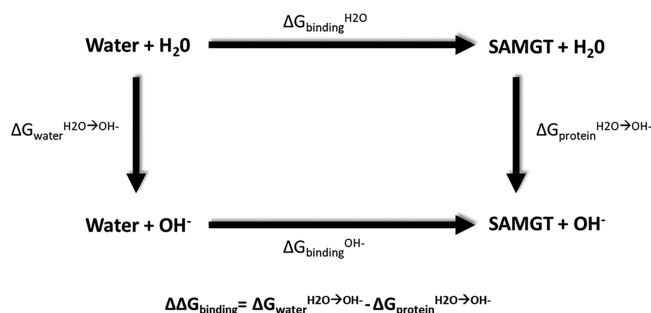
Figure 11. Thermodynamic cycle used to calculate the relative binding free energy for a hydroxide ion to the active site Mg^{2+} ion.

Partial charges and parameters for the hydroxide atoms were taken from the ATB server.[85,86] For the system with bulk solvent, a hydroxide ion was placed into a periodic box, which was then filled with 631 molecules of randomly oriented SPC water. For the system where the hydroxide is coordinating the Mg^{2+} ion, the snapshot corresponding to the lowest observed energy barrier was taken from the MD trajectory, in which a coordinating water molecule was replaced by the alchemical hydroxide. As hydrogens do not have any Lennard-Jones parameters in the Gromos54A7 force field, only a transformation of the electrostatic interaction parameters was necessary. The transformation of water to hydroxide was done in 20 linear steps, with a step size λ of 0.05 to ensure overlap between successive bins. The simulation length for each window was 5 ns for the transformation in bulk solvent. As this was an insufficient amount of sampling to obtain converged results for the transformation in the system where the hydroxide is coordinating the Mg^{2+} ion, simulation length was 100 ns for these calculations. All simulations were performed using the same settings as described for the classic MD simulations, but with a Langevin thermostat at a temperature of 310 K, with a collision frequency of 1 ps^{-1}. Both transformations were done in forward and backward directions. Because of significant hysteresis, the energy contributions of forward and backward calculations were averaged. To estimate the error on the free energy, block averaging was done to account for correlation within the MD trajectories.[87] As the hysteresis was larger than the statistical uncertainty, the hysteresis was used to estimate the error on the energy.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.0c00377.

Lipid II C1a−O4a distances, implied time scale plot for the 3-state Markov model of SAMGT, Free energy surface plot for the 3-state Markov model of SAMGT in function of the two TICA dimensions, Chapman-Kolmogorov test for the 3-state Markov model of SAMGT, QM region of the system for hybrid QM/MM calculations, plot of the activation energy against the sum of the shortest interaction distances between the donor lipid II diphosphate moiety oxygens and hydrogens of Arg_{148}, Lys140, and Lys153, structure and ring nomenclature of moenomycin A, and partial charges calculated for the phosphate moieties of the phospholipids (PDF)

Gaussian input files for the optimized stationary states of the wild-type protein and of the $E_{100}Q$ mutated protein (ZIP).

AUTHOR INFORMATION

Corresponding Author

Hans De Winter − *Department of Pharmaceutical Sciences, Laboratory of Medicinal Chemistry, University of Antwerp, 2610 Wilrijk, Belgium;* orcid.org/0000-0002-4450-7677; Email: Hans.DeWinter@uantwerpen.be

Authors

Kenneth Goossens − *Department of Pharmaceutical Sciences, Laboratory of Medicinal Chemistry, University of Antwerp, 2610 Wilrijk, Belgium;* orcid.org/0000-0003-4127-5724

Rui PP Neves − *LAQV@REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, 4169-007 Porto, Portugal;* orcid.org/0000-0003-2032-9308

Pedro A Fernandes − *LAQV@REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, 4169-007 Porto, Portugal;* orcid.org/0000-0003-2748-4722

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jcim.0c00377

ABBREVIATIONS

GTase   glycosyltransferase
GT51   glycosyltransferase family 51
SAMGT   *Staphylococcus aureus* monoglycosyltransferase
cMD   classical molecular dynamics
QM   quantum mechanics
MM   molecular mechanics
TS   transition state
IRC   intrinsic reaction coordinate

REFERENCES

(1) David, M. Z.; Daum, R. S. Community-associated methicillin-resistant Staphylococcus aureus: epidemiology and clinical consequences of an emerging epidemic. *Clin. Microbiol. Rev.* **2010**, *23*, 616−687.

(2) Taubes, G. The bacteria fight back. *Science* **2008**, *321*, 356−361.

(3) Thati, V.; Shivannavar, C. T.; Gaddad, S. M. Vancomycin resistance among methicillin resistant Staphylococcus aureus isolates from intensive care units of tertiary care hospitals in Hyderabad. *Indian J. Med. Res.* **2011**, *134*, 704−708.

