

Theoretical Insights into the Reaction and Inhibition Mechanism of Metal-Independent Retaining Glycosyltransferase Responsible for Mycothiol Biosynthesis

Juan I. Blanco Capurro,^{†,‡} Chad W. Hopkins,[§] Gustavo Pierdominici Sottile,^{||}
Mariano C. González Lebrero,^{⊥,⊗} Adrian E. Roitberg,^{*,#} and Marcelo A. Marti^{*,†,‡}

[†]Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Intendente Guiraldes 2160, C1428EGA, Ciudad Autónoma de Buenos Aires, Argentina

[‡]Instituto de Química Biológica de la Facultad de Ciencias Exactas y Naturales (IQUIBICEN), CONICET, Ciudad Universitaria, Intendente Guiraldes 2160, C1428EGA, Ciudad Autónoma de Buenos Aires, Argentina

[§]Department of Physics, University of Florida, Gainesville, Florida 32611, United States

^{||}Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Sáenz Peña 352, Bernal B1876BXD, Argentina

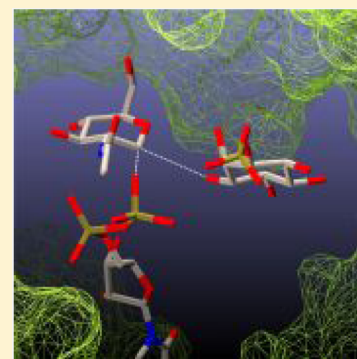
[⊥]Departamento de Química Inorgánica, Análítica y Química Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Intendente Guiraldes 2160, C1428EGA, Ciudad Autónoma de Buenos Aires, Argentina

[⊗]Instituto de Química Inorgánica, Materiales Ambiente y Energía (INQUIMAE), Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Intendente Guiraldes 2160, C1428EGA, Ciudad Autónoma de Buenos Aires, Argentina

[#]Department of Chemistry, University of Florida, Gainesville, Florida 32611, United States

Supporting Information

ABSTRACT: Understanding enzymatic reactions with atomic resolution has proven in recent years to be of tremendous interest for biochemical research, and thus, the use of QM/MM methods for the study of reaction mechanisms is experiencing a continuous growth. Glycosyltransferases (GTs) catalyze the formation of glycosidic bonds, and are important for many biotechnological purposes, including drug targeting. Their reaction product may result with only one of the two possible stereochemical outcomes for the reacting anomeric center, and therefore, they are classified as either inverting or retaining GTs. While the inverting GT reaction mechanism has been widely studied, the retaining GT mechanism has always been controversial and several questions remain open to this day. In this work, we take advantage of our recent GPU implementation of a pure QM(DFT-PBE)/MM approach to explore the reaction and inhibition mechanism of MshA, a key retaining GT responsible for the first step of mycothiol biosynthesis, a low weight thiol compound found in pathogens like *Mycobacterium tuberculosis* that is essential for its survival under oxidative stress conditions. Our results show that the reaction proceeds via a front-side S_Ni -like concerted reaction mechanism (D_{NAN} in IUPAC nomenclature) and has a 17.5 kcal/mol free energy barrier, which is in remarkable agreement with experimental data. Detailed analysis shows that the key reaction step is the diphosphate leaving group dissociation, leading to an oxocarbenium-ion-like transition state. In contrast, fluorinated substrate analogues increase the reaction barrier significantly, rendering the enzyme effectively inactive. Detailed analysis of the electronic structure along the reaction suggests that this particular inhibition mechanism is associated with fluorine's high electronegative nature, which hinders phosphate release and proper stabilization of the transition state.



■ INTRODUCTION

Glycosyltransferases (GTs), EC 2.4.x.y, are ubiquitous enzymes available in all forms of life, that establish natural glycosidic linkages between a carbohydrate and a wide range of molecules, yielding oligo- and polysaccharides, glycoproteins, glycolipids and cell wall components, antibiotics, macrolides, and nucleic acids. Generically, they catalyze the transfer of a carbohydrate moiety from an activated donor (the “glycosyl donor”) to a specific acceptor molecule, and whenever the donor is a diphospho sugar-nucleotide derivative (NDP), they are usually

called “Leoir enzymes”, in honor of the Argentinian Nobel laureate.¹

Structurally, GTs are currently classified into more than 90 CAZy families, and they usually adopt one of two folds.² The GT-A fold consists of two structurally different domains that yield a rather compact, globular structure, and usually binds

Received: October 10, 2016

Revised: December 7, 2016

Published: December 9, 2016

divalent cations in the active site, which play a key role in the catalytic reaction by stabilizing the glycosyl donor and aiding in leaving group departure; the GT-B fold consists of two domains with Rossmann-type folds connected by a hinge zone, and the active site locates right between the middle of these domains. In GTs, the glycosyl-transfer reaction may proceed with one of the two stereochemical outcomes regarding the configuration of the anomeric center, which may be either inverted or retained. Consequently, and given the fact that different enzymes perform their task with a strict stereochemical control, they have been categorized as either “inverting” or “retaining” GTs. Interestingly, the inverting/retaining classification of GTs was shown to be independent of the fold, since both GT-A and GT-B examples of inverting and retaining enzymes have been reported.^{1,2}

Inverting GTs have been traditionally the most studied,^{3–7} since inversion of the anomeric configuration often occurs in a straightforward process involving a concerted S_N2 -like single displacement mechanism (Figure 1A), in which the acceptor

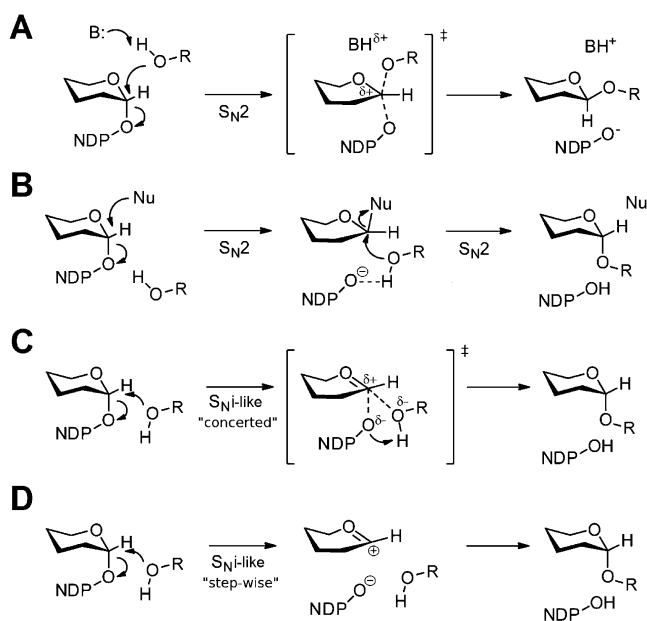


Figure 1. Various reaction mechanisms proposed for GTs: (A) inverting mechanism; (B) double-displacement retaining mechanism; (C) S_Ni -like concerted retaining mechanism (A_ND_N); (D) S_Ni -like stepwise retaining mechanism ($D_N^*A_{Nss}$).

molecule attacks the anomeric center from the opposite face of the leaving group. Also, in addition to stabilization of the transition state, inverting enzyme active sites provide a general base function, typically performed by a carboxylate residue, that activates the acceptor nucleophile.¹ Retaining GTs, on the other hand, have always been controversial. Original models suggested that retention was—by analogy with retaining glycosidases—the result of a double inversion, generated by two sequential S_N2 (Figure 1B): the first S_N2 occurring with an enzyme’s residue acting as a nucleophile that results in the formation of an “inverted covalent-enzyme intermediate”, after which a second S_N2 driven by the acceptor moiety would take place to yield the final product restoring the initial configuration. In other words, early proposals of retaining mechanisms consisted of two sequential inverting steps. Later on, as increasingly more retaining GT structures began to be

reported, the absence of properly located residues that could be responsible for the covalent-enzyme intermediates in many active sites led—as previously proposed for glycogen phosphorylase¹⁰—to the so-called “ S_Ni -like” mechanism model (also referred to in the literature as “front-side” or “single displacement”).^{1,8–11} In this model, the nucleophilic attack and leaving group departure occur in the same face of the donor sugar through an oxocarbenium ion-like transition state (Figure 1C) or intermediate (Figure 1D), and it is the NDP diphosphate group itself that acts as the base that deprotonates the acceptor. Given their ubiquity and relevance, understanding the GT mechanism is an active field of research. Moreover, unveiling GT catalyzed reaction mechanisms with atomic detail is a crucial aspect to understand the microscopic origin of their catalytic efficiency, with a potential impact in protein engineering and drug design projects, since they allow not only identification of key residues involved in the reaction and substrate recognition process but also the rational design of novel inhibitors that, for example, mimic transition state structures.^{12,13} In this context, computational methods offer both a powerful predictive tool and an invaluable complement to experimental data.

Density functional theory (DFT) based hybrid quantum mechanical/molecular mechanical (QM-MM) methods are among the best choices for studying enzymatic reaction mechanisms, since they allow one to get an accurate QM description of the chemical events that take place in the enzyme’s active site (breaking/forming of chemical bonds and charge fluctuation as a function of geometry), while treating the rest of the system and its environment using a less computationally expensive classically derived force field. We have a lot of experience in using these types of methods to describe enzymatic reaction mechanisms,^{14–19} and we have recently implemented a pure DFT method—using the Perdew, Burke, and Ernzerhof (PBE) exchange correlation functional—in the popular AMBER package which runs on powerful graphical processing units (GPU),^{20,21} allowing the user to determine free energy profiles (i.e., including entropic and dynamic contributions) of challenging reactions as those presented by GTs.

So far, the S_Ni -like mechanism has been successfully modeled by several groups in both GT-A and GT-B examples of retaining enzymes using computer simulations with different levels of theory and sampling schemes.^{22–24} However, despite accurately and consistently describing the S_Ni -like mechanism, some cases have been shown to favor a highly dissociative mechanism with an oxocarbenium-ion-like transition state²⁵ (IUPAC A_ND_N , Figure 1C), while others seem to prefer a stepwise mechanism with a short-lived oxocarbenium-ion intermediate^{26–34} (IUPAC $A_N^*D_{Nss}$, where the asterisk stands for “short-lived intermediate” and the “ss” stands for “solvent-separated ion pair”, Figure 1D). Despite that some key aspects of the reaction such as the particular order of bond forming and bond breaking events and the nature of the base that deprotonates the attacking $-OH$ have reached general consensus, the debate of whether the front-face mechanism takes place in a single step or as a stepwise reaction is the main question that shows that our understanding of the S_Ni -like mechanism is still incomplete. Moreover, to our knowledge, there is a lack of studies analyzing how the substrate analogues that act as inhibitors actually prevent the enzyme from performing its task. In this context, and to make our contribution to the field, in the present work, we analyzed

using our in-house developed QM(DFT-PBE)/MM method the reaction and inhibition mechanism of a key retaining metal-independent GT-B from Actinobacteria, MshA.³⁵

Mycothiol (Msh) is a low weight thiol compound found in Actinobacteria—Gram-positive bacteria—a phylum that includes pathogens like *Mycobacterium tuberculosis* and *Mycobacterium leprae*, whose function is analogous to that of Glutathione in Gram-negative bacteria and eukaryotes.³⁶ It plays a very important role in the bacillus pathogenicity, and its absence renders the bacteria highly sensitive to the immune system's nitrosative attack. Among the enzymes involved in mycothiol biosynthesis, glycosyltransferase “MshA” catalyzes the first step of the metabolic pathway, and has proved to have many key features that make it a potential therapeutic target.^{37–39}

MshA (family GT4) is a metal-independent retaining GT-B that catalyzes the transfer of *N*-acetyl-glucosamine (GlcNAc) from donor UDP-GlcNAc to acceptor 1-*L*-myo-inositol-1-phosphate (m-Ins-1P), to yield 3-phospho-1-*D*-myo-inositol-2-acetamido-2-deoxy- α -*D*-glucopyranoside (GlcNAc-Ins-P).^{36,37} On the basis of the available X-ray structure of *Corynebacterium glutamicum* MshA in complex with UDP and m-Ins-1P (PDB id: 3C4V) and steady state kinetic data,³⁵ the following general mechanism was proposed: m-Ins-1P's O3 oxygen performs a nucleophilic attack on GlcNAc's anomeric center (C1) to form the glycosidic bond, and it is concomitantly deprotonated by UDP's β -phosphate group (oxygen O3B) which acts as the leaving group. However, several questions remain unanswered, particularly related to the specific order of events, the potential presence of a stable ionic intermediate, and how catalysis is actually achieved. Furthermore, the available kinetic data with the substrate analogue UDP-(5F)-GlcNAc suggest that it acts as a competitive inhibitor,⁴⁰ but interpretation of the data is not straightforward, particularly in relation to how precisely the presence of the fluorine interferes in the glycosyl-transfer reaction. Thus, given the available data, and its relevance for future drug development projects, we strongly believe MshA is an ideal case for molecular simulation studies of the GT reaction and inhibition mechanisms, and have therefore undertaken its study. Our results show that MshA works through a concerted S_Ni-like mechanism, and that the key effect of the fluorine atom in the inhibitor is to prevent charge transfer that compensates the positive charge development in the anomeric carbon during transition state formation.