(4) Wilcox, M. H. MRSA new treatments on the horizon: current status. *Injury* **2011**, *42* (Suppl 5), S42−44.

(5) Galley, N. F.; O'Reilly, A. M.; Roper, D. I. Prospects for novel inhibitors of peptidoglycan transglycosylases. *Bioorg. Chem.* **2014**, *55*, 16−26.

(6) Sauvage, E.; Terrak, M. Glycosyltransferases and Trans-peptidases/Penicillin-Binding Proteins: Valuable Targets for New Antibacterials. *Antibiotics* **2016**, *5*, 12.

(7) Lovering, A. L.; Gretes, M.; Strynadka, N. C. Structural details of the glycosyltransferase step of peptidoglycan assembly. *Curr. Opin. Struct. Biol.* **2008**, *18*, 534−543.

(8) Punekar, A. S.; Samsudin, F.; Lloyd, A. J.; Dowson, C. G.; Scott, D. J.; Khalid, S.; Roper, D. I. The role of the jaw subdomain of peptidoglycan glycosyltransferases for lipid II polymerization. *Cell Surf.* **2018**, *2*, 54−66.

(9) Goldman, R. C.; Gange, D. Inhibition of transglycosylation involved in bacterial peptidoglycan synthesis. *Curr. Med. Chem.* **2000**, *7*, 801−820.

(10) Lovering, A. L.; de Castro, L. H.; Lim, D.; Strynadka, N. C. Structural insight into the transglycosylation step of bacterial cell-wall biosynthesis. *Science* **2007**, *315*, 1402−1405.

(11) Sung, M. T.; Lai, Y. T.; Huang, C. Y.; Chou, L. Y.; Shih, H. W.; Cheng, W. C.; Wong, C. H.; Ma, C. Crystal structure of the membrane-bound bifunctional transglycosylase PBP1b from Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 8824−8829.

(12) Yuan, Y.; Fuse, S.; Ostash, B.; Sliz, P.; Kahne, D.; Walker, S. Structural analysis of the contacts anchoring moenomycin to peptidoglycan glycosyltransferases and implications for antibiotic design. *ACS Chem. Biol.* **2008**, *3*, 429−436.

(13) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **2008**, *77*, 521−555.

(14) Terrak, M.; Nguyen-Disteche, M. Kinetic characterization of the monofunctional glycosyltransferase from Staphylococcus aureus. *J. Bacteriol.* **2006**, *188*, 2528−2532.

(15) Heaslet, H.; Shaw, B.; Mistry, A.; Miller, A. A. Characterization of the active site of S. aureus monofunctional glycosyltransferase (Mtg) by site-directed mutation and structural analysis of the protein complexed with moenomycin. *J. Struct. Biol.* **2009**, *167*, 129−135.

(16) Barrett, D.; Wang, T. S.; Yuan, Y.; Zhang, Y.; Kahne, D.; Walker, S. Analysis of glycan polymers produced by peptidoglycan glycosyltransferases. *J. Biol. Chem.* **2007**, *282*, 31964−31971.

(17) Terrak, M.; Ghosh, T. K.; van Heijenoort, J.; Van Beeumen, J.; Lampilas, M.; Aszodi, J.; Ayala, J. A.; Ghuysen, J. M.; Nguyen-Disteche, M. The catalytic, glycosyl transferase and acyl transferase modules of the cell wall peptidoglycan-polymerizing penicillin-binding protein 1b of Escherichia coli. *Mol. Microbiol.* **1999**, *34*, 350−364.

(18) Terrak, M.; Sauvage, E.; Derouaux, A.; Dehareng, D.; Bouhss, A.; Breukink, E.; Jeanjean, S.; Nguyen-Disteche, M. Importance of the conserved residues in the peptidoglycan glycosyltransferase module of the class A penicillin-binding protein 1b of Escherichia coli. *J. Biol. Chem.* **2008**, *283*, 28464−28470.

(19) Schwartz, B.; Markwalder, J. A.; Seitz, S. P.; Wang, Y.; Stein, R. L. A kinetic characterization of the glycosyltransferase activity of Eschericia coli PBP1b and development of a continuous fluorescence assay. *Biochemistry* **2002**, *41*, 12552−12561.

(20) Huang, C. Y.; Shih, H. W.; Lin, L. Y.; Tien, Y. W.; Cheng, T. J.; Cheng, W. C.; Wong, C. H.; Ma, C. Crystal structure of Staphylococcus aureus transglycosylase in complex with a lipid II analog and elucidation of peptidoglycan synthesis mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6496−6501.

(21) Ostash, B.; Walker, S. Moenomycin family antibiotics: chemical synthesis, biosynthesis, and biological activity. *Nat. Prod. Rep.* **2010**, *27*, 1594−1617.