COMPUTATIONAL METHODS

Starting Structure and Equilibration Protocol. Coordinates were retrieved from the Protein Data Bank (PDB id: 3C4V) for *Corynebacterium glutamicum* MshA in complex with UDP and m-Ins-1P.³⁵ Since this initial structure presented two empty gaps (residues 139–145; 266–273), they were completed by superimposition of all chain monomers available (including the ones from the apo structure, PDB id: 3C48). Final missing residues were added using the Modeler 9.9 software.⁴¹ The sugar donor, GlcNAc, was added onto this complete model using an in-house developed bias-docking approach, specially designed for docking of carbohydrate molecules⁴² running on the Autodock 4 software.⁴³ The complete ternary complex was then solvated with TIP3P water molecules in an octahedral box. The final solvated system consisted of over 36 000 atoms. All subsequent molecular dynamics simulations were performed with the AMBER ff99SB force field⁴⁴ for the protein and the GLYCAM-06j force field⁴⁵

for GlcNAc. Classical force field parameters for UDP and m-Ins-1P (Supporting Information, Figure S1) were generated with the Antechamber module of the AMBER suite, after geometry optimization and RESP charge derivation at the HF/6-31G* level with Gaussian g09, revision D.01.⁴⁶

The equilibration protocol consisted of 500-cycle runs of minimization with a 100 (kcal/mol)/Å² restraint constant applied to the protein in order to relax the solvent structure, followed by a 2000-cycle energy optimization in which the restraint was removed to lose initial unfavorable contacts. The system was then slowly heated to 300 K during a 200 ps simulation, with the Berendsen thermostat. Finally, pressure was equilibrated at 1 atm over 500 ps, to let the system reach the proper density. For all simulations we used the periodic boundary condition approximation, with the Ewald summation method with a 12 Å cutoff for nonbonded interactions, and the SHAKE algorithm for all hydrogen-containing bonds. Final production 100 ns MD simulations were performed at 300 K using the Langevin thermostat and a 2 fs time step, from which the last structure was selected for QM-MM simulations.

The 5F-GlcNAc inhibitor starting structure was generated by replacing the hydrogen atom “H5” with fluorine on the topology file of the starting QM/MM structure, and carefully equilibrating it to let the moiety and its surroundings relax properly.

QM-MM Umbrella Sampling Simulations. The QM region consisted of 64 atoms (Figure 2), including the whole

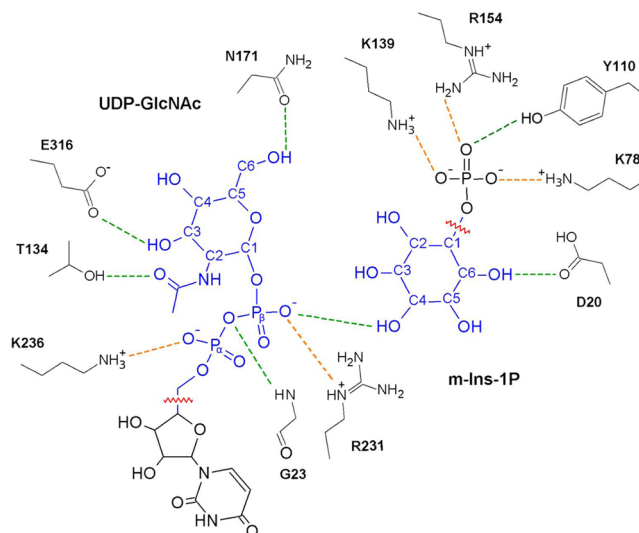


Figure 2. Schematic representation of MshA's active site. Colored in blue are the 64 atoms treated quantum-mechanically, and the red zigzag lines indicate where the two link atoms were placed. The main polar interactions involved in the anchoring of the reactants are represented with dashed lines, green for hydrogen bonds and orange for salt bridges.

GlcNAc donor, the diphosphate moiety of the UDP leaving group, the m-Ins-1P acceptor without the phosphate group, and two H link atoms. It should be noted that, although partitioning the QM and MM subsystems in a polar bond (such as the m-Ins-1P C1–O bond) is usually against general advice, we opted for this in order to avoid including the phosphate group in the QM subsystem. This inclusion would have required the incorporation of further side chain functional groups bearing

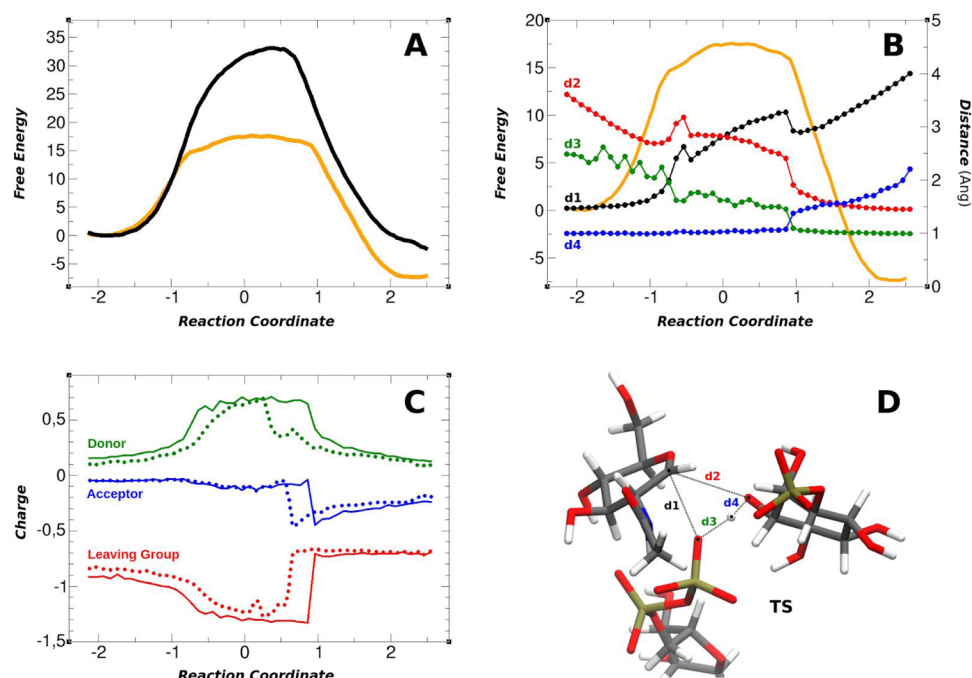


Figure 3. (A) Free energy profiles as a function of the reaction coordinate, at the DFT(PBE)/DZVP level for the natural substrate (orange) and the fluorinated substrate analogue (black). (B) Distances as a function of RC for the main atoms involved in the reaction center. (C) Mulliken populations for the most relevant charged centers for natural substrate (full lines) and for fluorinated substrate analogue (dotted lines). (D) Transition state representative structure.

compensating charges, thus significantly increasing the QM subsystem size and consequently the computational cost.

LIO is an implementation of DFT with Gaussian basis sets at the generalized gradient approximation level, which works with the AMBER code setting up a QM/MM application. All relevant parts of the computation of LIO are ported to GPU, obtaining a great impact on performance, including exchange and correlation, Coulomb, and QM/MM coupling terms.^{20,21} For this particular system under study, a common 1 ps long DFT QM/MM simulation on NVIDIA's Tesla K20X GPUs could take up to 10 h (~35 s per integration step, over a total of 1000 steps). All DFT QM/MM calculations in this work were performed with the Sander AMBER program and the LIO DFT program connected through Sander's external QM/MM interface,⁴⁷ with the PBE functional using the DZVP basis.

For umbrella sampling⁴⁸ calculations, reaction coordinate (RC) was defined as the combination of the two distances that represent, respectively, the bond hydrolysis " $d1 = d(\text{C1}-\text{O3B})$ ", and the bond formation at the anomeric center " $d2 = d(\text{O3}-\text{C1})$ ". The RC was sampled in increments of 0.1 Å intervals, yielding a total of 48 sampling windows, using a 200 (kcal/mol)/Å² force constant. Each window consisted of an initial 0.5 ps equilibration of the system, using always as a starting structure the last snapshot from the previous window. Then, each window simulation was extended for a least a further 6 ps, with the last 4 ps being used as production data. All simulations were carried out with a 1 fs time step and a QM/MM electrostatic cutoff of 16 Å, at 300 K (Langevin thermostat) and 1 atm (Berendsen barostat). The rest of the conditions were kept identical as mentioned in the above section for the classical production MD simulations (a sample Sander input file is provided in the Supporting Information, Figure S2). Free energy profiles (FEP) were calculated employing the WHAM program,⁴⁹ dividing the sampled data

into 100 bins and considering a convergence criterion of 0.0001.

For comparative purposes, we also determined the FEP and reaction mechanism with lower and computationally faster levels of theory for the QM subsystems. Specifically, we tested the following density function tight-binding (DFTB) AMBER implementations:^{50,51} DFTB2 (also known as SCC-DFTB), DFTB2-PA⁵² (which contains on-site third-order terms to improve proton affinities of phosphate compounds), and DFTB2-PR⁵² (a "reaction specific" improvement for phosphate hydrolysis reactions). DFTB2 parameters (Slater–Koster files) were retrieved from the dftb.org Web site. For all DFTB2 calculations, the "mio-1-1" parameter set was used.

RESULTS

Stability of the complete MshA ternary-complex classical MD simulation was carefully inspected both visually and by means of plotting the root mean squared deviation (RMSD) as a function of time for all the protein heavy atoms (Supporting Information, Figure S3). Results showed that the model was stable after a 100 ns run, so we could proceed with QM/MM calculations.

To determine the free energy profile (FEP) and underlying mechanism of MshA, we applied an umbrella sampling strategy using as a reaction coordinate the combination of the bond forming "m-Ins-1P O3–GlcNAc C1" and bond breaking "GlcNAc C1–β-phosphate O3B" distances. This RC allows the system not only to decide the order of events related to the bond breaking and bond formation on C1 but also allows HO3 proton transfer to occur unbiased. The corresponding FEP obtained at the DFT(PBE)-DZVP level is shown in Figure 3A, while key distances and Mulliken population plots are shown in Figure 3B and C, respectively. The profile shows that the reaction is exergonic by ca. 5 kcal/mol and presents a moderate

17.5 kcal/mol barrier. Analysis of the plots shows that, during the first stage of the reaction (RC between -2 and -1), O3 approaches the anomeric center up to ~ 2.7 Å (initial drop in d2). Afterward, the key process leading to the transition state (TS) clearly corresponds to the breaking of the “GlcNAc- β -phosphate” bond (d1), with the concomitant strengthening of the “ β -phosphate O3B-m-Ins-1P HO3” hydrogen bond. Release of the diphosphate as the reaction trigger is also evidenced in the Mulliken analysis, which shows that the diphosphate leaving group becomes negatively charged.

In the TS zone (Figure 3D, RC between 0 and 0.5), the most important geometrical change concerns the GlcNAc ring, which as expected instead of the initial 4C_1 chair shows a conformation that resembles more a 4H_5 half-chair, consistent with the formation of an oxocarbenium ion at C1. Further evidence for the ionic character of the TS comes from the C1 almost planar geometry (the planar angle is $>160^\circ$ in the TS region) and the significant increase in the GlcNAc positive charge. Also important is the fact that in the TS the attacking nucleophile still has the proton attached, although it is partially shared with the diphosphate.

The final stage of the reaction clearly depends on the concerted proton transfer from m-Ins-1P O3 to the β -phosphate (see d3 and d4) and formation of the O3-C1 bond (d2) and, last, complete detachment of the now protonated diphosphate. It is striking to observe the tight coupling between the bond formation and proton transfer, which occurs spontaneously, since—as already mentioned—the proton transfer is not part of the RC and thus no external force acts on it. Closer inspection of the substrate geometry in view of the observed mechanism also reveals two key interactions that help stabilize the building of negative charge on the UDP β -phosphate: first, the donor GlcNAc acetamide group hydrogen bond, and second, the acceptor m-Ins-1P hydroxyl group HO3 hydrogen bond. These interactions can be considered as a form of *substrate assisted catalysis*.^{5,30}

Overall, the results suggest a dissociative reaction, that can be assigned to a single-displacement concerted mechanism (D_NA_N), shown in Figure 1D, with the leaving group detachment as the key step, and a clear oxocarbenium-ion-like character centered on GlcNAc C1. As for TS identification, it can be seen that there is a rather flat and wide energy region corresponding to oxocarbenium-ion-like structures with very similar energies. Altogether, these properties make MshA a very interesting case for understanding how catalytic efficiency is achieved in GTs (see the Discussion).

To analyze the MshA reaction further, and particularly in relation to possible inhibition mechanisms, we decided to study the reaction but using UDP-(SF)-GlcNAc instead as an alternative substrate/inhibitor. The corresponding FEP profile—computed with the same strategy and level of theory—is also shown in Figure 3A. The data shows why UDP-(SF)-GlcNAc apparently acts as a competitive inhibitor. As expected, it binds exactly as the natural substrate, but despite that the transfer reaction is still exergonic, the free energy barrier is over 30 kcal/mol, resulting in an extremely slow reaction rate. Analysis of the underlying mechanistic steps (data not shown) evidences that for the fluorinated analogue mechanistic steps are the same as those previously described for the natural substrate, and therefore, their different behavior must result from subtle differences in how the TS state is reached and stabilized in each case.