(22) Ferguson, S. A.; Cook, G. M.; Montgomery, M. G.; Leslie, A. G.; Walker, J. E. Regulation of the thermoalkaliphilic F1-ATPase from Caldalkalibacillus thermarum. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 10860−10865.

(23) Ribeiro, A. J.; Ramos, M. J.; Fernandes, P. A. The catalytic mechanism of HIV-1 integrase for DNA 3′-end processing established by QM/MM calculations. *J. Am. Chem. Soc.* **2012**, *134*, 13436−13447.

(24) Pelletier, H.; Sawaya, M. R.; Wolfle, W.; Wilson, S. H.; Kraut, J. A structural basis for metal ion mutagenicity and nucleotide selectivity in human DNA polymerase beta. *Biochemistry* **1996**, *35*, 12762−12777.

(25) Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *J. Mol. Struct.: THEOCHEM* **1999**, *461*, 1−21.

(26) Maseras, F.; Morokuma, K. IMOMM: A new integratedab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states. *J. Comput. Chem.* **1995**, *16*, 1170−1179.

(27) 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)3)2+ H2Oxidative Addition. *J. Phys. Chem.* **1996**, *100*, 19357−19363.

(28) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.

(29) Becke, A. D. A new mixing of Hartree−Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372−1377.

(30) 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: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.

(31) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623−11627.

(32) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200−1211.

(33) Stevens, D. R.; Hammes-Schiffer, S. Exploring the Role of the Third Active Site Metal Ion in DNA Polymerase eta with QM/MM Free Energy Simulations. *J. Am. Chem. Soc.* **2018**, *140*, 8965−8969.

(34) Nakano, S.; Chadalavada, D. M.; Bevilacqua, P. C. General acid-base catalysis in the mechanism of a hepatitis delta virus ribozyme. *Science* **2000**, *287*, 1493−1497.

(35) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245−268.

(36) Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, *129*, 124105.

(37) Ribeiro, A. J. M.; Santos-Martins, D.; Russo, N.; Ramos, M. J.; Fernandes, P. A. Enzymatic Flexibility and Reaction Rate: A QM/MM Study of HIV-1 Protease. *ACS Catal.* **2015**, *5*, 5617−5626.

(38) Strajbl, M.; Sham, Y. Y.; Villà, J.; Chu, Z. T.; Warshel, A. Calculations of Activation Entropies of Chemical Reactions in Solution. *J. Phys. Chem. B* **2000**, *104*, 4578−4584.

(39) Warshel, A. *Computer modeling of chemical reactions in enzymes and solutions*; J. Wiley & Sons, Inc.: New York, 1991; Vol. *96*.

(40) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. How enzymes work: analysis by modern rate theory and computer simulations. *Science* **2004**, *303*, 186−195.

(41) Carugo, O. How large B-factors can be in protein crystal structures. *BMC Bioinf.* **2018**, *19*, 61.

(42) Barrett, D.; Leimkuhler, C.; Chen, L.; Walker, D.; Kahne, D.; Walker, S. Kinetic characterization of the glycosyltransferase module of Staphylococcus aureus PBP2. *J. Bacteriol.* **2005**, *187*, 2215−2217.

(43) Calixto, A. R.; Ramos, M. J.; Fernandes, P. A. Conformational diversity induces nanosecond-timescale chemical disorder in the HIV-1 protease reaction pathway. *Chem. Sci.* **2019**, *10*, 7212−7221.

(44) Hub, J. S.; de Groot, B. L.; Grubmuller, H.; Groenhof, G. Quantifying Artifacts in Ewald Simulations of Inhomogeneous Systems with a Net Charge. *J. Chem. Theory Comput.* **2014**, *10*, 381−390.

(45) Zhan, C.-G.; Dixon, D. A. First-Principles Determination of the Absolute Hydration Free Energy of the Hydroxide Ion†. *J. Phys. Chem. A* **2002**, *106*, 9737−9744.

(46) Huskey, W. P. *Isotope Effects in Chemistry and Biology 2006*; CRC Press/Taylor and Francis Group: Boca Raton, FL, 2006; Vol. *128*.

(47) Braddick, D.; Sandhu, S.; Roper, D. I.; Chappell, M. J.; Bugg, T. D. H. Observation of the time-course for peptidoglycan lipid intermediate II polymerization by Staphylococcus aureus monofunctional transglycosylase. *Microbiology* **2014**, *160*, 1628−1636.

(48) Zhang, Y.; Fechter, E. J.; Wang, T. S.; Barrett, D.; Walker, S.; Kahne, D. E. Synthesis of heptaprenyl-lipid IV to analyze peptidoglycan glycosyltransferases. *J. Am. Chem. Soc.* **2007**, *129*, 3080−3081.

(49) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **2004**, *32*, W665−667.

(50) Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 12−21.

(51) 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.