Detailed analysis of the FEP and distances and Mulliken population plots show that the main difference between natural and fluorinated substrate concerns the free energy needed to reach the TS region, which requires breaking of the phosphate-sugar bond and developing of the oxocarbenium character in the GlcNAc. The first noticeable difference between both substrates concerns the charge of the leaving diphosphate, which starts more negative in the natural substrate and draws charge earlier as the TS develop. This difference is clearly due to the presence of the fluorine attached to C5 because, while the initial charge in the C5-H(F) is about $+0.2e$ in the natural substrate, it is only $+0.1e$ in the fluorinated analogue. Moreover, when reaching the TS in the fluorinated analogue, the charge of the C5-F is significantly increased. Overall, the positive charge in the TS is similar in magnitude and distribution but is reached earlier in the natural substrate. In other words, and as expected, the fluorine atom increases the free energy barrier, since, due to its electronegativity, more energy is required to withdraw its charge allowing an optimal charge distribution across the whole GlcNAc moiety that promotes the diphosphate leaving group to be easily released.

DISCUSSION

Given their ubiquitous presence and relevance, glycosyltransferase reaction mechanisms have been extensively studied using theoretical and experimental chemistry methods. With the exception of GT-6 family members, for those retaining GTs which lack a putative residue that can act as a catalytic base, a front-side S_{Ni} -like mechanism is generally proposed. This mechanism was first proposed in 1937 to elucidate unusual stereochemical outcomes of alkyl halides.⁸ In the 1980s, it was reintroduced to explain the retention of the anomeric configuration stereochemistry in the solvolysis of α -glucosyl fluorides in mixtures of ethanol and trifluoroethanol,⁹ and was also used to explain the catalytic mechanism of an enzyme (glucogen phosphorylase) for the very first time.¹⁰ The first crystallographic X-ray structure of a retaining glycosyltransferase in complex with both complete donor and acceptor moieties was that of GT-A α -1,4-galactosyltransferase “LgtC” from *N. meningitidis*, reported in 2001 by Persson et al.¹¹ This authors were among the first to suggest that, given the absence of a putative active site base, a S_{Ni} -like mechanism with an “oxocarbenium-ion-like” transition state was the most probable scenario. A year later, Gibson et al. solved the structure of the first retaining GT-B, trehalose-6-phosphate synthase “OtsA” from *E. coli*, for which a similar mechanism was proposed.⁵³ Almost a decade after, separate groups carried for the first time full-enzyme QM/MM calculations on these structures. Despite using different strategies and sampling schemes, both groups were able to reproduce the S_{Ni} -like mechanism but at the same time arrived at different results. The work of Gomez et al. on LgtC²⁵ consisted of both constrained optimizations at the QM(DFT)/MM level to obtain a “potential energy profile” and also umbrella sampling at the SCC-DFTB level to obtain the corresponding “free energy profile”. Both results showed a rather concerted S_{Ni} -like mechanism with an oxocarbenium-ion-like transition state (A_ND_N). The work of Ardevol et al. on OtsA,²⁶ on the other hand, carried out Car-Parrinello DFT/PBE-based metadynamics simulations to construct a “free energy landscape” of the reaction, and their results suggested a stepwise S_{Ni} -like mechanism ($D_N^*A_{Nss}$), in which a short-lived oxocarbenium-phosphate ion pair intermediate was formed. Kinetic isotope labeling and linear free-energy relationship

investigation conducted the same year on OtsA by Lee et al.⁵⁴ gave experimental evidence and further support to the single-displacement S_Ni -like mechanism model, but they were unable to confirm whether the reaction proceeds via a stepwise or a concerted mechanism. After these two pioneer works on full-enzyme models of retaining GTs, many other cases were studied, becoming an active field of research.

Our results, performed for the metal-independent retaining GT-B MshA of *Corynebacterium glutamicum* at the QM(DFT-PBE)/MM level of theory using the umbrella sampling strategy, suggest a concerted and highly dissociative S_Ni -like mechanism (D_NA_N), similar to that observed for LgtC but with a rather flat high-energy region which makes it difficult for a precise TS structure identification. Also important is that in both cases proton transfer occurs late in the reaction. The obtained activation energy barrier of 17.5 kcal/mol is also in remarkable agreement with the experimental data from Vetting et al., who determined a K_{cat} value of 12.5 s^{-1} from steady state kinetics,³⁵ which can be assigned to an ~ 17.02 kcal/mol barrier using transition state theory and the Eyring equation, and in the range of the previously reported value of 23 kcal/mol for the above cited metal-independent GT-B OtsA *in silico* metadynamics study.²⁶

Furthermore, we also determined the FEP and underlying mechanism for the reaction of the fluorinated substrate analogue, previously reported as a competitive inhibitor.⁴⁰ Consistent with the experimental data, we obtained a significantly higher barrier (33 kcal/mol), which would result in a very slow reaction and thus explain the experimentally observed inhibition. Detailed analysis of the fluorinated analogue behavior shows that the most likely reason for the increase in the barrier is the electronegativity of the fluorine atom that impedes charge transfer to the leaving phosphate, which is further consistent with a dissociative concerted mechanism and oxocarbenium-ion-like TS.

An important aspect of this work concerns the level of theory used. Given the high computational cost of QM(DFT-PBE)/MM (or similar functionals, like B3LYP), and that significantly high free energy barriers were reported at the SCC-DFTB level for GTs LgtC (32.7 kcal/mol)²⁵ and EXTL2 (40.5 kcal/mol),³² it is important to assess their gain in accuracy compared to faster simplified methods. Figure 4 shows the FEP for MshA catalyzed reaction obtained with different semiempirical methods, namely, DFTB2, DFTB-PA, and DFTB-PR (see Computational Methods for details). The results for the different DFTB approximations show that all display higher barriers (up to 45 kcal/mol) and that there are very large differences among them. The lowest barrier is obtained with the

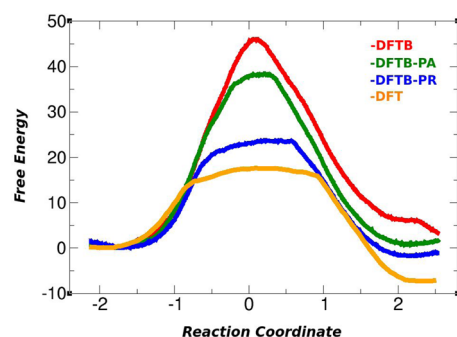


Figure 4. Free energy profiles at various DFTB levels.

DFTB-PR scheme (~ 25 kcal/mol), the one that has been specifically adjusted for phosphate transfer reactions, which again hints this event as the key step for the system under study. Detailed analysis of the underlying mechanism also shows that, in all DFTB versions, proton transfer occurs immediately after TS formation. Although further analysis of the reasons for these observations is beyond the scope of this publication, we suggest possible failure of DFTB methods to stabilize dissociative ion-pair-like intermediates.

Another interesting point concerns the computational method used. It should be noted that classic umbrella sampling MD as well as other free-energy-based approaches like metadynamics (dynamic QM/MM) are conceptually different and computationally more expensive than the static approaches based on geometrical restrained optimizations along a reaction coordinate, in which thermal effects on atomic motion and entropy are neglected and “potential energy” is calculated instead of “free energy”. Having had vast experience with static methods in the past,^{14–17} we are strong supporters of dynamic methods, since we are inclined to believe dynamic motions of proteins are strongly related to their function. In this context, it is interesting to note that, very recently, three different groups all conducted *in silico* studies on the same metal-dependent GT-A, the “polypeptide UDP-GalNAc transferase 2” (or “ppGalNAcT2”). One of them used metadynamics and obtained a stepwise mechanism,²⁷ while the other two used a static approach, with one obtaining a stepwise²⁹ and the other a concerted mechanism but with no identifiable TS due to a rather flat high-energy region.²⁸ Facing such a scenario, and although it is still difficult to give a definitive answer, the fact that both static and dynamic methods have rendered both types of mechanisms for the same GT-A (ppGalNAcT2) and that now two dynamic methods have yielded both mechanisms for two similar GT-B’s (OtsA and MshA) suggests that both mechanisms might be feasible.

As in almost any simulation method, there is a trade-off between accuracy and sampling, given a certain available computational time. In this work, we have decided to increase the method accuracy, using DFT quantum mechanics accelerated by GPUs, which allowed us to do roughly 6 ps per umbrella window. While this might seem like a short time, it is important to realize that, without the use of GPUs, even 6 ps would be completely out of reach. The 6 ps time is in fact applied to the degrees of freedom orthogonal to the reaction coordinate, so it mostly reflects the relaxation of the local environment to changes in electron density during the reaction. We already know that these changes are localized and not large, which means that they are reasonably sampled in this time scale. If the system were to undergo very large conformational changes, then of course 6 ps would not be enough and an alternative would be needed. As a final remark on the subject, although strong supporters of dynamic methods, we are also well aware that simulation times achievable at the QM(DFT)/MM level are still quite limited by the high computational cost they demand and, thus, there is still a lot to contribute to the field in order to obtain shorter times that enable the user to have more sampling of the configurational space.

CONCLUSIONS

In summary, our results show that the *Corynebacterium glutamicum* MshA reaction occurs through a concerted dissociative S_Ni -like mechanism (D_NA_N), with a flat high-energy region of oxocarbenium-ion-like structures representa-

tive of the transition state. Our data also suggest that the unstable dissociative nature of this TS requires the real DFT level of theory to be accurately described, and tight binding methods fail to describe accurately both the underlying molecular events and the resulting free energy along the reaction. Analysis of the fluorinated substrate analogue shows that it acts as an inhibitor due to a significant increase in the reaction activation energy barrier, motivated by the fluorine reluctance to transfer electron density to the leaving group.

■ ASSOCIATED CONTENT

■ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcc.6b10130.

Classical force-field parameter files in AMBER-compatible format for both MshA substrates, UDP and myo-inositol-1-phosphate; a sample Sander input file to run a typical QM/MM umbrella sampling window simulation using the external module LIO with AMBER; and the root mean square deviation of all protein heavy atoms as a function of time (step) for the classical MD simulation of the MshA ternary complex (PDF)

Special Issue Paper

This paper was originally submitted for the “J. Andrew McCammon Festschrift”, published as the August 25, 2016, issue of *J. Phys. Chem. B* (Vol. 120, No. 33).

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: roitberg@ufl.edu.

*E-mail: marti.marcelo@gmail.com.

ORCID

Adrian E. Roitberg: 0000-0003-3963-8784

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was primarily funded by grants CONICET- PIP 112 201101 00850, PICTO-2012-0057, and RAICES-Siembra 2014-2015 awarded to M.A.M. M.A.M. is a member of CONICET. The simulations were conducted with the Blue Waters supercomputer, supported by NSF Grants ACI-0725070 and ACI-1238993, and the Extreme Science and Engineering Discovery Environment (XSEDE), supported by NSF Grant OCI-1053575. The work in Blue Waters was supported by the NSF-PRAC allocation 1440031 to A.E.R. We would also like to thank the High-Performance Computing Center at the University of Florida.