(52) Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc Bioinf* **2016**, *54*, 5.6.1−5.6.37.

(53) *SPRUCE 1.0.0.2*, OpenEye Scientific Software. http://www.eyesopen.com: Santa Fe, New Mexico, 2019.

(54) White, D. C.; Frerman, F. E. Extraction, Characterization, and Cellular Localization of the Lipids of Staphylococcus aureus. *J. Bacteriol.* **1967**, *94*, 1854−1867.

(55) Kuhn, S.; Slavetinsky, C. J.; Peschel, A. Synthesis and function of phospholipids in Staphylococcus aureus. *Int. J. Med. Microbiol.* **2015**, *305*, 196−202.

(56) White, D. C.; Frerman, F. E. Fatty acid composition of the complex lipids of Staphylococcus aureus during the formation of the membrane-bound electron transport system. *J. Bacteriol.* **1968**, *95*, 2198−2209.

(57) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40*, 843−856.

(58) Goossens, K.; De Winter, H. Molecular Dynamics Simulations of Membrane Proteins: An Overview. *J. Chem. Inf. Model.* **2018**, *58*, 2193−2202.

(59) Piggot, T. J.; Holdbrook, D. A.; Khalid, S. Electroporation of the E. coli and S. Aureus membranes: molecular dynamics simulations of complex bacterial membranes. *J. Phys. Chem. B* **2011**, *115*, 13381−13388.

(60) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

(61) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. *Intermol Forces* **1981**, *14*, 331−342.

(62) 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.

(63) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.

(64) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(65) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182−7190.

(66) Reisser, S.; Poger, D.; Stroet, M.; Mark, A. E. Real Cost of Speed: The Effect of a Time-Saving Multiple-Time-Stepping Algorithm on the Accuracy of Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2017**, *13*, 2367−2372.

(67) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.

(68) Perez-Hernandez, G.; Paul, F.; Giorgino, T.; De Fabritiis, G.; Noe, F. Identification of slow molecular order parameters for Markov model construction. *J. Chem. Phys.* **2013**, *139*, 015102.

(69) Noe, F.; Clementi, C. Kinetic distance and kinetic maps from molecular dynamics simulation. *J. Chem. Theory Comput.* **2015**, *11*, 5002−5011.

(70) Prinz, J. H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J. D.; Schutte, C.; Noe, F. Markov models of molecular kinetics: generation and validation. *J. Chem. Phys.* **2011**, *134*, 174105.

(71) Doerr, S.; Harvey, M. J.; Noe, F.; De Fabritiis, G. HTMD: High-Throughput Molecular Dynamics for Molecular Discovery. *J. Chem. Theory Comput.* **2016**, *12*, 1845−1852.

(72) Sousa, S. F.; Ribeiro, A. J. M.; Neves, R. P. P.; Brás, N. F.; Cerqueira, N. M. F. S. A.; Fernandes, P. A.; Ramos, M. J. Application of quantum mechanics/molecular mechanics methods in the study of enzymatic reaction mechanisms. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2017**, *7*, e1281.

(73) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.

(74) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; SalomonFerrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *AMBER 2018*, University of California: San Francisco, 2018.

(75) 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.

(76) S. Fernandes, H.; Ramos, M. J.; M. F. S. A. Cerqueira, N. molUP: A VMD plugin to handle QM and ONIOM calculations using the gaussian software. *J. Comput. Chem.* **2018**, *39*, 1344−1353.

(77) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.

(78) Dennington, R.; Keith, T.; Millam, J. *GaussView*, version 5, Semichem Inc.: Shawnee Mission, KS, 2009.

(79) Field, M. J.; Bash, P. A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700−733.

(80) Fukui, K. The path of chemical reactions - the IRC approach. *Acc. Chem. Res.* **1981**, *14*, 363−368.

(81) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, non-covalent 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.

(82) Pereira, A. T.; Ribeiro, A. J. M.; Fernandes, P. A.; Ramos, M. J. Benchmarking of density functionals for the kinetics and thermodynamics of the hydrolysis of glycosidic bonds catalyzed by glycosidases. *Int. J. Quantum Chem.* **2017**, *117*, e25409.

(83) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(84) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. Guidelines for the analysis of free energy calculations. *J. Comput.-Aided Mol. Des.* **2015**, *29*, 397−411.

(85) Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J. Chem. Theory Comput.* **2011**, *7*, 4026−4037.

(86) Stroet, M.; Caron, B.; Visscher, K. M.; Geerke, D. P.; Malde, A. K.; Mark, A. E. Automated Topology Builder Version 3.0: Prediction of Solvation Free Enthalpies in Water and Hexane. *J. Chem. Theory Comput.* **2018**, *14*, 5834−5845.

(87) Flyvbjerg, H.; Petersen, H. G. Error estimates on averages of correlated data. *J. Chem. Phys.* **1989**, *91*, 461−466.