■ REFERENCES

- (1) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu. Rev. Biochem.* **2008**, *77*, 521–555.
- (2) Tvaroška, I. Structural Insights into the Catalytic Mechanism and Transition State of Glycosyltransferases Using ab initio Molecular Modeling. *Trends Glycosci. Glycotechnol.* **2005**, *17*, 177–190.
- (3) Kozmon, S.; Tvaroška, I. Catalytic Mechanism of Glycosyltransferases: Hybrid Quantum Mechanical/Molecular Mechanical Study of the Inverting *N*-Acetylglucosaminyltransferase I. *J. Am. Chem. Soc.* **2006**, *128*, 16921–16927.
- (4) Krupicka, M.; Tvaroška, I. Hybrid Quantum Mechanical/Molecular Mechanical Investigation of the β -1,4-Galactosyltransferase-I Mechanism. *J. Phys. Chem. B* **2009**, *113*, 11314–11319.
- (5) Tvaroška, I.; Kozmon, S.; Wimmerová, M.; Koča, J. Substrate-Assisted Catalytic Mechanism of *O*-GlcNAc Transferase Discovered by Quantum Mechanics/Molecular Mechanics Investigation. *J. Am. Chem. Soc.* **2012**, *134*, 15563–15571.
- (6) Tvaroška, I.; Kozmon, S.; Wimmerová, M.; Koča, A. QM/MM Investigation of the Catalytic Mechanism of Metal-Ion-Independent Core 2 β 1,6-*N*-Acetylglucosaminyltransferase. *Chem. - Eur. J.* **2013**, *19*, 8153–8162.
- (7) Kumari, M.; Kozmon, S.; Kulhánek, P.; Štěpán, J.; Tvaroška, I.; Koča, J. Exploring Reaction Pathways for *O*-GlcNAc Transferase Catalysis. A String Method Study. *J. Phys. Chem. B* **2015**, *119*, 4371–4381.
- (8) Cowdrey, W. A.; Hughes, E. D.; Ingold, C. K.; Masterman, S.; Scott, A. D. Reaction Kinetics and the Walden Inversion. Part VI. Relation of Steric Orientation to Mechanism in Substitutions Involving Halogen Atoms and Simple or Substituted Hydroxyl Groups. *J. Chem. Soc.* **1937**, 1252–1271.
- (9) Sinnott, M. L.; Jencks, W. P. Solvolysis of D-glucopyranosyl Derivatives in Mixtures of Ethanol and 2,2,2-trifluoroethanol. *J. Am. Chem. Soc.* **1980**, *102*, 2026–2032.
- (10) Klein, H. W.; Im, M. J.; Palm, D. Mechanism of the Phosphorylase Reaction. *Eur. J. Biochem.* **1986**, *157*, 107–114.
- (11) Persson, K.; Ly, H. D.; Dieckelmann, M.; Wakarchuk, W. W.; Withers, S. G.; Strynadka, N. C. Crystal Structure of the Retaining Galactosyltransferase LgtC from *Neisseria meningitidis* in Complex with Donor and Acceptor Sugar Analogs. *Nat. Struct. Biol.* **2001**, *8*, 166–175.
- (12) Errey, J. C.; Lee, S. S.; Gibson, R. P.; Fleites, C. M.; Barry, C. S.; Jung, P. M. J.; O'Sullivan, A. C.; Davis, B. J.; Davies, G. J. Mechanistic Insight into Enzymatic Glycosyl Transfer with Retention of Configuration through Analysis of Glycomimetic Inhibitors. *Angew. Chem.* **2010**, *122*, 1256–1259.
- (13) Raab, M.; Kozmon, S.; Tvaroška, I. Potential Transition-State Analogs for Glycosyltransferases. Design and DFT Calculations of Conformational Behavior. *Carbohydr. Res.* **2005**, *340*, 1051–1057.
- (14) Crespo, A.; Scherlis, D. A.; Martí, M. A.; Ordejon, P.; Roitberg, A. E.; Estrin, D. A. A DFT-Based QM-MM Approach Designed for the Treatment of Large Molecular Systems: Application to Chorismate Mutase. *J. Phys. Chem. B* **2003**, *107*, 13728–13736.
- (15) Crespo, A.; Martí, M. A.; Roitberg, A. E.; Amzel, L. M.; Estrin, D. A. The Catalytic Mechanism of Peptidylglycine α -Hydroxylating Monooxygenase Investigated by Computer Simulation. *J. Am. Chem. Soc.* **2006**, *128*, 12817–12828.
- (16) Capece, L.; Lewis-Ballester, A.; Yeh, S. R.; Estrin, D. A.; Martí, M. A. Complete Reaction Mechanism of Indoleamine 2,3-Dioxygenase as Revealed by QM/MM Simulations. *J. Phys. Chem. B* **2012**, *116*, 1401–1413.
- (17) Dumas, V. G.; Defelipe, L. A.; Petruk, A. A.; Turjanski, A. G.; Martí, M. A. QM/MM Study of the C–C Coupling Reaction Mechanism of CYP121, an Essential Cytochrome P450 of *Mycobacterium tuberculosis*. *Proteins: Struct., Funct., Genet.* **2014**, *82*, 1004–1021.
- (18) Ramírez, C. L.; Zeida, A.; Jara, G. E.; Roitberg, A. E.; Martí, M. A. Improving Efficiency in SMD Simulations Through a Hybrid Differential Relaxation Algorithm. *J. Chem. Theory Comput.* **2014**, *10*, 4609–4617.
- (19) Romero, J. M.; Martín, M.; Ramírez, C. L.; Dumas, V. G.; Martí, M. A. Efficient Calculation of Enzyme Reaction Free Energy Profiles Using a Hybrid Differential Relaxation Algorithm: Application to *Mycobacterium* Zinc Hydrolases. *Adv. Protein Chem. Struct. Biol.* **2015**, *100*, 33–65.
- (20) Nitsche, M. A.; Ferreira, M.; Mocskos, E. E.; González Lebrero, M. C. GPU Accelerated Implementation of Density Functional Theory for Hybrid QM/MM Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 959–967.

- (21) Zeida, A.; Babbush, R.; Lebrero, M. C.; Trujillo, M.; Radi, R.; Estrin, D. A. Molecular Basis of the Mechanism of Thiol Oxidation by Hydrogen Peroxide in Aqueous Solution: Challenging the SN2 Paradigm. *Chem. Res. Toxicol.* **2012**, *25*, 741–746.
- (22) Gómez, H.; Mendoza, F.; Lluch, J. M.; Masgrau, L. QM/MM Studies Reveal How Substrate–Substrate and Enzyme–Substrate Interactions Modulate Retaining Glycosyltransferases Catalysis and Mechanism. *Adv. Protein Chem. Struct. Biol.* **2015**, *100*, 225–254.
- (23) Tvaroška, I. Atomistic Insight into the Catalytic Mechanism of Glycosyltransferases by Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods. *Carbohydr. Res.* **2015**, *403*, 38–47.
- (24) Ardèvol, A.; Iglesias-Fernández, J.; Rojas-Cervellera, V.; Rovira, C. The Reaction Mechanism of Retaining Glycosyltransferases. *Biochem. Soc. Trans.* **2016**, *44*, 51–60.
- (25) Gómez, H.; Polyak, I.; Thiel, W.; Lluch, J. M.; Masgrau, L. Retaining Glycosyltransferase Mechanism Studied by QM/MM Methods: Lipopolysaccharyl- α -1,4-galactosyltransferase C Transfers α -Galactose via an Oxocarbenium Ion-like Transition State. *J. Am. Chem. Soc.* **2012**, *134*, 4743–4752.
- (26) Ardèvol, A.; Rovira, C. The Molecular Mechanism of Enzymatic Glycosyl Transfer with Retention of Configuration: Evidence for a Short-Lived Oxocarbenium-Like Species. *Angew. Chem., Int. Ed.* **2011**, *50*, 10897–10901.
- (27) Lira-Navarrete, E.; Iglesias-Fernandez, J.; Zandberg, W. F.; Companon, I.; Kong, Y.; Corzana, F.; Pinto, B. M.; Clausen, H.; Peregrina, J. M.; Vocadlo, D. J.; et al. Substrate-Guided Front-Face Reaction Revealed by Combined Structural Snapshots and metadynamics for the Polypeptide N-acetylgalactosaminyltransferase 2. *Angew. Chem., Int. Ed.* **2014**, *53*, 8206–8210.
- (28) Gómez, H.; Rojas, R.; Patel, D.; Tabak, L. A.; Lluch, J. M.; Masgrau, L. A Computational and Experimental Study of O-glycosylation. Catalysis by Human UDP-GalNAc polypeptide: GalNAc transferase-T2. *Org. Biomol. Chem.* **2014**, *12*, 2645–2655.
- (29) Trnka, T.; Kozmon, S.; Tvaroška, I.; Koča, J. Stepwise Catalytic Mechanism via Short-Lived Intermediate Inferred from Combined QM/MM MERP and PES Calculations on Retaining Glycosyltransferase ppGalNAcT2. *PLoS Comput. Biol.* **2015**, *11* (4), e1004061.
- (30) Gómez, H.; Lluch, J. M.; Masgrau, L. Substrate-Assisted and Nucleophilically Assisted Catalysis in Bovine α 1,3-Galactosyltransferase. Mechanistic Implications for Retaining Glycosyltransferases. *J. Am. Chem. Soc.* **2013**, *135*, 7053–7063.
- (31) Albesa-Jové, D.; Mendoza, F.; Rodrigo-Unzueta, A.; Gomollón-Bel, F.; Cifuentes, J. O.; Urresti, S.; Comino, N.; Gómez, H.; Romero-García, J.; Lluch, J. M.; et al. A Native Ternary Complex Trapped in a Crystal Reveals the Catalytic Mechanism of a Retaining Glycosyltransferase. *Angew. Chem., Int. Ed.* **2015**, *54*, 9898–9902.
- (32) Mendoza, F.; Gómez, H.; Lluch, J. M.; Masgrau, L. α 1,4-N-Acetylhexosaminyltransferase EXTL2: The Missing Link for Understanding Glycosidic Bond Biosynthesis with Retention of Configuration. *ACS Catal.* **2016**, *6*, 2577–2589.
- (33) Bobovská, A.; Tvaroška, I.; Kőňa, J. A Theoretical Study on the Catalytic Mechanism of The Retaining α -1,2-mannosyltransferase Kre2p/Mnt1p: the Impact of Different Metal Ions on Catalysis. *Org. Biomol. Chem.* **2014**, *12*, 4201–4210.
- (34) Bobovská, A.; Tvaroška, I.; Kőňa, J. Theoretical Study of Enzymatic Catalysis Explains Why The Trapped Covalent Intermediate in the E303C Mutant of Glycosyltransferase GTB was not Detected in the Wild-Type Enzyme. *Glycobiology* **2015**, *25*, 3–7.
- (35) Vetting, M. W.; Frantom, P. A.; Blanchard, J. S. Structural and Enzymatic Analysis of MshA From *Corynebacterium glutamicum*: substrate-assisted Catalysis. *J. Biol. Chem.* **2008**, *283*, 15834–15844.
- (36) Newton, G. L.; Buchmeier, N.; Fahey, R. C. Biosynthesis and Functions of Mycothiol, the Unique protective thiol of Actinobacteria. *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 471–494.
- (37) Fan, F.; Vetting, M. W.; Frantom, P. A.; Blanchard, J. S. Structures and Mechanisms of the Mycothiol Biosynthetic Enzymes. *Curr. Opin. Chem. Biol.* **2009**, *13*, 451–459.
- (38) Newton, G. L.; Unson, M. D.; Anderberg, S. J.; Aguilera, J. A.; Oh, N. N.; delCardayre, S. B.; Av-Gay, Y.; Fahey, R. C. Characterization of *Mycobacterium smegmatis* Mutants Defective in 1-d-myoinositol-2-amino-2-deoxy- α -d-glucopyranoside and Mycothiol Biosynthesis. *Biochem. Biophys. Res. Commun.* **1999**, *255*, 239–244.
- (39) Defelipe, L. A.; Do Porto, D. F.; Pereira Ramos, P. I.; Nicolás, M. F.; Sosa, E.; Radusky, L.; Lanzarotti, E.; Turjanski, A. G.; Marti, M. A. A Whole Genome Bioinformatic Approach to Determine Potential Latent Phase Specific Targets in *Mycobacterium tuberculosis*. *Tuberculosis (Oxford, U. K.)* **2016**, *97*, 181–192.
- (40) Frantom, P. A.; Coward, J. K.; Blanchard, J. S. UDP-(5F)-GlcNAc Acts as a Slow-Binding Inhibitor of MshA, a Retaining Glycosyltransferase. *J. Am. Chem. Soc.* **2010**, *132*, 6626–6627.
- (41) Eswar, N.; Eramian, D.; Webb, B.; Shen, M. Y.; Sali, A. Protein Structure Modeling with Modeller. *Methods Mol. Biol.* **2008**, *426*, 145–159.
- (42) Gauto, D. F.; Petruk, A. A.; Modenutti, C. P.; Blanco, J. I.; Di Lella, S.; Martí, M. A. Solvent Structure Improves Docking Prediction in Lectin-Carbohydrate Complexes. *Glycobiology* **2013**, *23*, 241–258.
- (43) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.
- (44) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ffl4SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (45) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: A Generalizable Biomolecular Force Field. *Carbohydrates. J. Comput. Chem.* **2008**, *29*, 622–655.
- (46) 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.; et al. *Gaussian 09*, revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.
- (47) Götz, A. W.; Clark, M. A.; Walker, R. C. An Extensible Interface for QM/MM Molecular Dynamics Simulations with Amber. *J. Comput. Chem.* **2014**, *35*, 95–108.
- (48) Valleau, J. P.; Torrie, G. M. In *Modern Theoretical Chemistry*; Berne, B. J., Ed.; Plenum Press: New York, 1977; Vol. 5, Statistical Mechanics, Part A.
- (49) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
- (50) Walker, R. C.; Crowley, M. F.; Case, D. A. The Implementation of a Fast and Accurate QM/MM Potential Method in Amber. *J. Comput. Chem.* **2008**, *29*, 1019–1031.
- (51) de M. Seabra, G.; Walker, R. C.; Elstner, M.; Case, D. A.; Roitberg, A. E. Implementation of the SCC-DFTB Method for Hybrid QM/MM Simulations within the Amber Molecular Dynamics Package. *J. Phys. Chem. A* **2007**, *111*, 5655–5664.
- (52) Yang, Y.; Yu, H.; York, D.; Elstner, M.; Cui, Q. Description of Phosphate Hydrolysis Reactions with the Self-Consistent-Charge Density-Functional-Tight-Binding (SCC-DFTB) Theory. 1. Parameterization. *J. Chem. Theory Comput.* **2008**, *4*, 2067–2084.
- (53) Gibson, R. P.; Turkenburg, J. P.; Charnock, S. J.; Lloyd, R.; Davies, G. J. Insights Into Trehalose Synthesis Provided by the Structure of the Retaining Glucosyltransferase OtsA. *Chem. Biol.* **2002**, *9*, 1337–1346.
- (54) Lee, S. S.; Hong, S. Y.; Errey, J. C.; Izumi, A.; Davies, G. J.; Davis, B. G. Mechanistic Evidence for a Front-Side, S_Ni-Type Reaction in a Retaining Glycosyltransferase. *Nat. Chem. Biol.* **2011**, *7*, 631–637.